DOCUMENT RESUME

ED 293 242 EC 202 091

TITLE	Early Intervention Research Institute. Final Report.
INSTITUTION	Utah State Univ., Logan. Early Intervention Research Inst.
SPONS AGENCY	Department of Education, Washington, DC.
PUB DATE	1 Oct 87
NOTE	353p.; Prepared by the Developmental Center for Handicapped Persons.
PUB TYPE	Reports - General (140)
EDRS PRICE	MF01/PC15 Plus Postage.
DESCRIPTORS	Comparative Analysis; *Cost Effectiveness; Developmental Disabilities; *Disabilities; Downs Syndrome; Early Childhood Education; Educational Research; Hearing Impairments; Infants; *Intervention; Language Handicaps; Paraprofessional School Personnel; Parent Participation; Preschool Education; *Program Effectiveness; Program Evaluation; Special Health Problems; Synthesis; Young Children
IDENTIFIERS	*Early Intervention; Single Subject Research Design

ABSTRACT

An overview is provided of a 5-year project conducted by the Early Intervention Research Institute (EIRI) which focused on the effects and costs of early intervention for young handicapped children. The bulk of the document consists of research reports of 14 investigations. (1) meta-analysis of the early intervention efficacy literature; (2) cost-effectiveness of half-day and full-day preschool special education programs; (3) early intervention with hearing impaired children; (4) cost-effectiveness of paraprofessionals versus professionals in early intervention; (5) cost-benefit analysis of the Abecedarian program; (6) quantitative synthesis of single-subject research; (7) cognitive intervention with developmentally delayed preschoolers; (8) effects of differing levels of parent involvement on preschoolers' development and behavior problems; (9) intervention for infants with intraventricular hemorrhage; (10) cost-effectiveness of parent and clinic early intervention for children with language handicaps; (11) analysis of special education early childhood projects approved by the Joint Dissemination Review Panel; (12) economics of early intervention; (13) follow-up study on intraventricular hemorrhage; (14) Down Syndrome parent involvement. Activities of the EIRI in improving the quality of efficacy research, influencing early childhood legislation, and making practitioners and researchers more aware of the importance of collecting both costs and effects data in the early intervention field are also discussed. (JW)

TABLE OF CONTENTS

ABSTRACT

The Early Intervention Research Institute was funded for a five-year period (1982-87) to investigate the efficacy and cost-benefits of early intervention for handicapped preschoolers.

During the first year of the workscope, the institute had as its major priority, conducting a comprehensive integration of the research literature using recently-developed meta-analysis techniques (Glass, 1976); and developing and applying a model for conducting high-quality cost-effectiveness analyses of early intervention. Based on the results of the first year, additional primary research studies (including longitudinal studies) were conducted focusing on those handicapped populations for which few or no cost or efficacy data were available. Specific target populations (including ages, subcategories, and severity of handicap) for the Years 2-5 research thrusts were identified based on the results of the Year 1 meta-analysis.

During its five-year workscope, the institute conducted a total of 14 investigations aimed at expanding the efficacy knowledge base.

The major accomplishments of the institute were:

1. Influencing the early intervention field to improve the quality of efficacy research.

2. Impacting on legislation (P.L. 99-457), which will increase the number of infants and young children served and improve the quality of present services.

3. Developing and disseminating cost protocols designed to improve the collection of cost data.

4. Developing and disseminating a new methodology for integrating single subject intervention studies.

During the five-year period, the institute provided training for 32 graduate students, disseminated 3,054 institute products, and answered 1,331 requests for information.

INTRODUCTION

According to the White House Conference on Handicapped Individuals (1977), "early intervention" is any program designed to accomplish one of three primary purposes:

1. Prevent further progression of a disability or handicapping condition.
2. Produce actual improvement in a handicapping condition.
3. Introduce helping procedures in situations where the handicapping condition is already established.

During the last 20 years, millions of dollars have been spent on the development, implementation, and evaluation of early intervention programs. A major contributor to the amount of resources devoted to early intervention programs for the handicapped has been the Handicapped Children's Early Education Program (HCEEP). Originally funded by Congress in 1968, HCEEP began with 24 demonstration programs in 1969. Since that time, literally hundreds of demonstration projects have been developed, and many of these projects have been widely replicated. For example, the Portage Project, one of the earliest HCEEP demonstration projects, has now been officially replicated in over 170 sites in the United States, and its materials have been used by hundreds of other programs. In the past 12 years, HCEEP demonstration projects have served nearly all categories of handicapping conditions including: trainable mentally retarded, educable mentally retarded, specific learning disabilities, deaf-blind, deaf/hard of hearing, visually handicapped, seriously emotionally disturbed, speech impaired, other health impaired, orthopedically impaired, at risk, developmentally delayed, and multihandicapped. The ages of the children served have ranged from birth to 8 years, with over 50% being 3 years and under, and nearly 90% being 5 years and under. Many states have mandated preschool programs for handicapped children, and substantial numbers of early intervention programs exist in most other states as well.

Resources for early intervention are not limited to the federal government. Deweerd (1981) noted that for each child who received early intervention services from federally-funded demonstration projects, four children were being served by other agencies through a combination of service programs and replication models.

The massive resources being devoted to early intervention have resulted in many calls for evaluation of the worth of such programs As Black and Hutinger (1981) noted:

> What benefits will it provide? What will it cost? Can we afford it? Whenever a school board, or a superintendent, or the head of some other public or private agency is considering whether to support a program aimed at a needed community service, these questions are basic. The effects of the austere economic climate of the 1980s has increasing impact on all aspects of public education and human services. Early education for the handicapped...is no exception. (p. 1)

Evaluation (i.e., efforts to determine the worth or value) of early intervention programs have taken many different forms--both large and small. For example, in 1975, a third party evaluation contract was awarded to Battelle Institute of Columbus, Ohio, to evaluate the impact of HCEEP funded demonstration programs. One hundred twenty-nine randomly selected children in 29 projects were tested in areas including: social, motor, cognitive, and communicative skills. Based on these data, the evaluators concluded that across all categories of handicapping conditions, children made 1-1/2 to 2 times greater gains than they would have been expected to make without the benefit of the project. Additionally, 97% of those parents interviewed perceived positive changes or improvements which they attributed to the project (Stock et al., 1976). Deweerd (1981) concluded that another indicator of the worth of HCEEP funded demonstration projects was the fact that in 1979, 85% of the initial demonstration projects had secured funds to continue their programs and that the level of funding had increased. Literally hundreds of other research studies for both HCEEP funded and other early intervention programs have collected data to determine the "worth" of such programs.

Unfortunately, the results and conclusions from such evaluations have been disturbingly discrepant. For example, there is growing agreement among practitioners that early intervention promises significant resolution or amelioration of some of the most persistent and expensive problems which educators face. According to Jordan, Hayden, Karnes, and Wood (1979):

> Programs providing early educational and therapeutic programming to meet the needs of young handicapped children and their families are reducing the number of children who will need intensive or long-term help. The importance of reaching handicapped children early and working to help them reach their full potential cannot be overemphasized. With early help, the sooner the better, these children can often function at higher levels than has been dreamed possible in prior years. (p. 26)

However, the promise and benefits of early intervention have not been universally accepted. As Hodges and Sheenan (1978) pointed out, "no consistent picture of success emerged from the early childhood education efforts of the 1960s. Although modest or robust immediate gains from structured programs were frequent, just as frequently, these gains eroded after the children left the experimental programs" (p. 4). Gottfried (1973) concluded that:

> Gains in cognitive and intellectual functioning attributable to preschool training were found by some projects but not others at the time of school entrance. However, there were no reports of substantial persistent gains beyond the third grade. Those studies which conducted school-age follow-up studies uniformly reported disappointing long-term results. (p. 286)

Even though the results of research should guide policy and practice, research on early intervention, when considered as a total body of evidence, has been somewhat confusing. Some researchers have reported success; others, failure. Some have suggested that early intervention is effective, but only for specific subgroups of children. Thus, even though the concept of early intervention has been heartily endorsed by individual practitioners and state and federal funding agencies (Swan, 1980), the research evidence is not at all clear. Even more important, the factors which account for the variation in research results have not been identified.

As primary research articles investigating the effectiveness of early intervention have accumulated, practitioners and policy makers have increasingly called for an effective integration of the knowledge which is being produced. In theory, the results of both basic and applied research on a given topic, such as early intervention, should culminate in increased knowledge and improved practice. In reality, however, the very important step of integrating the findings of the completed research on the effectiveness of early intervention into conclusions which affect practice and influence policy has not occurred.

The problems in making sense of the research on the worth of early intervention programs stems from problems in two areas. To determine worth, one must simultaneously consider both benefits or effects of the program and the costs of the program. Stated differently, how much worth something has depends on "what you get" for "what you pay." The major problems with past research in these two areas--effects and costs--are summarized below.

Problems Determining the "Effects" of Early Intervention

Problems in determining the effects of early intervention fall into three main categories: (a) techniques for summarizing apparently discrepant results, (b) the narrowness of the measures considered, and (c) inadequate consideration of long-term effects from early intervention.

Techniques for summarizing apparently discrepant results. In recent years, more and more researchers have realized that commonly-used techniques for summarizing the results of completed research were inadequate (Glass, 1976; Jackson, 1980; Light & Smith, 1971). As Glass (1976) pointed out:

> We need more scholarly effort concentrated on the problem of finding the knowledge that lies untapped in completed research studies. We are too heavily invested in pedestrian reviewing where verbal synopses of studies are strung out in dizzying lists. The best minds are needed to integrate the staggering number of individual studies. This endeavor deserves higher priority now than adding a new experiment or survey to the pile. (p. 4)

The typical approach among social scientists to reviewing and integrating the literature on a given topic follows one of two routes. In both approaches, a group of easily accessible articles from fairly prominent journals or other publications are listed. In the first approach, the reviewer offers a verbal synopsis of the 20 to 40 research articles, and often concludes that the existing research is inconclusive--sometimes researchers find one thing, sometimes, another. A call is then made for additional research using better techniques and more precise methodology so that the truth of the matter can be discovered. In the second approach, the reviewer begins with a similar group of articles, but eliminates all but a few because of supposed design or analysis flaws. The findings of the remaining "acceptable" studies (frequently the work of the reviewer or his/her colleagues) are presented as the truth of the matter.

Both approaches to integrating and understanding previously completed research in the social sciences have serious inadequacies. Almost always, the articles selected for the review are only a small, non-representative fraction of the total research on the particular topic, and thus ignore a significant body of information. In addition, the "definitive" study almost never exists. Obviously better design and analysis procedures are desirable, but it is not at all unusual for a series of well-designed studies on the same topic in the social sciences to yield conflicting results.

The narrowness of measures considered. IQ is the most widely used measure in evaluating special education programs--including early intervention programs (Zigler & Trickett, 1978). This emphasis on IQ is understandable because it is related to so many theoretically and practically significant behaviors (Kohlberg & Zigler, 1967; Mischel, 1968). However, when a measure of IQ is used as the sole or even the primary measure of program effect, serious problems can occur. As Zigler and Balla (1982) pointed out, there are several other equally important child variables. The most important of these variables include motivation and personality factors (Ramey & Campbell, 1979; Sameroff, 1975; Zigler, 1966; 1969), physical health and well being (North, 1979; Zigler & Trickett, 1978), and academic achievement. Zigler and Balla (1982) also emphasized the importance of considering family variables such as the effect of early intervention on siblings (Gray, 1977; Klaus & Gray, 1968), and mother's verbal behavior, self-confidence, and child-rearing practices; and societal variables such as drop-out rates, prevalence of delinquency, and adult marital status. Failure to consider variables such as these may miss some of the most important "effects" of early intervention.

Inadequate consideration of long-term effects. Longitudinal studies to determine the effect of a program are critical in considering programs for handicapped individuals. Because developmental progress and changes in behavior are slower to occur for such individuals, results of a program may not become apparent for an extended time. Variables that are not measured in the short run (such as motivation) may stimulate subsequent development in the child. Such changes mav lead to "sleeper" effects that only become apparent after the passage of time (Seitz, Apfel, & Rosenbaum, 1981; Sheehan & Keogh, 1982).

Problems Determining the "Costs" of Early Intervention

Just as "effect" determination has frequently been inadequately done in research on early intervention, so also the consideration of costs--the other side of the coin--has suffered from its own problems. The most serious problem has been the lack of any consideration of cost data in studies which purport to assess the worth of early intervention programs. Ignoring cost data completely (which happens in most of the research) implicitly assumes that resources are unlimited--a position clearly at odds with reality. But even those studies that have considered costs have always suffered from one or both of the following problems (Bedger, 1974; Frakes, 1981; Frohreich, 1973; Kakalik, Furry, Thomas, & Garney, 1981).

Failure to consider all sources of cost. Most studies which have considered cost of special education programs have only used the project budget as their source of cost data (see for example Black & Hutinger, 1981). Such an approach is overly simplistic and misleading. Instead of simple project budgets, cost should be defined in economic terms as being the value of the resource that would be available for alternative use if a service were not provided (Conley, 1973; Levin, 1981). Using this definition, the value of such variables as contributed time (e.g., volunteer aides, parents working with children at home, teacher overtime) and shared resources (e.g., buildings, utilities, transportation) would need to be considered. Failure to consider all cost data can result in false conclusions.

Failure to consider effects in relation to costs. Almost all cost analyses in special education have failed to go beyond an estimate of "per child cost." What is needed in cost analyses is a simultaneous consideration of both costs and effects. The most effective program may not be the most cost-effective program. For example, consider an actual study of computer-assisted instruction (CAI) that concluded that seven minutes a day of drill and practice on a computer terminal provided gains in arithmetic scores equivalent to 25 minutes a day of teacher instruction (Suppes & Morningstar, 1969a; 1969b, as cited in Levin, 1981). Unfortunately, the study failed to point out that such a CAI program would cost four to five times as much per student as it would cost to use traditional classroom instruction. In fact, as Levin (1975, pp. 90-91) pointed out, seven minutes a day of CAI would have required an additional 25% of the school's instructional budget, while enough additional classroom drill and practice to achieve similar gains would only have required an additional 6% of the budget.

Determining the Worth of Early Intervention

The problems cited above in determining the worth, or in other words, "the efficacy and cost-benefit," of early intervention have been pervasive, but they are not intractable. Techniques do exist for (a) the effective summarization of existing research, (b) conducting comprehensive cost-effectiveness analyses, and (c) conducting primary research which considers effects on a broad range of variables over an extended time.

During its five-year workscope, EIRI developed and implemented a plan for conducting long-range programmatic research activities focused on determining the efficacy and cost-benefits of early intervention. These research activities focused particularly on the optimal duration and intensity of educational services for children and families having significantly different characteristics. During the first year, the major efforts of EIRI focused on conducting a comprehensive integration of the research literature using recently-developed meta-analysis techniques (Glass, 1976); and developing and applying a model for conducting high-quality cost-effectiveness analyses of early intervention. Based on the results of the first year, additional primary research studies (including longitudinal studies) were conducted focusing on those handicapped populations for which few or no cost or efficacy data were available. Specific target populations (including ages, subcategories, and severity of handicap) for the Years 2-5 research thrusts were identified based on the results of the Year 1 meta-analysis.

INSTITUTE MISSION, GOALS, AND OBJECTIVES

The priority area selected for the Early Intervention Research Institute (EIRI) was that of "efficacy and cost benefits of early intervention." Consistent with the following description of this priority area given in RFP 82-040, the overall mission of EIRI was to expand the knowledge base and improve our understanding of the efficacy and cost-effectiveness of early intervention for handicapped children.

> The purpose of this research program is to investigate the effectiveness and associated costs of early education and related services for infants and children with different kinds and severities of handicapping conditions. Research should also address the optimal duration and intensity of educational services for children and families having significantly different characteristics. A research program in this priority area should include the collection of original (new) research data and the analysis of research data already reported in the professional literature. Further, new data collection should be aimed at handicapped populations for which few or no cost or efficacy data are available. (Department of Education, RFP 82-040)

In carrying out this mission, EIRI had the following goals:

1. Integrate the findings and conclusions from previously conducted research on early intervention to determine what is known, what gaps exist, and where future research should focus. Update this review annually and integrate the findings from this update with the institute's own ongoing work.

2. Conduct an integrated program of early intervention research (including longitudinal research) focused on the most important problems and issues encountered in delivering early intervention in typical service settings.

3. Disseminate information about the institute's findings and products to a broad audience of professionals and families concerned with early intervention for the handicapped.

4. Train graduate students and research assistants in research techniques and effective methods of intervention applicable to preschool handicapped populations.

5. Formally evaluate the impact of the institute's findings and products on the field of early intervention.

6. Solicit input, criticism, and feedback from a broad constituency (Advisory Committee members and others) to ensure that the institute's direction and procedures are appropriately focused and being carried out in such a way as to result in the broadest possible impact of institute findings and accomplishments.

During the first year of the institute, the goals listed above were addressed through a series of three related research projects and a variety of other activities. The findings and results from these research projects provided a foundation for the research projects in future years. During the project's five years, 14 major investigations were conducted.

1. A meta-analysis of the early intervention efficacy literature.
2. A cost-effectiveness analysis of half- and full-day programs.
3. A longitudinal study of early intervention with hearing impaired children.
4. A cost-effectiveness analysis of paraprofessional vs. professional in early intervention for the handicapped.
5. A longitudinal cost-benefit analysis of the abecedarian project.
6. Synthesis of single subject research (development of a methodology).
7. Improving pre-academic skills in developmentally delayed preschoolers through the use of a highly structured cognitive intervention program.
8. The effects of three levels of parental involvement on preschoolers with developmental delays and behavior problems.
9. A prospective study of infants with intraventricular hemorrhage.
10. Parent and clinic intervention for children with language handicaps: A cost-effectiveness analysis.
11. An Analysis of Special Education Early Childhood Projects Approved by the Joint Dissemination Review Panel.
12. The economics of early intervention: What do we really know?
13. Intraventricular hemorrhage follow-up study.
14. Down syndrome parent involvement study.

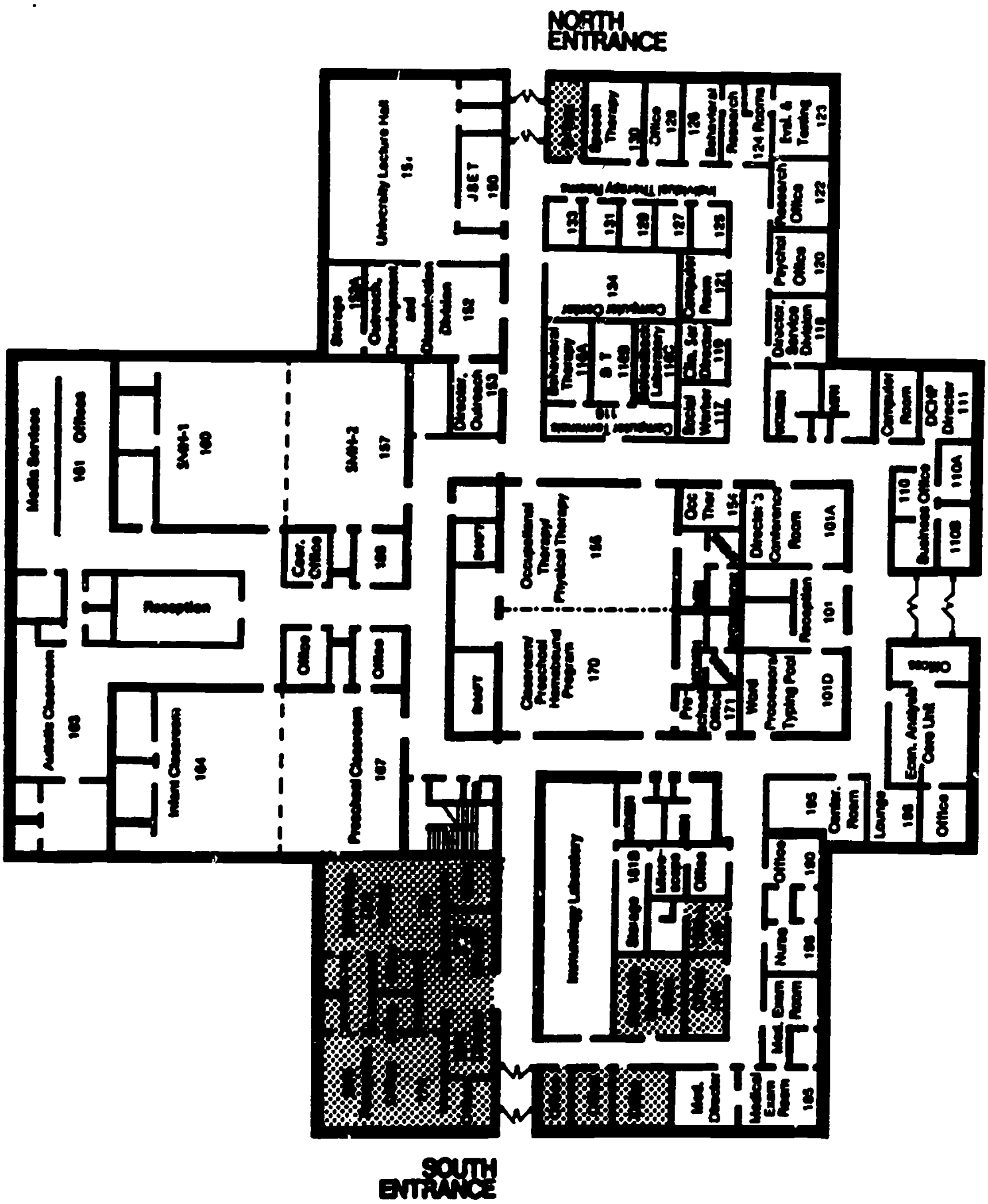
Detailed study descriptions follow on pages 16 to 288.

ADMINISTRATIVE ORGANIZATION

The Early Intervention Research Institute was administered through the Developmental Center for Handicapped Persons (DCHP) facility as the major research unit within the DCHP. The institute operated under the general organizational structure of the DCHP.

The DCHP facility was constructed in 1972 utilizing funding from PL 88-164, state, and local money. The DCHP is located on the campus of Utah State University and has as its major objectives: (a) interdisciplinary training of personnel who will work with the developmentally disabled, (b) demonstration of exemplary services for the handicapped, (c) research and evaluation activities designed to better understand and prevent handicapping conditions, and (d) dissemination of research findings and "best practices" with regard to the handicapped. The DCHP has an interdisciplinary staff of approximately 200 individuals from the fields of special education, medicine, cytogenetics, cellular biology, psychology, child development, communicative disorders, physical therapy, occupational therapy, sociology, research methodology, and instructional technology. The DCHP provided space and facilities which were initially designed for the purpose of conducting research in keeping with this proposal. This space was made available to house the Early Intervention Research Institute as shown in the floor plan on the following page.

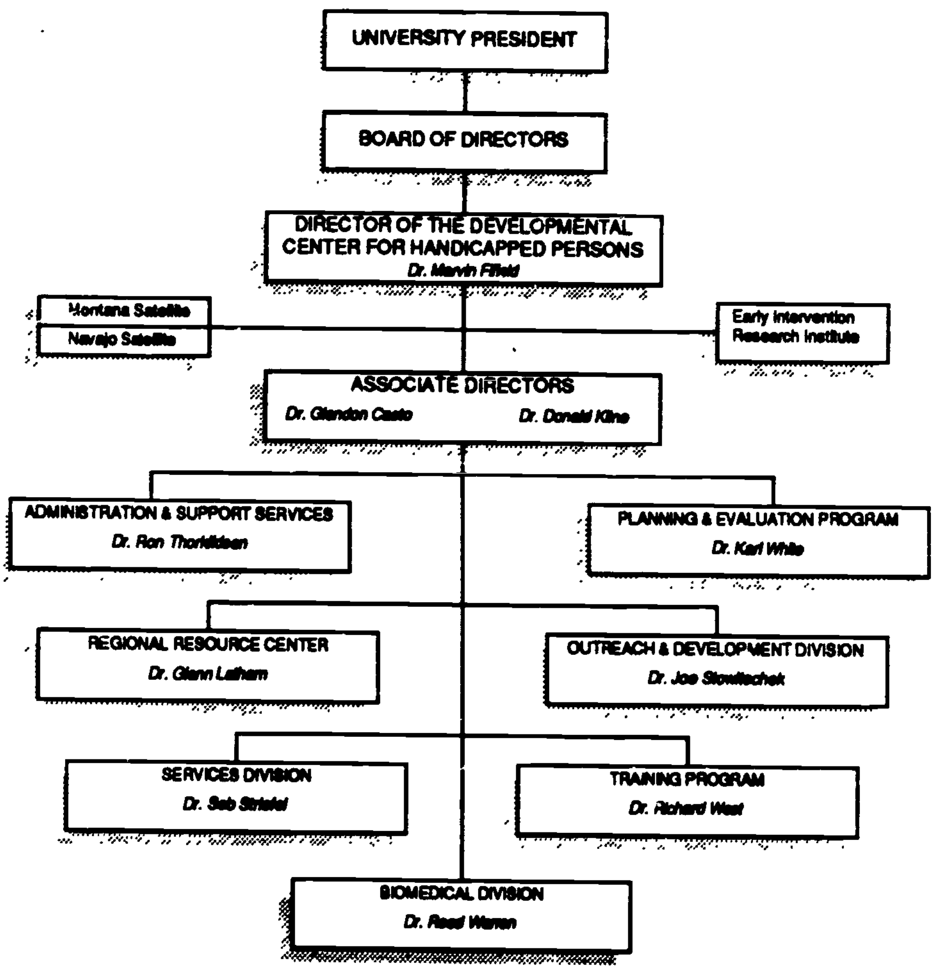
Support and commitment to the Developmental Center for Handicapped Persons and its various components from Utah State University Administration has been extensive and continuous since 1970. The President of the University was personally instrumental in securing matching state and local funds when the center was constructed. Budget requests for increased state support for the program have always been supported by the University Administration. The Office of the Vice President for Research provided funding for the initial acquisition of equipment and program development money for the Biomedical Laboratory and the Early Childhood Research Program. The DCHP is one of three major research, training, and service units on the



NORTH ENTRANCE
University Lecture Hall
Media Services
Offices
Reception
Autistic Classroom
Infant Classroom
Preschool Classroom
Occupational Therapy/ Physical Therapy
Immunology Laboratory
Business Office
Typing Pool
SOUTH ENTRANCE

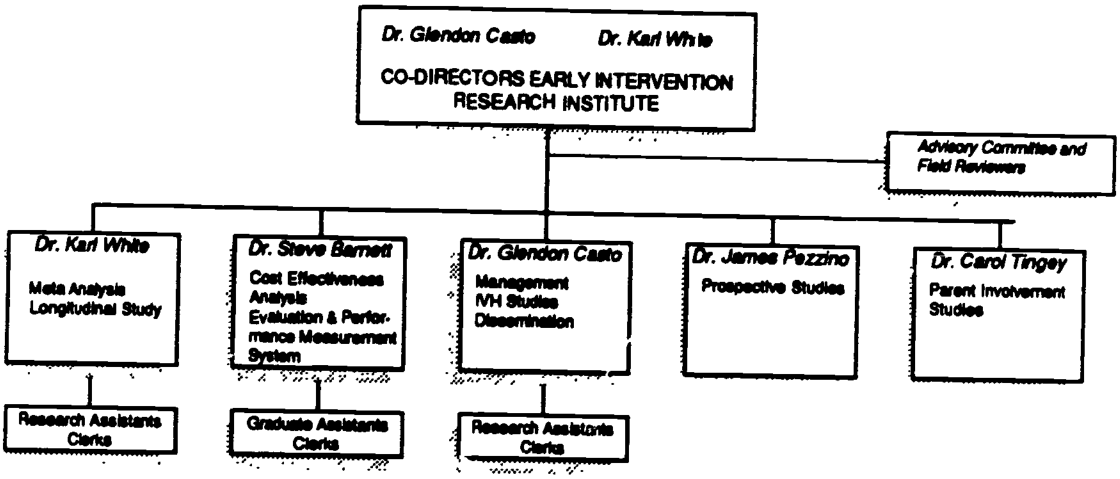
university campus. It is viewed by the University Administration as fulfilling a commitment to provide research, trained professionals, and services for handicapped citizens in the state and nation. It is noteworthy that during the last legislative session, the Utah State Legislature appropriated $50,000 to provide core support for the institute so that its activities might continue when the federal funding period is over.

The organizational structure of the Early Intervention Research Institute within the DCHP is depicted below. The director was responsible to the Director of the DCHP and directly to the Board of Directors of the Center. The DCHP Board of Directors is appointed by the President of the University as an independent policy board to oversee the affairs of the DCHP and its various programs. The board is made up of university administrators, community leaders, and consumers.



As can be seen from the institute's organizational structure below, individual members of the senior staff were assigned primary responsibility for each of the institute's major tasks. Solid lines indicate the primary governance structures for personnel evaluation, quality control, and accountability. The functioning of the entire insti e was initially under the direction of Dr. Casto, who received input from the advisory groups and was accountable to the DCHP Director and Board of Directors. At the end of Year 1, Dr. Karl White was named Institute Co-Director and continued in this capacity from that time on.

The present organizational structure and governance relationship of the EIRI itself are shown below.



STUDY DESCRIPTIONS

INVESTIGATION 1:
META-ANALYSIS OF EARLY INTERVENTION RESEARCH

Evaluation (i.e., efforts to determine the worth or value) of early intervention programs have taken many different forms--both large and small. For example, in 1975, a third-party evaluation contract was awarded to Battelle Institute of Columbus, Ohio, to evaluate the impact of early intervention demonstration programs funded by the Handicapped Children's Early Education Program of the U.S. Department of Education (HCEEP). One hundred twenty-nine randomly selected children in 29 projects from all over the U.S. were tested in areas including social, motor, cognitive, and communicative skills. Based on these data, the evaluators concluded that across all categories of handicapping conditions, children made 1-1/2 to 2 times greater gains than they would have been expected to make without the benefit of the project. Additionally, 97% of those parents interviewed perceived positive changes or improvements which they attributed to the project (Stock et al., 1976). Deweerd (1981) concluded that another indicator of the worth of HCEEP-funded demonstration projects was the fact that in 1979, 85% of the initial demonstration projects had secured funds to continue their programs and that the level of funding had increased. Literally hundreds of other research studies for both HCEEP-funded and other early intervention programs have collected data to determine the "worth" of such programs.

Unfortunately, the results and conclusions from such evaluations have been disturbingly discrepant. For example, there is growing agreement among practitioners that early intervention promises significant resolution or amelioration of some of the most persistent and expensive problems which educators face. According to Jordan et al. (1977):

> Programs providing early educational and therapeutic programming to meet the needs of young handicapped children and their families are reducing the number of children who will need intensive or long-term help. The importance of reaching handicapped children early and working to help them reach their full potential cannot be overemphasized. With early help, the sooner the better, these children can often function at higher levels than has been dreamed possible in prior years. (p. 26)

However, the promise and benefits of early intervention have not been universally accepted. As Hodges and Sheenan (1978) ~~pointed~~ out, "no consistent picture of success emerged from the early childhood education efforts of the 1960s. Although modest or robust immediate gains from structured programs were frequent, just as frequently, these gains eroded after the children left the experimental programs" (p. 4). Gottfried (1973) concluded that:

> Gains in cognitive and intellectual functioning attributable to preschool training were found by some projects but not others at the time of school entrance. However, there were no reports of substantial persistent gains beyond the third grade. Those studies which conducted school-age follow-up studies uniformly reported disappointing long-term results. (p. 286)

Even though the results of research should guide policy and practice, research on early intervention, when considered as a total body of evidence, has been somewhat confusing. Some researchers have reported success; others, failure. Some have suggested that early intervention is effective, but only for specific subgroups of children. Thus, even though the concept of early intervention has been heartily endorsed by individual practitioners and state and federal funding agencies (Swan, 1980), the research evidence is not at all clear. Even more important, the factors which account for the variation in research results have not been identified.

As primary research articles investigating the effectiveness of early intervention have accumulated, practitioners and policy makers have increasingly called for an effective integration of the knowledge which is being produced. In theory, the results of both basic and applied research on a given topic, such as early intervention, should culminate in increased knowledge and improved practice. In reality, however, the very important step of integrating the findings of the completed research on the effectiveness of early intervention into <u>conclusions which affect practice and influence policy</u> has not occurred.

Problems with Typical Efforts to Integrate Research Findings

In recent years, more and more researchers have realized that commonly used techniques for summarizing the results of completed research were inadequate (Glass, 1976; Jackson, 1980; Light & Smith, 1971). As Glass (1976) pointed out:

> We need more scholarly effort concentrated on the problem of finding the knowledge that lies untapped in completed research studies. We are too heavily invested in pedestrian reviewing where verbal synopses of studies are strung out in dizzying lists. The best minds are needed to integrate the staggering number of individual studies. This endeavor deserves higher priority now than adding a new experiment or survey to the pile. (p. 4)

The typical approach among social scientists to reviewing and integrating the literature on a given topic follows one of two routes. In both approaches, a group of easily accessible articles from fairly prominent journals or other publications are listed. In the first approach, the reviewer offers a verbal synopsis of the 20 to 40 research articles, and often concludes that the existing research is inconclusive--sometimes researchers find one thing, sometimes, another. A call is then made for additional research using better techniques and more precise methodology so that the truth of the matter can be discovered. In the second approach, the reviewer begins with a similar group of articles, but eliminates all but a few because of supposed design or analysis flaws. The findings of the remaining "acceptable" studies (frequently the work of the reviewer or his/her colleagues) are presented as the truth of the matter.

Both approaches to integrating and understanding previously completed research in the social sciences have serious inadequacies. Almost always, the articles selected for the review are only a small, non-representative fraction of the total research on the particular topic, and thus ignore a significant body of information. In addition, the "definitive" study almost never exists. Obviously, better design and analysis procedures are desirable, but it is not at all unusual for a series of well designed studies on the same topic in the social sciences to yield conflicting results.

Meta-Analysis Procedures

The problems which have been experienced with trying to integrate the existing literature on the effectiveness of early intervention are pervasive, but they are not intractable. Over the last decade, substantial efforts has been devoted to improving techniques for integrating the results of previous research (Glass, McGaw, & Smith, 1981; Hunter, Schmidt, & Jackson, 1981; Light & Pillemer, 1982; Rosenthal, 1978). Out of these efforts has evolved a set of procedures known as meta-analysis which have much potential for effectively summarizing the results of previous research.

Briefly described, conducting a meta-analysis requires the location of either all studies or a representative sample of studies on a given topic, converting the results or outcomes of the studies to a common metric, coding the various characteristics of studies that might have affected the results, and then using correlational and descriptive statistical techniques (both univariate and multivariate) to summarize study outcomes in a way that allows the examination of covariation of study characteristics with outcomes. In his critique of previous efforts to integrate the findings of research in the social sciences, Jackson (1980) concluded that the "meta-analysis approach is a very important contribution to the social science methodology. It is not a panacea, but it will often prove to be quite valuable when applied and interpreted with care" (p. 455).

Since its introduction, the meta-analysis approach has been used to review and integrate research findings on a wide variety of topics, including the relationship of class size to achievement (Glass & Smith, 1979), the relation of socioeconomic status and academic achievement (White, 1982), the efficacy of stimulant drugs for treating hyperactivity (Kavale, 1980; White & Myette, 1982), the effectiveness of training and reinforcement on standardized test results (Taylor & White, 1981), and the effectiveness of sensorimotor training with handicapped children (Kavale, 1982). In all, over 100 meta-analysis studies have been completed and reported. Although not all previous meta-analyses have been well done, it is clear that the meta-analysis techniques are being accepted as a useful methodology by substantial numbers of professionals.

It should be noted that some educational researchers have raised questions about the use and interpretations of meta-analysis (ERS, 1980; Eysenck, 1978; Gallo, 1978; Mansfield & Bussee, 1977; Shaver, 1979; Simpson, 1980). Some have questioned the results of a specific meta-analysis; others have raised cautions or concerns about the methodology per se. Most of these criticisms and cautions have been responded to in the literature (Glass, 1978; 1980; Glass & Smith, 1978; Glass et al., 1981). The most important point that such concerns have demonstrated is that meta-analysis, like all other research procedures, is not a fail-safe approach. However, the meta-analysis methodology, if properly implemented, has excellent potential as a tool for integrating existing research.

Purpose and Objectives

The purpose of this study was to apply meta-analysis techniques to as many research studies on the efficacy of early intervention as could be identified. All primary research studies were included in the meta-analysis which: (a) reported research on the efficacy of an intervention program designed to improve the cognitive, social/emotional, or life skills of handicapped, at-risk, or disadvantaged children; (b) began before children were 66 months old; and (c) were designed and reported so that an estimate of program impact could be calculated. Such estimates of impact were included from experimental, quasi-experimental, and pre/post designs.

The specific objectives of the study included:

1. To determine what past research reveals about the effectiveness of early intervention, including what factors and study characteristics (e.g., age of child, type of intervention, nature of the dependent variables, involvement of the family) covary with and possibly influence study outcomes.

2. To prioritize and focus future research efforts by identifying those research questions which need further investigation and replication as opposed to those questions which have already been sufficiently investigated, documented, and replicated.

The remainder of this section will briefly examine the adequacy of previous reviews of the early intervention literature to establish a foundation for the work described herein, describe the procedures used in the meta-analysis, and report the results of the first 156 studies included in the analysis. Findings in this report should be viewed as tentative, since additional studies are now being coded for future inclusion.

Analysis of Previous Reviews of Early Intervention Research

As in any systematic process of scientific inquiry, it was important, before beginning the meta-analysis of early intervention research, to examine previous efforts to accomplish the same goals. Such a "review of the literature" (in this case, an analysis of previous efforts to integrate early intervention research) served two main purposes. First, an analysis of previous reviews was necessary to determine whether there was a need for another review of the literature (e.g., was previous work methodologically sound; did sufficient evidence exist, i.e., primary research studies, exist to answer the questions of interest; was there substantial evidence which had not been included in previous reviews?). Secondly, an examination of previous work is important to plan for future work by establishing an appropriate point of departure and identifying the strengths and weaknesses of past investigations so that the former can be built upon and the latter avoided.

Previous Reviews Included in Analysis

A computer-assisted search of ERIC, Psychological Abstracts, CEC Abstracts, Dissertation Abstracts, Social Science Research, SSIE Current Research, and Index Medicus was conducted to identify previous reviews of the literature which dealt with (a) preschool or young children, (b) some form of intervention or treatment, and (c) handicapped, disadvantaged, or at-risk populations. Sixty-four review articles were identified by this search. A coding sheet was used to collect information about each review based on the following questions:

1. Does the reviewer critique previous reviews and explain how his/her review will differ from, expand, or replicate previous work?

2. Does the reviewer describe the procedures used to locate or delimit primary research studies used in the review?

3. What is the actual number of efficacy of early intervention studies used in the review to draw conclusions?

4. How did the author represent the results or findings of individual efficacy studies?

5. How did the reviewer consider data about how concomitant variables might covary with outcomes?

6. What variables were suggested by the reviewer as variables which might affect the effectiveness of early intervention (e.g., low vs. high SES subjects; or age at which intervention begins)?

7. What were the conclusions of the authors about common methodological weaknesses in the primary research included in the review?

8. What were the major conclusions of the review?

The 64 review articles included in the analysis were published between 1966 and 1982 in a variety of educational, psychological, and medical journals, as well as government reports, ERIC documents, and textbooks. The 64 reviews cited a total of 630 primary research studies to draw conclusions about the efficacy of early intervention.[1] Surprisingly, there was very little overlap in the primary research studies cited from review to review, as shown in Table 1. For example, 466 studies were cited in only one review, and only one study was cited in as many as 24 of the 64 reviews.

[1]Although 630 efficacy of early intervention "studies" were counted, these studies were reported in 1,027 unique articles. Often, there were multiple articles written about the same study. Primary research articles written by the same authors were considered to be from the same study, unless there was contrary evidence.

Table 1

Frequency with which Primary Research Studies on Efficacy of Early Intervention Were Cited by 64 Reviewers

Number of Reviewers Who Cited	Number of Studies
1	466
2	84
3	22
4	26
5	14
6-7	6
8-13	8
14-18	3
24	1
Total number of primary studies cited by reviewers	630

The particular research studies which were cited most frequently in these 64 reviews are shown on the following page in Table 2. The primary focus of most reviews was on disadvantaged populations, and 18 of the reviews did not consider handicap populations. The number of early intervention efficacy studies cited in each of the reviews ranged from 9 to 74, with a median of 16.5.

Is there a Need for Another Review of the Early Intervention Research Literature?

One of the most obvious evidences of need for another review of the early intervention research literature was the fact that although hundreds of early intervention efficacy studies were identified in this analysis, the average number of studies cited in existing reviews was only 16.5. Although some of the later efficacy studies would not have been available for earlier reviews, the correlation between

Table 2

Primary Research Studies Most Frequently Cited by Reviewers of Early Intervention Research Literature

Research Study and/or Principal Author(s)	No. of Reviewers Citing Study	Representative References
1. Weikart/Perry Preschool Project	24	Weikart (1967, 1968) Weikart et al. (1978)
2. Karnes	18	Karnes et al. (1969) Karnes et al. (1970)
3. Gray & Klaus/Early Training Project	17	Gray & Klaus (1965) Gray & Klaus (1970)
4. Skeels & Skodak	17	Skeels (1965) Skodak & Skeels (1949)
5. Heber & Garber/Milwaukee Project	13	Heber & Garber (1975)
6. Bereiter & Engelman/Direct Instruction	13	Bereiter & Engelman (1966)
7. Kirk	10	Kirk (1973)
8. Gordon/Florida Parent Project	9	Gordon (1968)
9. Caldwell	9	Caldwell (1967) Caldwell (1974)
10. Ramey/Abecedarian Project	8	Ramey (1974) Ramey & Campbell (1979)
11. Levenstein/Verbal Interaction Project	8	Levenstein (1976)
12. Hodges	8	Hodges & Spicker (1967)

year of publication for each of the reviews and number of efficacy studies cited was -.10. Thus, the failure to cite more efficacy studies does not appear to be a function of the number of articles available. The small number of efficacy studies cited, along with the failure to specify the criteria for inclusion/exclusion in most previous reviews, raises serious questions not only about the generalizability of conclusions, but also about the objectivity of the reviews. With hundreds of articles available on the efficacy of early intervention, one could probably find a dozen articles to support any point of view. The fact that so few studies are cited in most reviews is disturbing.

Another major problem with previous reviews is the way in which results of primary research studies are reported. Seventy-eight percent of the 1,500 citations of efficacy studies in the 64 reviews reported only that "differences" were found between experimental and control groups, or that the study demonstrated that the intervention was "effective" or "ineffective." The problems with such reporting are evidenced by the following typical statement taken from Stone (1975, p. 17):

> A number of intervention techniques have been reported to be of value to the developmentally delayed child. Among these are perceptual training (Frostig & Horne, 1964)...increasing the child's exposure to a variety of stimuli (Koegel, 1970), and increasing the discriminative aspects of individual stimulus (Horowitz, 1968).

When the outcomes of previous studies are reported in this manner, it is impossible for the reader to know whether differences between groups are educationally significant, statistically significant, or trivial. Consequently, it is difficult to know how much confidence to place in the conclusions of the reviewer.

Another important weakness in existing reviews of early intervention is the lack of attention to how subject or study characteristics may covary with results. For example, do studies which report interventions with very young children as subjects generally find larger benefits than studies which report interventions with older preschool children; or do studies with mildly handicapped children result in larger differences than studies with moderately or severely handicapped children? Seventy-five percent of the reviews either failed to consider the covariation of concomitant variables with outcomes or based conclusions about such covariation on less than 20% of the efficacy studies cited.

A less serious but nonetheless important weakness with existing reviews was their failure to consider previous reviews of the literature. Of the 64 reviews coded in our analysis, only two cited more than two previous reviews, critically described the procedures and conclusions of those reviews, and described how their review would differ from or improve on previous work. Forty-nine of the 64 reviews failed to cite any previous reviews of the literature. The failure to acknowledge and build upon the work of others is an important weakness that potentially impairs the quality of future work.

In summary, there are a number of important methodological weaknesses in previous reviews of the early intervention research literature. The number of efficacy studies cited in any given review is relatively small and probably not representative of the research which has been conducted. Techniques for examining the magnitude of outcomes and the covariation of subject and study characteristics have been inadequate. Little attention has been paid to earlier work which would permit a systematic building on the findings of others. Given these weaknesses, the amount of primary research which has been conducted to determine the efficacy of early intervention and the millions of dollars which are spent yearly to provide early intervention to handicapped, disadvantaged, and at-risk children underscores the need for high-quality integrative reviews of the literature. If properly done, such a review would provide important information to policy makers, program administrators, researchers, and practitioners about whether and how to implement early intervention programs.

Planning for Future Work

Table 3 lists the overall conclusions reached by reviewers in the 64 reviews considered. As shown in panel "A," most reviewers concluded that early intervention is generally effective if properly implemented. Specific benefits attributed to early intervention (see panel "B") included cognitive, academic, social, and attitudinal growth for the target child and improved functioning of the parents and the siblings.

Table 3

Conclusions About the Overall Effectiveness of Early Intervention

# and % of Reviews Drawing Conclusion		"A" GENERAL CONCLUSIONS
47	73.4%	o Early intervention is generally effective if properly implemented
9	14.1%	o Early intervention is effective, but only in special situations
7	10.9%	o Evidence about early intervention effectiveness is inconclusive
1	1.6%	o Early intervention is generally not effective
# of reviews drawing conclusions		**"B" SPECIFIC BENEFITS OF EARLY INTERVENTION**
11		o Increases IQ
9		o Improves academic achievement
7		o Enhances social skill
7		o Improves self-concept and emotional health
6		o Improves parents' behavior and attitudes
5		o Improves functioning of siblings
5		o Results in fewer children placed in special education programs
3		o Results in fewer children retained at grade level
3		o Improves language development
		"C" LONGITUDINAL EFFECTS OF EARLY INTERVENTION
5		o Gains made in early intervention programs are maintained
16		o Gains made in early intervention are not maintained
3		o Evidence about long-term maintenance is contradictory and more research is needed

Even though most reviewers concluded that there was sufficient evidence to document the immediate benefit of early intervention, there was much less support for long-term benefits. Of those 23 reviews in which the longitudinal effects of early intervention were considered, only 5 (22%) concluded that the gains attributable to early intervention programs were maintained, 15 (65%) concluded that gains were not maintained, and 3 (13%) concluded that there was not sufficient evidence to draw conclusions.

Table 4 lists the most frequently cited variables which might be associated with or influence the success of early intervention. Table 5 shows the conclusions most frequently drawn by reviewers as they relate to variables cited in Table 4. Not every variable listed in Table 4 is represented in Table 5 because many reviewers

Table 4

Variables Suggested by Previous Reviewers as Potential Mediating Variables and the Number of Different Reviews in Which Each was Cited[a]

# of reviews citing variable	Intervention Variables	# of reviews citing variable	Subject Variables
23	Degree of parental involvement	2[illegible]	Age at which intervention begins
24	Degree of structure in intervention	20	Socioeconomic status (SES)
14	Training/competence/attitude of intervenor	17	Degree of environmental stimulation/deprivation in home setting
12	Nature of intervention (e.g., philosophical orientation or type of curriculum)	14	Parent/child relationship, and whether family is intact
10	Use of operant conditioning principles	8	Nutritional level, health care, or immunization
9	Parents' attitude and motivation	8	Severity of handicap
9	Degree to which instructional level is appropriate for target child	8	Race
		7	Sex
8	Amount of intervenor/child interaction	6	IQ level of child before intervention
8	Parent modeling of correct behavior aside from explicit intervention involvement	5	Type of handicap
		5	Previous preschool experience
7	Degree of individualization		
6	Intervenor/child ratio		
6	Continuity between preschool/school programs		
6	Site intervention (center vs. home)		
5	Clarity of program goals		

[a]Only variables which were suggested by five or more reviewers are included in this table.

Table 5

Conclusions About How Mediating Variables Are Related to, or Influence, Intervention Effectiveness

(Numbers in parentheses indicate the number of reviewers taking that position)

Mediating Variable	Pro	Con
Degree of parental involvement	Interventions that involve parents are most economical and most effective (12)	
Age at intervention	The earlier the age at which intervention begins, the greater the gains (14)	Similar gains result from successful programs regardless of age of entry (5)
Critical age	Efforts at intervene after the critical period becomes progressively less effective (8)	There is no indication of a critical period in which early intervention is most effective (3)
Degree of structure in the intervention	More structured intervention programs result in greater gains (12)	Degree of structure in the intervention is not related to intervention effectiveness (1)
Nature of intervention	Curriculum type per se is unrelated to intervention effectiveness. However, more comprehensive curricula (including cognitive, behavioral, and social-emotional components) are more effective (6)	
Training/competence/ attitude of intervenor	Better trained, more competent interventionists result in more effective programs (4)	
Length of intervention	Longer programs result in greater gains (4)	Length of intervention is unrelated to child gains (4)
Center vs. home-based	Home and center-based programs, if well implemented, are equally effective (4)	
Individualization	Individualized intervention is more effective	
Socioeconomic status (SES)	Low SES children make greater gains in gross motor skills, and high SES children make greater gains in IQ (6)	
Race	Race is unrelated to intervention effectiveness (1)	Black children gain significantly more from early intervention than white children (2)
Severity of handicap	Severity of handicap substantially influences program success (2)	Severity of handicap is unrelated to program success (1)
Sex	Boys make greater gains than girls on some outcomes (1)	Gains are unrelated to sex of the child (1)

cited a concomitant variable as important, but did not draw specific conclusions about that variable. The most frequently drawn conclusion of the 64 reviews considered were related to the involvement of parents, the age at which intervention begins, and the degree of structure in the intervention program. As can be seen in Table 5, there was a fair degree of disagreement among reviewers about the influence of many of the variables cited.

This information does much in planning for another review of the early intervention research literature. First, these data emphasize that any additional efforts to integrate the research on early intervention needs to focus on both immediate and long-term benefits,needs to examine outcomes in a variety of areas (IQ, academic achievement, social skills, self-concept, functioning of parents and other family members, etc.), and needs to examine the covariation with study outcomes of a variety of subject (e.g., age at which intervention begins, socioeconomic status, race, sex, etc.) and intervention (e.g., degree of parental involvement, degree of structure in intervention, training of intervenor, etc.) variables. The results of this analysis identify those variables which have been suggested most frequently as well as variables which have been cited infrequently but may still be important.

Summary

The analysis of previous reviews of the early intervention research literature definitely established the need for another integrative review. Given the large number of existing early intervention efficacy studies, the meta-analysis techniques described below seem like a potentially valuable set of procedures for making sense of this large data base. The methodological weaknesses identified in previous reviews underscores the need for conducting another review. The conclusions of previous reviews, both in terms of immediate and long-term benefits of early intervention, and the subject and study characteristics which are reported to covary with intervention effectiveness, identifies the key information which needs to be collected and interpreted in conducting such a review.

Procedures

Included in this section is a description of (a) the procedures used in selecting and identifying early intervention efficacy studies to be included in the meta-analysis, (b) the procedures used in developing the coding system and conventions, and (c) the procedures for coding the articles included.

Identifying Studies to be Used in the Meta-Analysis

Efforts were made to include any study of the efficacy of early intervention with handicapped, at-risk, or disadvantaged children which began before 66 months of age and provided information which could be used in estimating the benefit of the intervention program. Estimates of benefit were derived from pre-post, true experimental, and quasi-experimental designs. Single subject research designs have not been included at this point because the type of data yielded by such designs is difficult to incorporate with more traditional group designs in a meta-analysis data set. However, various alternatives are currently being explored that will enable us to utilize this valuable data set as the results of the meta-analysis are expanded during 1983-84.

The first step in identifying articles was a computer-assisted literature search conducted at the Utah State University library through the DIALOG system. This computer-assisted search was done of the ERIC, Psychological Abstracts, CEC Abstracts, Dissertation Abstracts, Social Science Research, SSIE Current Research, and Index Medicus data bases. Very broad guidelines were set deliberately for this search in an effort to include as many studies as possible so that appropriate studies would not be missed. This search resulted in the identification of 1,402 articles which were then sorted by staff members into the approximately 800 articles which reported efficacy studies and those which reported other information about early intervention. Each article was then screened to determine if it reported information on an early intervention program which began before 66 months of age for subjects which were handicapped, at-risk, or disadvantaged and provided some data from which an estimate of the magnitude of program effect could be estimated. Articles which passes this initial screening were then put in the "To be Coded" file. Articles which were rejected at this stage were independently checked by another staff member to make sure that relevant articles were not excluded.

It is interesting to note that the computer-assisted search was not a very effective means of identifying articles to be included in the meta-analysis. Of the almost 1,800 articles obtained thus far in the meta-analysis effort, only 305 (less than 20%) came from the computer-assisted search. Most of the articles that have been identified were obtained through references of other articles already in the files.

In addition to the computer-assisted search and the bibliographic searches of articles already obtained, letters were sent to each of the HCEEP demonstration and outreach project directors and to all members of the EIRI Advisory Committee and field reviewers asking them to identify additional studies of early intervention efficacy that may not have been identified in our search.

Once articles were obtained for the meta-analysis, a very specific set of procedures was followed in preparing them for coding, following them through the coding process, and preparing data from the coding for analysis.

Development of Coding System and Conventions

A coding system was developed to collect information about each article included in the meta-analysis. Information collected about each study included:

- A description of the <u>subjects</u> included in the research,
- The type of <u>intervention</u> used,
- The type and quality of <u>research design</u> employed,
- The type of <u>outcomes</u> measured and procedures used, and
- The <u>conclusions</u> reached by the study.

The specific items included on this coding sheet were identified using the analysis of previous reviews so that variables which other authors suggested as important were included. In addition, coding systems used in previous meta-analyses were examined and useful features incorporated. This first draft of the coding sheet was then "pilot tested" by members of the meta-analysis team on eight different articles. Several revisions of the coding system were done during this process.

For each item on the coding sheet, conventions were written which provided operational definitions for coding. Because it was impossible to specify every eventuality that would be encountered in coding studies, coders were also instructed to provide "coding clarifications" for items they coded for which the conventions were not a good "fit." These coding convention clarifications were discussed with the entire staff each week in staff meeting, and corrections and revisions to codings were made based on that information.

One of the most important pieces of information collected about each study was the estimate of program effect. Two types of "Effect Sizes" were collected wherever possible: a standardized mean difference Effect Size and a variance Effect Size. A standardized mean difference Effect Size was obtained for every study. The standardized mean difference was defined as the $(X_E - X_C) \div SD_C$ (Glass, 1976). This standardized mean difference Effect Size measure converts all scores to a standardized score which has similar meaning across different types of variables. For example, an Effect Size of 1.0 on a measure of IQ indicates that the average person in the experimental group is 1 standard deviation or 15 points above the average person in the control group. An Effect Size of 1.0 on a reading test has approximately the same meaning, although it may be 25 points or 5 points depending on the metric of the test being used. For every measure, an Effect Size of 1.0 represents 1 standard deviation difference between the average score of each group and indicates that the average person in the experimental group would score at the 84th percentile of the control group (assuming normal distribution).

Using a standardized metric for outcome avoids problems of interpretation due to statistical artifacts which are dependent on sample size and allows the covariation of outcome and study/subject characteristics to be examined more completely. Unfortunately, means and standard deviations were not reported in all studies. In these cases, formula for converting _F_ statistics, _t_ statistics, analysis of variance tables, regression equations, and proportions to Effect Sizes were used (see Glass, McGaw, & Smith, 1981).

Procedures for Coding Studies

As noted earlier, written procedures for coding each study were developed. SEveral procedures were used to increase the accuracy and consistency of coding and are worth emphasizing here. First, after each article was coded, the coder would take the article and the coding sheets to another member of the team and provide a very brief synopsis of the type of design used in the study, which outcomes were coded, and which information in the article was used to compute the effect sizes. The "checker" would then check the logic of which outcomes had been selected and independently calculate effect sizes for those outcomes. This independent calculation would then be checked against the written computations which were done by the original coder. In addition, the checker would examine key variables on the coding sheet, check that every blank of the coding sheet was filled in, and make sure that the "checklist" on the first page of the coding sheet had been properly completed. At that point, if mistakes had been found, the issue would be resolved with the original coder, and then the coding packet would be turned in.

In addition to this checking of every article, interrater consistency checks were done for 10 articles included in the meta-analysis. The results of these interrater consistency checks are shown in Table 6. As can be seen, the average "exact" agreement (i.e., the most conservative estimate) on coding was 86.3 across the 10 studies. Not counting it as a disagreement when one coder chose to leave an item blank and another coder chose to make an educated estimate of an item, this

Table 6

Summary of Interrater Consistency Checks

References	DATE	EXACT	ONE STEP	ONE STEP & BLANKS	% ES IN COMMON	KARL	GLENDON	CIE	DAVID	DENNIS	GARY	MARGO	DUANE
Brassell, W. R., Dunst, C. J. (1978). Fostering the object construct: Large scale intervention with handicapped infants. American Journal of Mental Deficiency, 82, 507-510.	2/11/83	80.4	82.0	84.1	100	X	X			X	X		
Gavrin, J., & Sacks, L. S. (1963). Growth potential of preschool aged children in institutional care: A positive approach to a negative condition. American Journal of Orthopsychiatry, 33, 399-408.	2/19/83	92.3	92.9	94.8	24	X	X	X	X				
Blank, M., & Solomon, F. (1968). A tutorial language program to develop abstract thinking in socially disadvantaged preschool children. Child Development, 39, 379-389.	3/8/83	88.46	90.17	92.10	100	X			X	X	X		
Carlsen, P. M. (1975). Comparison of the occupational therapy approach for healing the young cerebral palsied child. American Journal of Occupational Therapy, 29, 267.	3/30/83	85.03	86.62	87.90	100	X		X	X		X		
Harris, S. R. (1981). Effects of neurodevelopmental therapy on motor performance of infants with Down syndrome, Dev. Medchild Neurol., 23, 477.	4/1/83	84.21	85.48	87.16	86	X	X			X			
O'Connell, J. C., & Farran, D. C. (1980). The effects of day care intervention on the use of intentional communication behaviors in socioeconomically depressed infants. Paper presented at the Biennial Southeastern Conference of Human Development, Alexandria, Virginia (ERIC Document Reproducation Service No. ED 195 359).	4/14/83	90.10	91.17	93.46	100	X	X		X				
Banta, T. W., Higginbotham, L., & Levin, M. (1979). Evaluation of East Tennessee's child health and development project. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA (ERIC Document Reproducation Service No. ED 175 174).	4/15/83	85.03	86.78	8/.74	100	X	X	X		X			
Scott, R. (1974). Research and early childhood: The Home Start Project. Child Welfare, 53, 112-119.	4/20/83	90.73	92.43	90.70	67	X		X			X		
Scarr-Salapatek, S., & Williams, M. L. (1972). A stimulation program for low birth weight infants. American Journal of Public Health, 62, 662-667.	6/28/83	79.2	80.0	82.5	75	X	X	X		X	X		
Fuschillo, J. C. (1968). Enriching the preschool experience in children from age 3: The evaluation. Children, 15, 140-143.	9/17/83	87.90	88.54	90.96	56	X	X				X	X	X
		86.5	87.7	89.2	93.0	10	7	5	4	5	6	1	1
		Median Value							Total #				

figure increased slightly to 89%. The median value for the number of effect sizes chosen in common by various coders was 93%. The relatively high interrater consistency and the fact that checks of the computation and coding procedures were made for every article suggest a high level of consistency across the articles included in the meta-analysis.

Finally, it was noted earlier that because means and standard deviations were not always reported in the article, alternative computational formula needed to be used. Since the logic of the standardized mean difference effect size is based on the use of standard deviation of the control group, it was felt to be important to check the degree to which the use of alternative effect size computational formula might bias results. This was done by computing alternative effect sizes whenever an article provided enough information to compute an effect size using the means and standard deviation of the control group and to compute an effect size in other ways. The effect sizes used in the meta-analysis reported in the results section was always based on the means and standard deviation of the control group when that information was available. However, in those cases where alternative forms of information were available, the average effect size was extremely close (usually within less than .05 of a unit).

Results and Discussion

The results of the investigation thus far have proven to be both enlightening and provocative. They have confirmed some commonly accepted positions, called others into question, and identified the most total absence of empirical data for either supporting or refuting others. The results section will summarize several of the major findings and outline the implications of those findings for conducting further research (more detailed summaries of these same data are given by Casto, White, & Taylor, 1983; White & Casto, 1985).

Characteristics of the Data Set

As shown in Table 7, the data set consists of 2,266 effect sizes from 326 different studies. Data from experimental/control studies, A vs. B studies, and single-subject design studies were analyzed separately because of the fundamentally different questions they address. There are 1,121 effect sizes from intervention vs. control studies. Most of those (906) are from studies with disadvantaged children, 215 are from studies with handicapped children, and 85 from studies with medically at-risk children. In terms of the handicapped population, the vast majority of effect sizes come from studies which considered the effects of intervention with either mentally retarded children or with groups of children exhibiting a mix of handicaps, but predominantly mild to moderate mental retardation.

Table 7

Average Effect Size for Intervention Versus Control Early Intervention Efficacy Studies for Subgroups of Data

	Handicapped		
	ES	S_{es}	N_{es}
All studies	[illegible]8	.05	215
Only good-quality studies	.40	.13	13
Only good-quality studies with immediate posttest	.43	.15	20

ES = mean effect size

S_{es} = standard error of the mean for ES

N_{es} = number of ESs on which a calculation is based

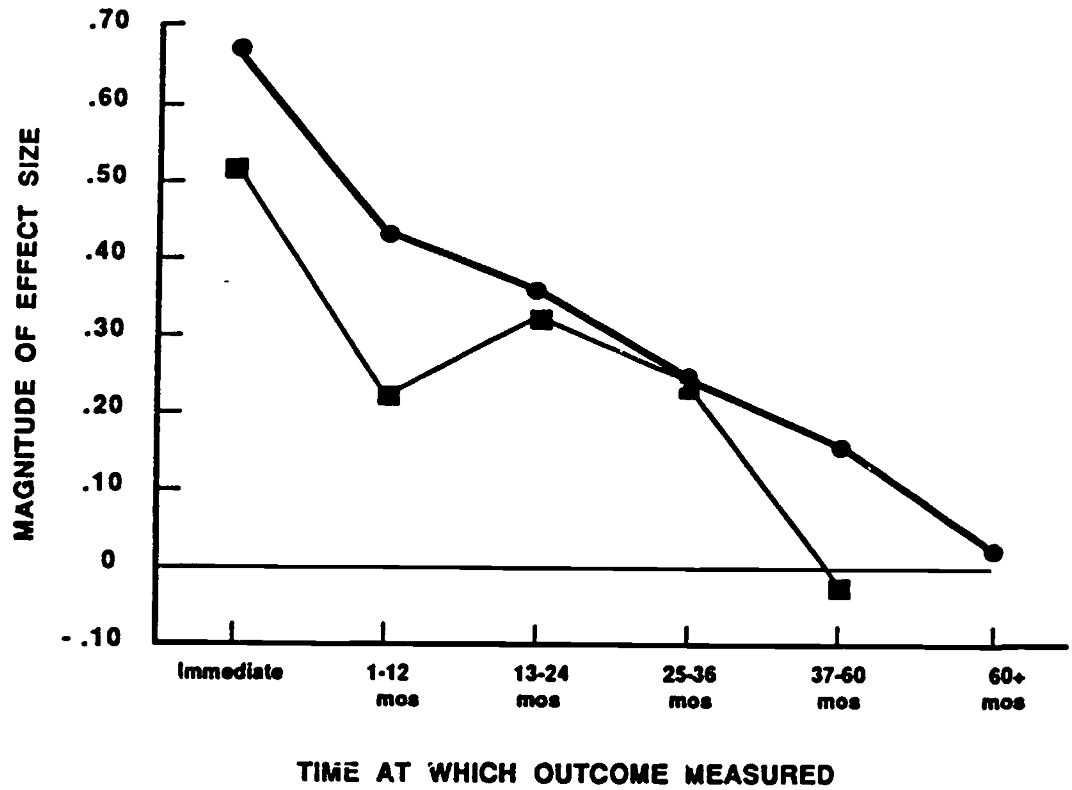
Included in the data set are also 984 effect sizes from intervention A vs. B studies which have examined such questions as earlier vs. later, intervention intensity, and degree of parental involvement. We have also examined the results of 75 single-subject design studies, most of which have focused on increasing social interaction, correcting conduct disorders, or developing language competence.

The data set includes studies reported from 1937 to 1984, with 70% since 1970. Most of those studies were reported in educational or psychological journals, but significant numbers came from medical journals, books, unpublished documents through ERIC, and government reports. IQ has been the most frequently measured outcome, accounting for over 40% of all outcomes. There has also been substantial reporting of language, academic functioning, and motor functioning, but relatively scarce consideration of social-emotional growth or family outcomes. Over 60% of the effect sizes were measured immediately at the conclusion of the intervention, and only 11% were measured more than 36 months after intervention concluded, all of those with disadvantaged samples.

Overall Effects of Intervention

Turning now to what can be concluded from the data about the efficacy of early intervention, we have considered the evidence about early intervention effectiveness separately for studies done with disadvantaged, at-risk, and handicapped children because we believe that the children in each group are so different that combining the data would be misleading. Unfortunately, in many previous reviews of early intervention efficacy literature, conclusions about the effects of a particular type of intervention for handicapped children have been based primarily on data from studies with disadvantaged children (see White, Bush, & Casto, 1985). Although there is certainly some limited applicability for certain issues, the general practice is highly suspect.

Most of the available evidence about immediate and long-term benefits of early intervention is for disadvantaged children. As shown in Figure 1, the best estimate



all studies	# ES's # of studies	546 (99)	110 (24)	77 (24)	28 (7)	53 (9)	82 (7)
good studies	# ES's # of studies	153 (31)	21 (9)	23 (10)	15 (6)	13 (6)	

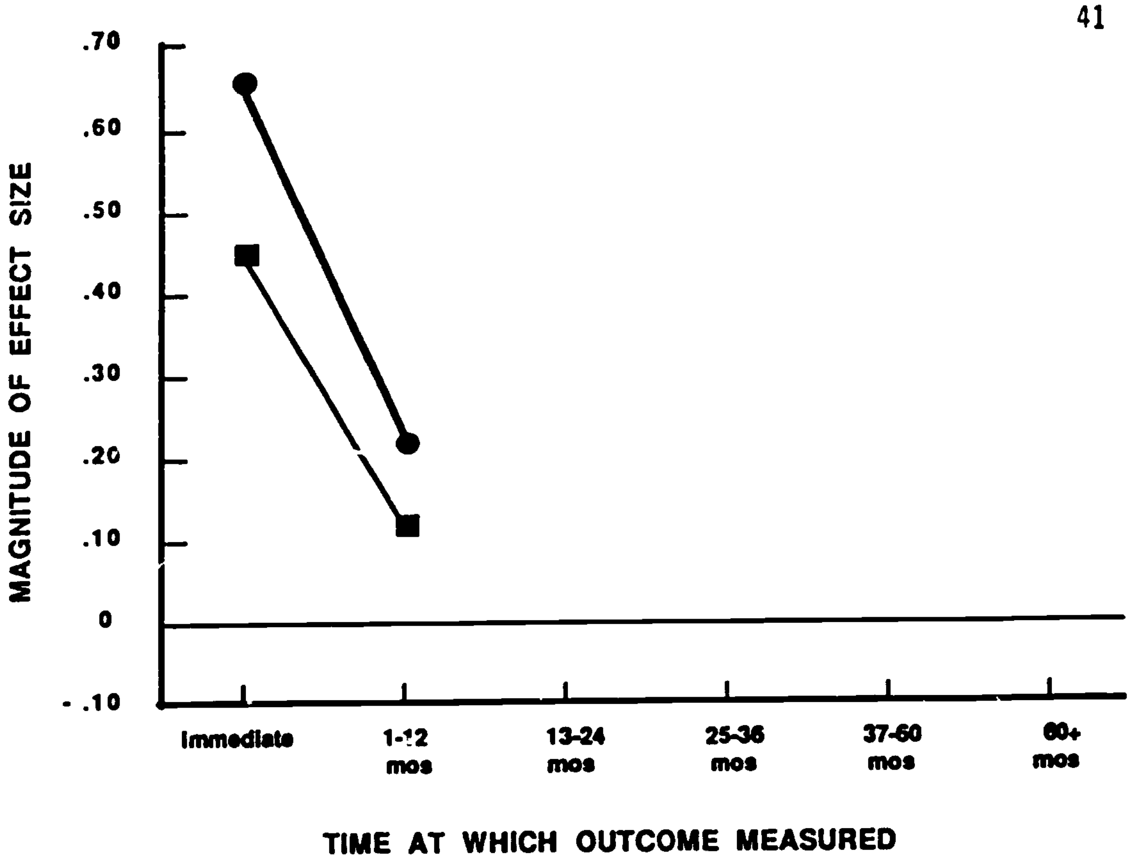
of the immediate effect of early intervention for disadvantaged children is approximately 1/2 a standard deviation. In other words, for IQ measures, this represents a gain of about 8 points; for motor functioning, an improvement from the 30th to the 50th or the 10th to the 22nd percentile; for reading achievement at the second grade, a gain of approximately 10 months worth of reading. Those are substantial effects that are of obvious clinical importance. The magnitude of these immediate effects is similar for every domain, and, generally speaking, for programs using different philosophical approaches conducted in different settings with involvement of different types of intervenors.

Over time, there is a clear tendency for the measurable effects of intervention with disadvantaged children to be substantially reduced. Several caveats are important in interpreting that data. First, when the results are limited to studies of good methodological quality, these results are based on relatively few data, and there are some significant exceptions to the tendency for data to wash out over time. For example, the recently reported Perry Preschool Project data (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984), which was conducted with disadvantaged children, is not yet included in this data set because of the recency of its publication. This is a well-designed study which reports sub tantial effects for early intervention on variables, such as high school graduation rates, employment, and teenage pregnancy rates. On the other hand, it is important to point out that there have been other reasonably well-designed studies that have failed to find long-term effects in s of the same areas.

A second caveat is th presently available long-term data are predominantly IQ and academic achievement data. As more studies collect data, such as that reported in the Perry Preschool Project, a different picture may emerge.

A third problem is that many people mistakenly assume that long-term effects are essential in demonstrating the efficacy of early intervention. If the only object of early intervention for disadvantaged children is to permanently change the measured IQ of those children, then these data are discouraging. However, the abundant evidence for other important short-term benefits should not be ignored.

For handicapped children, as shown in Figure 2, the data is much easier to interpret because there is so little of it. Also, in contrast with data about the efficacy of early intervention with disadvantaged children where 25% of the available data comes from studies of high methodological quality, only 16% of the data for handicapped children comes from studies of high methodological quality. Furthermore,



		Immediate	1-12 mos	13-24 mos	25-36 mos
all studies	# ES's	179	17	8	1
	# of studies	(65)	(8)	(3)	(1)
good studies	# ES's	20	13		
	# of studies	(11)	(2)		

there is no follow-up data collected more than 12 months after the intervention was completed from studies of high methodological quality. The best estimate for immediate effects is about 4/10 of a standard deviation. However, this estimate is based on only 20 effect sizes from 11 different studies. The estimate of program impact when all studies are included is almost double (.72), suggesting that some of the perception about the magnitude of benefits for handicapped children is based on data which are of questionable validity. However, the fact remains that there is evidence of a strong and replicable immediate effect for handicapped children based on studies of only good methodological quality. Unfortunately, those who claim that early intervention for handicapped children results in long-term impact are arguing in the absence of data.

Effects of Mediating Variables

As noted earlier, White et al. (1985) identified a number of mediating variables suggested by previous reviewers as contributing to the effectiveness of early intervention. Data from our analysis of previous research are also useful in determining whether these frequently advocated positions can be empirically supported. Because of space limitations, only a few examples will be given, and the analyses on which these conclusions are based are much more detailed than reported here (a more complete description is contained in Casto et al., 1983). Furthermore, these analyses are limited to early intervention studies with disadvantaged children since so few data which address the above questions are available from good-quality studies with handicapped children.

Involvement of parents in intervention programs. One of the most frequent conclusions in the early intervention efficacy literature is that programs which involve parents are more effective than programs which do not (e.g., Bronfenbrenner, 1974; Comptroller General, 1979; Goodson & Hess, 1975; Hewett, 1977; Weikart et al., 1973). Although intuitively logical, we have been unable to find strong empirical support for this position. As shown in Table 8, when values are adjusted for age at which intervention began, time of measure, and quality of dependent variable, the average of 684 effect sizes from 80 studies inn which parents were not used at all or only used to a minor degree was .42. The average of 200 effect sizes from 27 studies in which parents were utilized as the major or only intervenor was .41. When effect sizes are limited to only high-quality studies, there is still very little difference between programs which utilize parents extensively and those which do not.

Similar results were obtained when data were examined for whether the program was center-based, home-based, or combination home- and center-based; whether parents or parents and children were the target of the intervention; and the degree to which the intervention program intended to involve parents.

Table 8

Average Effect Sizes for Different Levels of Parent Participation in the Intervention

	PARENT PARTICIPATION	ADJUSTED for differences on quality of outcome measure and time of measurement					
		ALL STUDIES				GOOD STUDIES	
		ES	S_{es}	N_{es}	($n_{studies}$)	ES	(N_{es})
DISADVANTAGED	Minor or not at all	.52	.03	684	(80)	.40	(171)
	Major or only	.42	.04	200	(27)	.51	(54)
HANDICAPPED	Minor or not at all	.72	.07	137	(48)	.38	(17)
	Major or only	.59	.09	70	(27)	.43	(6)

The most direct information about the effect of parental involvement was obtained from nine studies which had made direct comparisons between different levels of parental involvement (Abbott & Sabatino, 1975; Bidder, Bryant, & Gray, 1975; Gordon, 1969; Karnes, Teska, & Hodgins, 1970; McCarthy, 1968; Miller & Dyer, 1975; Nedler & Sebra, 1971; Radin, 1971; Ramey & Bryant, 1983). As can be seen in Table 9, when all 134 effect sizes from these studies were considered, there is a slight advantage for programs which involve parents more extensively (ES = .08). However, these fincings are heavily influenced by the Gordon (1969) study, which found an average advantage of .18 for interventions which involved parents. Although the methodological quality of the Gordon study is quite good, many of the ESs comparing different levels of parental involvement were confounded with age at start, program duration, and setting. Thus, the results from Gordon (1969) should be viewed cautiously. The other eight studies yielded an average effect size of .06, favoring programs which did not involve parents.

Table 9

Average Effect Sizes for Within-Study Comparisons of Different Levels of Parental Involvement

DEGREE OF PARENTAL INVOLVEMENT[a] (More vs. Less)	ES	S_{es}	N_{es}
All Comparisons			
Parent vs. No Parent or More vs. Less	.08	.05	134
Gordon Study Only			
More Involvement vs. Less Involvement	.18	.06	70
All Comparisons Except Gordon Study			
Parents vs. No Parent or More vs. Less	-.06	.09	64

[a]ESs from nine studies

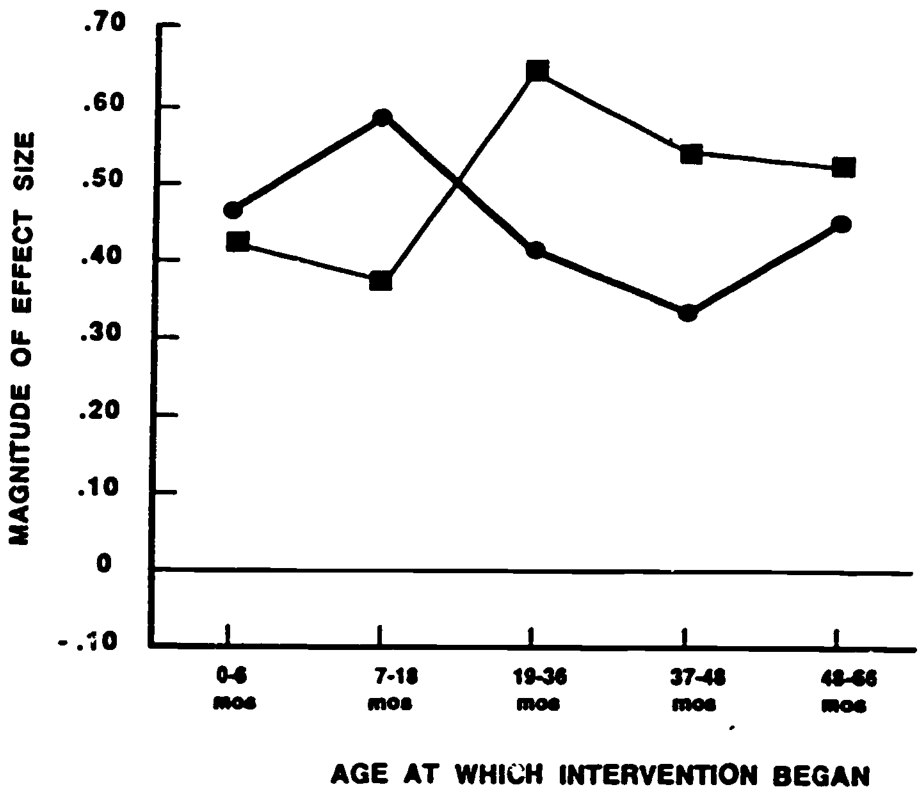
Taken together, these data suggest that programs for disadvantaged children which involve parents extensively can be effective, but it does not appear that they are any more effective than programs which do not involve parents.

Admittedly, this is counter to the intuitively logical position advocated by many people. Before drawing conclusions about whether parents should be involved in early intervention programs, it is important to note the limitations of the available data First, most of the arguments in support of involving parents in early intervention programs have come from studies done with disadvantaged and at-risk children (e.g., Bronfenbrenner, 1974; Gordon, 1969; Rescorla, Provence, & Naylor, 1982). Children from such families often come from large families with high incidence of single-parent families, poverty, and other stressors, and low levels of parent education--all of which may hinder effective parent participation. Thus, it may be that effective tests of parent involvement have not been done. Secondly, many of the outcomes included in this meta-analysis (over 40%) are from measures of IQ. It may well be that the involvement of parents leads to gains in other areas which simply have not been investigated. Finally, very few of the effect sizes (less than 2%) came from studies in which the investigators verified that parents were actually involved to the degree intended. Thus, it may be that some investigators were examining intended rather than actual parent involvement. These caveats notwithstanding, there are no data at this time to confirm the widely held belief that involvement of parents leads to more effective intervention.

Age at which intervention begins. Another frequently stated position in the early intervention literature is that the earlier the child is involved in a program, the more effective the program will be (e.g., Bronfenbrenner, 1974; Comptroller General, 1979; Garland, Swanson, Stone, & Woodruff, 1981). In spite of the popularity of this position, these analyses provide only meager empirical support from intervention versus control studies for this position. As shown in Figure 3, average effect sizes for studies comparing experimental to control group children beginning at different ages are very similar after adjustments are made for time at which the outcome was measured and quality of the dependent variable.

As shown in Table 10, the 17 ESs from five studies which made direct comparisons of starting children at two different ages with all other variables held constant (Braun & Caldwell, 1973; Caldwell & Smith, 1970; Gordon, 1969; Jason, 1975; Morris & Glick, 1977) show .04 of a standard deviation advantage for those children which begin later. Other studies have examined the effect of age at start, but have been substantially confounded with other variables such a duration or setting (Beller, 1969; Gordon, 1969; Scott, 1974; Strickland, 1971). These "confounded" studies shown an average of .16 favoring children who began earlier.

Taken together, these data raise questions about the commonly held position that "earlier is better." Unfortunately, there is very little evidence available, and most of it is for disadvantaged rather than handicapped children. Furthermore, none



all studies	# ES's # of studies	158 (18)	79 (17)	110 (19)	226 (25)	315 (34)
good studies	# ES's # of studies	73 (8)	22 (5)	7 (4)	88 (11)	37 (17)

Table 10

Average Effect Size for Within-Study Comparisons of Age at Which Intervention Begins[a]

	ES	S_{es}	N_{es}
Effect sizes from unconfounded studies	-.04	.08	17
Effect sizes confounded with other variables (e.g., intensity, setting)	.16	.06	101

[a]Data based on eight studies

of the direct comparisons came from studies which high ratings of methodological quality. Most disturbing is that in spite of the frequently stated position that "earlier is better," very few empirical studies have even addressed the issue of time at which intervention begins.

Conclusions

Where does all of this leave us with respect to early intervention for handicapped, disadvantaged, and at-risk children? First, the paucity of available empirical data, particularly from well-designed studies, for many of the most important questions is disconcerting. Contrary to the conclusions of many previous reviewers and textbook authors, there simply is not enough information to be confident about the long-term impact of early intervention with handicapped children, and evidence in support of many of the commonly held positions about mediating variables (e.g., parental involvement or age at start) is either non-existent or contradictory.

Second, there is strong support for the immediate positive effects of intervention with disadvantaged children and emerging support for long-term benefits. Furthermore, data are beginning to accumulate which support the immediate benefits of intervention for handicapped children. Although there is not yet evidence of long-term benefits of handicapped children, it is clear that handicaps identified before age 5 seldom disappear in later years. Thus, the obvious costs and effects of not intervening suggest that some type of intervention is necessary--the real question is what type of intervention at what age.

Third, it is important for us to keep science in its proper perspective. As pointed out by the National Academy of Science:

> As the growth in public programs that benefit children has accelerated, [questions about how those services can be provided most equitably and efficiently] have become more insistent, and controversies among elected officials, practitioners, and parents concerning public policies affecting children have become more intense and widespread. Because these issues are both

value-laden and political, they will not be 'resolved' by research alone. (Hayes, 1982)

But recognizing the limited role that research does play in establishing public policy and still being supportive of and advocating for better services for children does not excuse sloppy science.

One of the clearest findings from the analysis of previous investigations of the efficacy of early intervention is the need for more extensive and higher quality research.

A number of other investigators have lamented the ambiguity caused by methodological flaws in early intervention efficacy research (Dunst & Rheingrover, 1981; Simeonsson, Cooper, & Scheiner, 1982). Fortunately, some of the most serious problems with previous studies are reasonably resolvable. First, in spite of the difficulties involved, the use of randomized designs in early intervention efficacy research is neither unethical nor impossible. Random assignment to groups is especially feasible and advantageous in those cases where the number of families in need of services far exceeds the capacity of service agencies to provide comprehensive services or in those cases where alternative treatment programs are being considered, such as half-day vs. full-day programs.

In many cases, the alleged ethical compromises of "withholding treatment" do not exist because research can be implemented in areas where very little treatment now exist, or in some cases where we have no information about which alternative treatment is actually better. For example, for parents of some handicapped children, respite may be more important and beneficial than increased involvement. Or parents may actually prefer fewer instead of more frequent home visits.

There are also numerous instances where the use of randomized designs has demonstrated the lack of efficacy for treatments which were though to be beneficial, but actually were not. One of the clearest examples of this is the administration of high concentrations of oxygen to premature infants. As is well known, during the late 1940s and early 1950s, hospitals began administering concentrated doses of oxygen to low birthweight babies to reduce the negative complications of frequent respiratory distress experienced by such infants. Almost by accident it was discovered that in hospitals where this practice occurred, there was a dramatic increase in retrolental fibroplasia which frequently led to blindness. Research was not immediately implemented, however, because many believed it would be unethical to withhold something as necessary as oxygen for infants suffering from respiratory distress. When randomized trials were done, it was conclusively established that the oxygen was causing an increased incidence of blindness. As early intervention researchers, we should continually ask ourselves whether our hesitancy to conduct randomized trials comparing different forms of intervention may also be delaying progress.

A second, relatively simple procedure, which would substantially improve the quality and credibility of early intervention efficacy research, would be to use data collectors who are unaware of the group membership of subjects. Only 21% of the effect sizes included in our analyses came from studies in which the data collector was definitely "blind." Unfortunately, the educational and medical literature is rife with examples of ineffective or even harmful treatments that have been advocated by well-intentioned people who believed that their treatment was making a difference, when, in fact, it was not. We need not look far to see examples, such as dietary treatment of hyperactivity, Doman Delacato therapy for learning disabled children, or even blood letting in the not too distant past, of people seeing what they expected to see. The use of "blind" data collectors would eliminate this serious threat to the credibility of study results.

Another serious problem with previous research is that virtually all of the existing early intervention efficacy research has failed to determine the extent to which the intended treatment was actually implemented. For example, in programs which intended to utilize parents as intervenors in their child's program, our analysis of previous research identified virtually no instances of evidence that parents actually did become involved to the degree intended by the program designer. Unless such information is obtained, there is a real danger that comparisons are being made between programs which were intended to be different, but which were not. The failure to verify that intended treatments were actually implemented may be responsible in part for the failure of previous research to detect differences between alternative intervention programs.

Finally, there needs to be increased attention to making sure that there is symmetry between the outcomes which are assessed and the goals of the intervention. The specific nature of the intervention program should in large part dictate the outcome measures that are selected. Because resources for research are always limited, investigators should put first priority on measuring those outcomes which are most central to what their intervention is expected to accomplish. For example, some interventions have focused primarily on enhancing social and emotional functioning, but have limited their assessment to measures of IQ. Because there is a substantial interrelationship between the multiple lines of development, differences in IQ may have been found. But these differences are probably much weaker than would have been the case had measures been taken on those behaviors and competencies which were directly targeted by the intervention.

The poet T. S. Elliot once asked, "Where is the knowledge we have lost in information?" Early intervention efficacy researchers are faced with an abundance of information. The challenge with which we are faced is to find the knowledge contained in all that information which will lead to the most effective and efficient services for handicapped, disadvantaged, and at-risk children. To do so will require sustained attention to analyzing and understanding the hundreds of research studies that have already been completed, as well as continuing to conduct high-quality research about which approaches are most efficacious for which subgroups of children.

PRODUCTS

Number in Product List	Title
2	Meta-Analysis Training Manual
16	Low Birthweight Review
30	Efficacy Overview
49	Conducting Longitudinal Research
50	Language Intervention
55	Strain and Smith Reply
56	Dunst and Snyder Reply
58	Selecting Outcome Measures
59	Efficacy of Early Intervention
60	Socially Withdrawn Children
61	Analysis of Special Educat[illegible] Texts
65	The importance of Structu[illegible]
67	Early Intervention Efficacy
73	Environmentally At-Risk Infants
74	Preventive Intervention
79	Learning From Previous Reviews
80	Parent Involvement
81	Age At Start
83	Behavior Disordered Children
93	Studies with At-Risk Children
96	Efficacy Studies
105	Head Start Research
106	Down Syndrome Review
111	Program Intensity and Duration
114	Is Earlier Better?

INVESTIGATION 2:

COST-EFFECTIVENESS ANALYSIS OF HALF-DAY AND FULL-DAY PRESCHOOL SPECIAL EDUCATION PROGRAMS

Policy makers at both the state and local levels often must make decisions regarding educational programs for which existing research provides little guidance. In such cases, it is sensible to ask if their decisions can be informed by data from local programs. Local program data can be obtained at relatively low cost and typically is viewed as more relevant than data from another region or state. This section illustrates the use of local data to examine cost and effectiveness for a decision in which both are important considerations.

The expansion of early childhood education for handicapped and disadvantaged children is an issue that currently receives considerable attention at the state level. In considering how to expand these services, a major consideration is length of the program day. Policy makers can choose between half- and full-day program models. The research literature provides little evidence regarding the relative effectiveness of half- and full-day programs. In addition, it is frequently supposed that half-day programs will cost just about half of what full-day programs will if schools schedule two sessions per day. Thus, there is a strong temptation to conclude that half-day programs can serve almost twice as many students at any given level of funding without any significant reduction in effectiveness.

In this investigation, cost-effectiveness analysis is applied to data from half- and full-day programs for preschool-age handicapped children in neighboring school districts. A simple, nonexperimental design was used to collect data on program effects. Most data on costs were obtained from local budgets. Although our data were collected in conjunction with a larger research project that obtained an extensive and extremely detailed body of data, only relatively modest data collection efforts are required for the policy analysis model presented.

Sample

The sample for study was composed of nine half-day preschool programs serving handicapped children in Sioux City, Iowa, and nine full-day preschool programs serving handicapped children in districts surrounding Sioux City. Most of the children in these programs had their primary handicapping condition identified as mental retardation (51%). The next most frequent primary handicap was communication disorder (28%). The remaining children (21%) had as primary handicaps: physical impairment, learning disability, behavior disorder, and hearing impairment. The distribution of primary handicaps did not differ between half- and full-day programs.

A matched sample of 78 children (39 from each type of program) was selected for the study. Children were matched within primary handicapping condition on three variables: chronological age, months of prior preschool program experience, and developmental level at program entry. There were no statistically significant differences ($p > .10$) between the half- and full-day samples on any of the matching variables. The children were age four at the beginning of the school year in which data were collected. Average prior preschool program attendance was 10.8 months.

Programs

The half- and full-day programs were similar in design, although different in duration. They used the same basic curriculum and provided individual as well as group instruction. Each classroom contained about six children and was staffed by a teacher and an aide. Both programs were conducted five days per week; children in the half-day program attended for about three hours per day. Teachers in the half-day program typically taught two sessions per day. Teachers in both types of programs visited the children's home occasionally to discuss child progress and to suggest home activities for parents to implement with their children as a supplement to classroom activities. In addition, both programs employed therapists to provide individual motor and speech/language instruction and to help teachers implement appropriate activities in these areas more generally.

Data

The program's education effects were assessed by administering tests at the beginning and end of the school year. These tests were The Minnesota Child Development Inventory (MCDI) (Ireton & Thywing, 1974), and the Early Childhood Continuum of Assessment, Programming, Evaluation, and Resources (CAPER) (Carran, 1983). The MCDI is a standardized measure of children's development based on the mother's observations. The CAPER is a criterion-referenced measure of children's master of skills. Together, these tests provide information regarding development in cognitive, language, motor, social, and self-help domains.

Cost data were collected and computed using a modification of the system described by Taylor, White, and Pezzino (1984). In the present study, however, time tracking data were based on a multiple sampling of intervenors' activities rather than on one-time post hoc estimates of time by activity for a "typical" week. Cost data were collected from three sources. First, over the course of the program year, all intervenors provided four weeks worth of time tracking data which documented how much time they spent by activity: direct instruction, preparation, travel, parent contact, testing, inservice, and administration. Second, school district personnel were interviewed in person to obtain cost data for personnel, equipment, facilities, and transportation. School district personnel often provided us with copies of budget documents as a result of the interviews. Third, parents provided a one-week time tracking sample documenting the amounts of time they spent in program-related activities. These activities included meeting with program staff, transporting children to and from programs, and implementing educational programs at home with their children.

Cost Estimation

The program's costs were analyzed by estimating both costs of the program to the school districts and "donations." Costs to school districts were subdivided into personnel and non-personnel categories. Personnel costs consist of salaries, benefits, and the employer's share of social security taxes. Non-personnel costs encompass facilities, equipment and facilities, children's transportation, and reimbursed travel costs for itinerant staff. Donations is the category for resources used by programs, but not purchased by them. They include classroom volunteers and parents' contributions of materials and time.

Costs to School Districts

Personnel costs were calculated for all persons involved with the preschool programs. This includes not only the teachers, aides, and therapists providing direct services, but the support personnel as well. Support personnel include administrative staff (principals, bookkeepers, secretaries), food service employees, bus drivers, and teachers providing specific activities (art, music, physical education). In most cases, the cost of support personnel attributable to any particular classroom can only be estimated roughly. For example, a classroom's support service cost could be estimated based on their percentage of the total school population. The error introduced by such estimation is likely to be quite limited because only a small percentage of cost is involved. However, care must be taken to determine which support services are received. For example, preschool programs may not be provided with bus service or may not have art and music instructors.

Non-personnel costs can be divided into the categories of supplies, equipment, and facilities. These are progressively more difficult to estimate. By supplies, we mean items that are purchased and consumed regularly, whether they be paper and pencils, bus fuel, utilities, or insurance. The cost of supplies is readily determined from school budgets and easily apportioned into the programs studied.

Equipment costs fall somewhere in between supplies and facilities in their characteristics, since equipment may last several years before having to be replaced. However, if the replacement rate is fairly constant across programs and over time, it may be sufficiently accurate to estimate equipment costs from current expenditures. Obviously, this would not be appropriate for programs requiring unusually expensive and durable equipment. We estimated equipment costs based on inventories for each classroom and estimated price, depreciation, and "salvage" value for each item.

Facility costs were estimated from average local rental values for similar facilities. Often, this procedure will not be satisfactory for decision makers. Typically, there are disagreements regarding what facilities are "similar" to schools. Moreover, space for program expansion may vary greatly in cost or be a constraint rather than a variable. For example, there may be a certain number of unused classrooms that are available at practically no cost. Building on a few extra classrooms might cost less per classroom than if new property must be acquired and entire new buildings constructed. Decision makers must take a comprehensive view of facility's costs that goes beyond estimates based on the few classrooms studied.

Donations

Donations include all goods and services that have costs to the community (local, state, or national), but have no costs or reduced costs to the programs. Donations are not necessarily voluntary gifts. Programs may seek to reduce their own costs by shifting them to other public programs, levels of government, or even to private citizens. From the perspective of the community as a whole, analyses that omit these costs can be misleading. Two of the most common donations to preschool programs are the time of parents and non-parent volunteers. Although preschool programs do not pay for this time, there is often a significant cost associated with it. From a community perspective, the time of parents and other volunteers is not free because they could be doing something else worthwhile.

Parent time can represent a significant percentage of the cost of preschool special education programs. Often, programs stress parent involvement and encourage parents to conduct program-related activities at home. In addition, parents sometimes must provide transportation for their children to attend a program. Parents are likely to view most of this time as a cost of obtaining the preschool program for their child. Time costs may be especially important for parents of handicapped children if their children require greater financial and time resources generally.

The cost of parent time depends on individual's opportunity costs; that is, the alternatives they forgo in order to participate in a program's activities. The opportunity cost to parents may be less time on the job, working on household tasks, caring for another child, or in leisure activities. In none of these cases is time taken away free; it is difficult to value, however. We calculated two estimates that we believe are likely to bracket parent costs. The first estimate valued their time at $4.38 per hour (salary plus benefits), the cost of a paraprofessional. Most of the parents could probably find employment at this rate. It has the added advantage that it also provides an estimate of the minimum cost of replacing parent inputs to the program, although paraprofessionals are obviously not perfect replacements for parent time. The second estimate valued parents' time at $10 per hour, which is roughly the average compensation earned by fathers of children in this study. Transportation costs were estimated based on time cost plus 23¢ per mile for expenses.

As with parent time, valuation of non-parent volunteer time depends on the volunteer's alternatives. Where people are truly volunteers and would prefer to donate their time than to engage in other activities, cost is zero. The analyst must be careful, however, to consider the possibility that other influences are involved, as may be the case when programs use "student volunteers" who receive school credit. In addition, if the use of volunteers is not feasible in a large-scale program expansion, estimates for the cost of similar non-volunteer services should be included. The value of non-parent volunteer time was estimated to be $4.38 per hour. As with parent time, it provides a lower-bound estimate of opportunity cost and a convenient measure of the minimum cost to the programs of paid workers.

With all categories, care must be taken that cost comparisons are not misleading because circumstances unrelated to basic program differences make costs higher for one program than the other. For example, by chance, one program might have a few children who require extraordinarily expensive equipment or services for special conditions. One program might have much higher transportation costs because it is in an area with lower population density. In some cases, it will be difficult or even impossible to make things comparable. For example, costs may be higher because a program has better facilities or more experienced teachers. Such program differences may increase effectiveness as well as cost. In these cases, the best approach is to calculate a range of costs under different assumptions.

Results

No differences in educational effectiveness between half- and full-day programs were found. Multiple regression analysis was used to test the effects of program type on posttest score, controlling for pretest score. Regressions were performed for summary scales and subscales of both instruments used (CAPER and MCDI). No statistically significant effects of program type on posttest scores were found overall or in any specific area of development.

Children in both half- and full-day programs showed significant gains in their cognitive and language development scores. Gains in these domains are especially important since 79% of the children studied had mental retardation or communication disorder as their primary handicapping condition. CAPER cognitive and expressive language scores from the beginning and end of the program year (September to May) are reported in Table 11. The magnitudes of these gains compare favorably with those

Table 11

Mean CAPER Cognitive and Expressive Language Pre- and Posttest Scores (Standard Deviation)

Measure	All	Full-Day	Half-Day
Cognitive			
Post	80.01 (22.33) n = 78	80.31 (22.49) n = 39	79.71 (22.47) n = 39
Pre	68.63 (20.44) n = 76	68.74 (20.68) n = 39	68.51 (20.47) n = 37
Expressive Language			
Post	72.06 (24.51) n = 68	73.03 (24.53) n = 36	70.97 (24.84) n = 32
Pre	57.76 (23.59) n = 67	61.31 (24.25) n = 36	53.65 (22.73) n = 31

All pre-post differences are significant ($p < .01$)

No full-day vs. half-day differences are significant ($p < .01$)

produced by other preschool programs for children with handicaps. If these gains represent program effects, then both programs would appear to be successful.

We have access to detailed data on instructional services and other resources used in the two types of programs. This enabled us to estimate cost per child for each program based on the amount of resources devoted to each child. For comparability, costs were estimated using a uniform set of resource prices representing averages for the school districts studies. For example, average teacher costs at each level of experience and qualifications were used. The resulting cost estimates for both programs are presented in Table 12. The school district costs for the half-day program are about 63% of the costs of the full-day program.

Table 12

Costs of School Districts (Per Child) for Half- and Full-Day Preschool Programs

Cost Variable	Half-Day	Full-Day
Personnel	$4,339	$6,882
Facilities	791	1,198
Other	210	337
TOTAL	**$5,340**	**$8,417**

Personnel costs are the largest portion of costs. Personnel costs for the half-day programs are only about half as much as for the full-day programs, with one exception. The time tracking data indicate that half-day students received about 2/3 the amount of motor and language therapist and other specialized service personnel time received by full-day students. Thus, in this category, half-day costs substantially exceeded half the full-day. Non-personnel costs for the half-day programs were also more than half the amount for full-day. For the most part, this is because some items, like transportation cost, do not vary with length of day. However, the half-day costs are higher to some extent because their costs for administration and facilities are higher for reasons unrelated to the length of day.

The estimated costs of donations are reported in Tables 13 and 14, along with school districts' costs. Donations included parents involvement in the program, parent-provided transportation, and non-parents volunteer. By far the most important donation was parent involvement. Parents reported spending an average of 542 hours per child over the school year working with their children at home and meeting with teachers. At $4.38 per hour, the value of this time is nearly half the budgeted cost of the half-day program. At $10, it exceeds the budgeted cost of the half-day

Table 13

Value of "Donations" (Cost Per Child) for Half- and Full-Day Preschool Programs With Time Values at $4.38 Per Hour

Cost Variable	Half-Day	Full-Day
Donations		
Parent Involvement	$2,570	$2,176
Parent Transportation	113	269
Other Volunteers	59	21
School Districts' Cost	5,340	8,417
TOTAL	**$8,082**	**$10,883**

Table 14

Value of "Donations" (Cost Per Child) for Half- and Full-Day Preschool Programs With Time Valued at $10 Per Hour

Cost Variable	Half-Day	Full-Day
Donations		
Parent Involvement	$ 5,868	$ 4,968
Parent Transportation	211	503
Other Volunteers[a]	59	21
School Districts' Cost	5,340	8,417
TOTAL	**$11,478**	**$13,909**

[a]Valued at $4.38 per hour.

program. The cost of parent involvement was greater for the half-day program because half-day program parents spent more time working at home with their children.

Parent transportation costs and non-parent volunteer time proved to be negligible for the programs we examined. Almost all children rode school buses, and few volunteers were used. Differences in these costs between the two types of programs were minor and unrelated to the basic differences in program. Other preschool programs might be substantially different in regard to transportation and volunteers, however, and it would be .mprudent to ignore their costs generally.

When donations, primarily parent involvement, are accounted for, the relative costs of half- and full-day programs change somewhat. The half-day programs are roughly 75% to 85% of the cost of full-day programs, depending on whether time is valued at $4.38 per hour or $10 per hour. the increase in the estimated relative cost of half-day programs is only partially due to the greater parent involvement in half-day programs. Even if donations of the two types of programs had been equal, adding them in would increase the ratio of half- to full-day costs.

We have considered both educational effectiveness and community-wide cost, but we have not provided a complete picture of the programs and their consequences. In particular, we did not measure all of the benefits. Both types of programs provide child care. The full-day programs provided 3-2/3 hours more child care per day than did the half-day programs. The value of these extra hours of child care is difficult to estimate. We have no information on how much the parents might have been willing to pay. The conventional wisdom is that parents of handicapped children have difficulty obtaining child care, and desire more than they can obtain at the going rate for child care. To investigate the potential importance of the child care benefit, we tried two plausible estimates. If this time is valued at $2 per hour, its value would be about $1,300 over a school year; at $5, its value would be $3,300.

Implications

The study presented in this investigation demonstrates the potential usefulness of cost-effectiveness analysis as a tool for state and local policy makers. Our cost estimates suggest that simply assuming half-day preschool programs for handicapped children were half the cost of full-day programs would have been misleading. Based on the costs paid by the schools we studied, half-day preschool programs could only serve about 1.6 times the number served by half-day programs with a given budget. Policy makers who instituted half-day programs under the 50% cost assumption might be surprised by short fall in numbers served and the eventual cost of serving the entire eligible population.

Cost-effectiveness analysis is also a broadening process that expanded the policy maker's horizon beyond program budgets. All costs to the community must be considered. In this case, parents' efforts to implement activities at home as part of their child's program entailed considerable costs. These costs were substantially higher for the half-day program. It may be that half-day programs maintain effectiveness while reducing cost to the schools partly because they shift some of the educational costs to the parents. All effects must be considered. Educational effectiveness is not the only potential benefit from the preschool programs, though it sometimes seems to be the only one recognized by schools. The value of the additional child care provided by the full-day programs is potentially a significant amount relative to cost. Although the estimates for parent time cost and value of child care are hardly accurate enough to simply add up, they indicate important consequences that will contribute to satisfaction with the program type chosen.

The desirability and importance of the costs of parents and child care provision depends on the goals and values. School policy makers may believe that parents <u>should</u> bear a large portion of the cost for their child's education. Most people would agree that parents <u>should</u> be involved in their child's education, though how much and in what way will be debated. Similarly, the importance of full-day care as a means to allow women to participate in the labor office is a matter of values. Thus, a significant by-product of cost-effectiveness analysis may be a clarification of goals and values.

This investigation also demonstrated the potential ease with which a cost-effectiveness analysis can be conducted. The study described was nonexperimental, and most of the data required was readily available from the schools without special efforts. Typically, some tests of child progress are administered annually. Most of the detailed staff time and program cost data was not needed to produce reasonably accurate cost estimates. We did not need time tracking data to determine that half-day programs spent a greater portion of their funds on individual therapy than did full-day programs. That could be determined from school records. Non-personnel costs are relatively small, and except for facility's cost, easily estimated. Facility's cost estimates probably rely more on judgement about options than on data for the programs examined. Facility planning and cost per child varies greatly depending on specific circumstances like excess capacity, the use of portable or temporary structures, and demographic trends.

The importance of parents' time costs add one difficulty to the analysis. Costs to parent could not be determined from school records. Policy makers seeking such information would have to conduct parent surveys to obtain this information. Such surveys are relatively easy and low-cost to conduct. However, we have some serious concerns about the data our own survey yielded.

The level of program-related time use reported by parents seems remarkably high; an average of two hours per day. Unfortunately, the question we asked parents about the amount of time spent implementing a program at home was ambiguous. Specific activities were not reported. The ambiguity is difficult to avoid, because some parents may view themselves as implementing program activities in almost everything they do with their children.

Another concern with the parent time data is that the higher time cost of the half-day programs may not be a characteristic of half-day programs. In our judgement, it seems reasonable that parents and staf would seek to supplement the shorter day with more time spent at home, and, thus, are more available for parents to work w h them. The difference could have been incidental, however, in choosing the matched sample, child and program characteristics were considered, parent characteristics were not. Thus, we have less confidence in making influences about parents.

The specific results and implications discussed above are of limited interest beyond the school districts we studied, but the policy analysis model demonstrated should be of more general interest. At the school district or regional level, this type of cost-effectiveness analysis might significantly improve the information available for program planning. Familiarity with a good introduction to economic evaluation (Levin, 1983; Thompson, 1980) would provide sufficient expertise to conduct such an analysis. The necessary data can be obtained with relatively modest additional efforts. Although "blind" reliance on nonexperimental program comparisons might easily lead to incorrect inferences, we believe that common sense and knowledge about the program would enable policy makers to adjust their conclusions appropriately. At the state level, decisions have more wide-spread consequences, and resources for research are more extensive. Thus, a stronger research design and a larger, more diverse sample of programs, children, and families. The same basic model of cost-effectiveness analysis can be employed, however.

Products

Number in Product List	Title
37	Cost-Effectiveness Analysis
87	Half- and Full-Day Programs

INVESTIGATION 3:

LONGITUDINAL STUDY OF EARLY INTERVENTION WITH HEARING IMPAIRED CHILDREN

Introduction

This research project was done to investigate the long-term impact of early intervention on hearing impaired children. This study will be reported in five sections. The first section will include a pr blem statement and a discussion of the basic research questions. The next section will present a review of literature on early intervention programs for hearing impaired children. The third section will be a discussion of research methods and procedures, and will include a description of the research model, sample selection, selection and development of measures, testing, and data analyses procedures. The fourth section will present the results of the study and a discussion. Finally, implications of the research project and recommendations for future research will be discussed.

Problem Statement and Research Questions

Problem Statement

Prelingual hearing impairment affects a relatively large number of children each year (approximately 1 in 1,500 births). A hearing impairment is particularly devastating to the child during the first few years of life when language acquisition occurs (Clark & Watkins, 1978; Northern & Downs, 1974). To ameliorate this serious problem, many early intervention programs for hearing impaired children have been established throughout the country during the past few years.

One of the most successful and widely disseminated of these programs is the SKI*HI program, which has been approved by the Joint Dissemination Review Panel for national dissemination. The SKI*HI model has been used with over 6,000 children during the last nine years, and is currently being used with over 1,500 children in) sites throughout the country.

The validation of SKI*HI for national dissemination was based on data from a quasi-experimental design where 33 children who received treatment prior to 30 months of age were compared to 27 children who were identified after 30 months of age and had not yet received treatment. Comparability of the two groups was established on degree and type of hearing loss, age, and other demographic variables. Comparison of the two groups demonstrated that the group with earlier intervention was significantly better on use of residual hearing, auditory development, receptive and expressive language, and parental involvement with their child's early education (Clark, 1979).

Unfortunately, there are no data on these hearing impaired children to determine the long-term effects of early home programming on them. Since the untested assumption upon which these intervention programs are operating is that early gains will be maintained and will impact on other areas, there is vital need for research to be conducted on the long-term impact of early intervention on hearing impaired children.

Any serious effort to examine the cost-effectiveness of early intervention for the hearing impaired would have to consider the effect of these long-term outcomes. Therefore, since such longitudinal data are lacking, it is impossible to fully justify continuance of the national and local resources being used in early intervention programs for the hearing impaired.

Perhaps knowledge of the long-term impact of early intervention is most important for hearing impaired children and their families. Longitudinal data are needed to help these deaf youngsters and their families know if they are receiving services that provide positive impact on their lives beyond treatment time.

Research Questions

The main purpose of this study was to investigate the long-term impact of home intervention on hearing impaired children. The basic research question that emerged was: "Do hearing impaired children who received home intervention earlier in their lives perform better than hearing impaired children who did not receive home intervention earlier in their lives on measures of language, academic achievement, and psycho-social behaviors?" In addition to this most basic question, two other important questions emerged: "Do children who receive home intervention before age 2-1/2 perform better than children who did not receive intervention until after age 2-1/2 on measures of language, academic achievement, and psycho-social behaviors?" "Do children who received no home intervention but attended preschool perform better than children who did not receive home intervention and did not attend preschool on measures of language, academic achievement, and psycho-social behaviors?"

Review of Literature

During the last 15 years, there have been only a very few studies done on the long-term impact of early intervention for hearing impaired children. The large majority of these studies have investigated the effects of child-oriented, center-based programs (preschools or nursery schools) on hearing impaired children. Only a very few studies have investigated the impact of parent programs on hearing impaired children and none of these studies have look specifically at home (versus center-based) parent programs such as the SKI*HI program.

This review of literature contains first a discussion of the studies that have been done on the long-term effects of center-based, child-oriented programs on hearing impaired children. Next, a discussion of studies done on the impact of parent programs on hearing impaired children will be presented.

Child-Oriented Intervention with Hearing Impaired Children

The studies on the long-term effects of nursery and preschool programs on young hearing impaired children are inconclusive. Research done primarily during the 1960s did not yield conclusive evidence for positive sustained impact of preschool intervention. Craig (1964) administered comprehensive batteries of speechreading and reading tests to 151 children at the Western Pennsylvania School for the Deaf and the American School for the Deaf (Connecticut) who had attended preschool earlier in their lives. He also tested a control group of 101 children from the same institutions who had not attended preschool. He found no statistically significant differences between the experimental and control groups after the children had been in the primary grades for three to four years. Similar results were found by Phillips (1963), who tested 9-year-old severely and profoundly hearing impaired children from eastern United States schools for the deaf, including the Lexington School (New York) and the American School for the Deaf (Connecticut). No statistically significant differences between the experimental preschool group and the control no-preschool group were found on measures of arithmetic achievement, language achievement, and socialization.

Vernon and Koh (1970) compared children who had experienced three years of oral preschool (John Tracy Preschool Program) to children with no preschool who had: (a) oral home environments, and (b) manual communication home environments. Groups were matched on age and IQ. There were 23 subjects in the experimental group and 23 subjects in each of the two control groups. Participation in preschool did not seem to be the determining factor of later academic achievement advantages. At age 18, children who experienced an oral preschool program did not score statistically significantly higher than the no-preschool children from oral home environments on the Stanford Achievement Test. However, the experimental preschool children scored statistically significantly lower than the no-preschool children from manual communication home environments on the Stanford subtests of paragraph meaning and reading.

Balow and Brill (1975) did a follow-up study of the Vernon and Koh research. They studied 264 John Tracey Preschool graduates who were attending the California School for the Deaf at Riverside. This sample was considerably larger than the 23 subjects used in the Vernon and Koh study. The Tracy graduates were compared to other students at the Riverside School who had not had preschool programming. The John Tracy graduates scored statistically significantly higher on the Weschler Adult Test than the control group. An analysis of covariance showed that a statistically significant difference in achievement remained when the effects of IQ were controlled.

Moores, Weiss, and Goodwin (1978) conducted a six-year longitudinal study on preschool programs for deaf children. Subjects included hearing impaired children who had attended seven different preschools which emphasized different communication methodologies. The hearing impaired children were shown to have almost identical scores to hearing children in the standardization sample of the Illinois Test of Psycholinguistic Abilities and the reading subtest of the Metropolitan Achievement Test Primer Battery. However, communication success as measured by the Receptive Communication Scale (a tool developed by the research team) depended on the type of preschool program in which the children had participated. Children scored highest who had been in speechreading and signing preschool programs. These children were followed by those who had experienced speech and finger spelling preschool programs; these were followed by children who had been in preschool programs utilizing speech and audition. Children scored lowest who had been in programs utilizing auditory receptive communication only.

Intervention Directed to Parents of Hearing Impaired Children

Most of the studies done on the long-term impact of parent intervention on hearing impaired children have involved center-based programs of parental instruction. Parents have received training in clinic settings or demonstration home settings in how to provide meaningful language stimulation for their hearing impaired children.

Lowell (1967) studied hearing impaired children whose parents had received training in a demonstration home while the children were 1 to 3 years of age. After the parents completed the program, the language growth of the children was monitored. Using the modified Boone Scales of Linguistic Encoding and Decoding, two groups of experimental children showed statistically significant gains for months after the program was completed. Two control groups of children who had been enrolled in traditional nursery school and whose parents did <u>not</u> receive instruction did not show statistically significant improvement.

Ewing and Ewing (1964) found that deaf children whose parents had received center-based guidance were linguistically superior to children whose parents did not have the benefit of such training. Gains for the experimental children were statistically significantly greater than for the control children in the articulation of spoken English, use of colloquial English, vocabulary, spontaneity of vocalization, and variety of pitch and intonation during the first three to four years of elementary school. Parents who had received training were judged to be more cooperative with the school in continuing the child's home language training than control parents. Teacher interviews were utilized to obtain this information.

Horton (1976) studied six hearing impaired second grade children whose parents had received training in the Mama Lere Demonstration Home. The Children were 0 to 3 years at the time of intervention. Two control groups were also studied: (a) five hearing impaired second grade children whose parents had not received instruction (but who had been fit with hearing aids at a median age of 4 years), and (b) six hearing second grade children who were in the same school as the intervention group. The severity of hearing loss for the experimental and control hearing impaired groups was not statistically different. Fifty consecutive utterances produced by the children in each of the above groups were analyzed according to Lee's Developmental Syntax Types. The findings revealed that the language competence of the experimental group was not statistically different from the hearing control group. However, there were statistically significant differences between the experimental and no-intervention hearing impaired groups favoring the experimental group. For example, the intervention group produced, on an average, 75% of their utterances on the sentence level, compared to only 32% for the no-intervention control group. Only 8% of the intervention group's utterances were of the noun type (immature construction) compared to 19% noun-type construction usage in the control group. In the intervention group, 79% of the utterances were mature verbal constructions, while only 19% of the utterances in the control group were of this type.

In another study, Horton (1976) compared six hearing impaired second grade children whose parents had received training in a demonstration home to 53 hearing second grade children. The Metropolitan Achievement Test was given to both groups. The mean percentile ranks for both groups were virtually equivalent in the area of reading. The hearing impaired children scored slightly lower on the math subtest.

Studies on the long-term effects of home visit programs for parents of hearing impaired children (such as the SKI*HI Program) have not been reported in the literature during the last 15 years. Lack of research on such programming is most unfortunate because:

1. Parent-oriented programs have been shown to have longer lasting positive effects on children than child-oriented programs without parental participation (Bronfenbrenner, 1974).

2. Home programs are claimed to be superior to clinic or demonstration programs because:

   a. The home is the parents' and child's natural environment.

b. Intervention in the home allows for utilization of natural prime times for language stimulation (such as bath time, getting child dressed, etc.) (Clark & Watkins, 1978; Shearer & Shearer, 1976).

c. In home programming, parents do not need to be dressed and go out to a center. Nearly 100% attendance was reported by Watkins (1971) in the Utah home visit parent infant program.

d. Studies done on home visit parent programs for other handicaps such as visual impairment and mental retardation reveal that these programs are more cost-effective than center-based parent programs (Macy & Carter, 1980).

It is evident, then, that research on the long-term effects of home visits parent programs (such as SKI*HI) is greatly needed.

Summary

The scanty research available on long-term effects of preschool programming for hearing impaired children is inconclusive. Studies done during the 1960s reveal that children who experience preschool do not score higher on academic achievement than control children. However, in later studies, it is shown that children who attended preschool are comparable to hearing controls or superior to hearing impaired controls on some academic measures. Some research indicates that the types of preschool program (favoring sign language utilization) may be a more important indicator of later academic success than participation in preschool per se.

Studies have been done on the long-term impact of center-based programs for parents of young hearing impaired children. Children whose parents have been in these programs show greater language competence and academic achievement in the first few primary grades than children whose parents have not participated in such programs. Research on the long-term impact of home visit parent programs (such as the SKI*HI model) is not available.

Methods and Procedures

Research Model

The research model used in the study was a longitudinal research design which studied the relationship over time of home intervention to language, academic achievement, and psycho-social performance of hearing impaired children. The design was similar to the Stanley and Campbell ex post facto design because treatment administration (home intervention) had already occurred and current performance levels were assessed. The general statistical model used was analysis of covariance and multiple comparison procedures. Specific measures were taken to control for threats to internal and external validity that were inherent in the research model as shown in Table 15.

Sample Selection

In order to examine the issues of home intervention vs. no-home intervention, early vs. late home intervention, and preschool vs. no preschool, four research groups were selected.

Group 1: Children who had a home intervention program (SKI*HI) before age 2-1/2 and who attended preschool.

Group 2: Children who had a home intervention program (SKI*HI) after age 2-1/2 and who attended preschool.

Group 3: Children who did not receive home intervention ages 0-5 but who did attend preschool.

Group 4: Children who did not receive home intervention ages 0-5 and who did not attend preschool.

The subjects for Groups 1 and 2 were children who participated in a study done by Clark and Covert (Clark, 1979). In this study, 33 children who had an average 9 months of treatment before age 2-1/2 were compared to 27 children who had no treatment until age 2-1/2.

In this current research study, the early treatment children in the Clark and Covert study were matched to the late treatment children in that study on the

Table 15

Controls for Threats to Internal and External Validity

INTERNAL VALIDITY	
1. History:	a. Control group used. (Likely same historical factors operated on experimental and control children so history non-differential). b. Factors suspected of differential influence were either matched or, if highly correlated with dependent variables, were treated as covariates: 1) amount and type of preschool 2) amount and type of therapy
2. Maturation	a. Control group used. (Likely same maturation factors in operation for control and experimental children so maturation non-differential). b. Factors suspected of differential influence were either matched or, if highly correlated with dependent variables, were treated as covariates: 1) age 2) bouts with middle ear infections
3. Testing	a. There were no pretest effects on child scores.
4. Instrumentation	a. Diagnosticians "blind" to group membership of children. b. Fixed instrument used. c. All tests given during same two-week period.
5. Regression	a. Groups were not selected on basis of extreme prescores.
6. Mortality	a. Attrition rates were slightly different between Groups 1 and 2 (30% and 15%, respectively). However, no reason to suspect differential attrition (such as children moving out of state).
7. Differential Selection	a. Factors suspected of being different for groups matched or, if highly correlated with dependent variable, treated as covariates: 1) hearing loss 2) other handicaps 3) current school placement 4) index of social position (occupation and education of parents) 5) age of parents 6) number of parents 7) hearing status of parents 8) amount of treatment (for Groups 1 and 2) 9) lapsed time since treatment (for Groups 1 and 2)
EXTERNAL VALIDITY	
1. Interaction of Testing and Treatment	a. Pretests were not given so possibility did not exist of subjects' responses (as a result of pretest effects) being non-generalizable to untested populations.
2. Interaction of Testing and Treatment	a. Study purports generalizability of results only to populations of hearing impaired children with characteristics similar to children in this study.
3. Reactive Effects	a. Children too young during treatment to be subject to reactive effects. b. At testing time, all children simply informed they were to participate in some activities to see how well they were doing in school. Therefore, John Henry Effect (subject attempt to prove or disprove treatment theory) not likely a problem. c. Since hearing impaired children are regularly tested, Hawthorne effects (improved or worsened performance as a result of "test taking") minimized.
4. Multiple Treatment Interference	a Series of treatments were not given to possibility did not exist of one treatment distorting another treatment, making test results of any one treatment ungeneralizable to other treatment applications.
5. Generalizability to Other Treatment	a. Treatment given by different parent advisors.
6. Generalizability to Other Measures	a. Multiple measures used.
7. Generalizability to Other Times (beyond immediate posttreatment)	Since an important purpose of this study was to determine generalizability of treatment effect (beyond immediate posttreatment time), this was not a threat. However, study purports generalizability of treatment effects only to those times beyond treatment of children included in this study.

variables of hearing loss, age, existence of other handicaps, and preschool attendance. Attrition attributed to parent refusal to include the children in the study, out-of-state moves, and unsuccessful matching, resulted in a final N of 23 in both groups. These children had received treatment earlier in their lives in the form of the SKI*HI model. This model contained a child identification component, regular weekly home visits by a professional to the child's home, and medical, audiological, and psychological ancillary services. At the time of the current study, children in Groups 1 and 2 were in 31 schools scattered throughout the State of Utah.

Children in Groups 3 and 4 were selected from a pool of sites that did not have a home intervention program in existence long enough to yield "graduates" currently 6-13 years of age. Four sites were selected from this pool:

1. Tennessee School for the Deaf, Knoxville, Tennessee.
2. Alabama Institute for the Deaf, Birmingham, Alabama.
3. Memphis Oral School for the Deaf, Memphis, Tennessee.
4. Local school districts in Utah and Idaho (Cache County School District, Logan, Utah; Logan City School District, Logan, Utah; Preston School District, Preston, Idaho.

Children from these sites were carefully matched with the children in Groups 1 and 2 on four variables listed above. A total N of 96 (23 in each of the four groups) resulted.

Selection and Development of Measures

Fifteen SKI*HI Model impact areas were defined, and a group of professionals who work with the model were asked to rate how the impact areas are directly affected by the intervention program. They related nine impact areas as most important. Outcome variables for these nine impact areas were then defined and included child receptive and expressive language, communication, academic achievement, speech, social-emotional adjustment and self-concept, parent attitudes, communication, and hearing aid management.

Potential instruments to measure these outcome variables were next carefully researched. It was determined that appropriate measures were not available for parent attitudes, communication, and hearing aid and communication management. So instruments to measure these variables were developed specifically for this study. Commercially available measures were obtained for the other outcome variables.

Table 16 contains a list of the outcome variables and the instruments used to measure those variables.

Table 16

Outcome Variables and Measures

Outcome Variable	Measures
1. Receptive language	1. Carrow Test of Auditory Comprehension of Language
	2. Peabody Picture Vocabulary Test
2. Express language	3. Lee's Developmental Sentences Scoring
	4. Expressive One-Word Picture Vocabulary Test (Gardner)
3. Communication	5. Communication Inventory and Teacher Rating (developed for this study)
4. Academic achievement	6. Woodcock-Johnson Psycho-educational Battery: Part II. Tests of Achievement (Reading, Math, Written Language)
5. Speech	7. Arizona Articulation Proficiency Test
6. Social-emotional adjustment, self-concept	8. Meadow-Kendall Social-Emotional Assessment Inventory for Deaf Students
7. Parent attitudes	9. SKI*HI Parent Attitude Scale (developed for this study)
8. Parent management of hearing aid	10. Parent Questionnaire (developed for this study)
9. Parent communication	11. Parent Questionnaire (developed for this study)

Testing

Clearance was obtained to test the human subjects in this study from the Utah State University Institutional Review Board. Clearance was also obtained to conduct child testing from the parents of each child in the study and from the administrator of the school each child was currently attending.

Eight diagnosticians were recruited and were given three days of training at Utah State University to administer all the measures. The diagnosticians were graduate students in Communicative Disorders and Special Education at Utah State University who knew sign langauge and had experience or course work in psychometrics.

The diagnosticians conducted child testing at 37 schools in Utah, Idaho, Alabama, and Tennessee. All testing was done in a two-week period of time. Each child received two 1-1/2 hour test sessions over two days.

Parent Attitude Scales and Parent Questionnaires were sent to all parents of the children in the study. Eight-four percent of the parent attitude scales were returned, and all information was returned on the parent questionnaires.

Data Analysis

Tests were scored for use in the data analysis in three major ways. First, commercially available tests were scored according to test protocols. Second, scoring procedures for instruments specifically designed for this study were devised, and these measures were then scored. Finally, videotaped language sample and articulation tests were transcribed and scored according to instruction manuals. Because scoring of the Arizona articulation test required some subjective judgments as to the correctness of sound production, an interrater reliability study was done on 15 of the children in the research project. This study yielded a reliability coefficient of .96.

In order to answer the basic research questions about differences between groups of children who received home intervention vs. no home intervention, early home intervention vs. late, and preschool vs. no preschool, analyses of covariance and multiple comparison procedures were used to determine these group differences. In addition, effect sizes to determine educational significance of the research data were determined.

In order to perform these analyses, test scores obtained above on 36 dependent variables were entered onto computer coding sheets, along with 22 potential covariates. The potential covariates were obtained from school record information, the Parent Questionnaire, and from an analysis of treatment data on children in Groups 1 and 2 who participated in the Clark and Covert 1979 study. Coded data were then transferred to computer disk for analysis.

Potential covariates were then correlated with the dependent variables. Six covariates, with most of the dependent variables at a level of .3 or higher, were subsequently selected as the final covariates to be controlled in the analyses. They included hearing loss, age, existence and severity of other handicaps, age of mother, Hollingshead and Redlich (1957) Index of Social Position (derived from the parent education and occupation items on the Parent Questionnaire), and number of childhood middle ear infections. Multiple Rs were obtained to determine the relationship of each dependent variable to the covariates collectively. The larger the relationship (multiple R), the more need was evidenced to covary on the six factors.

Next, overall differences among groups were obtained by performing a univariate analysis of covariance with multiple covariates. In this analysis, group differences were determined for each dependent variable while covarying on the six covariates. A multivariate analysis of covariance was also performed. Dependent variables were categorized into the four logical groups of language/communication, academic achievement, psycho-social behaviors, and parent attitudes. Group differences were then obtained for each dependent variable category while covarying on the six covariates. This measure provided additional indication that group differences existed and confirmed group differences for individual measures within dependent variable categories.

Analyses of specific group differences were next performed on comparisons that were considered of primary importance because they dealt with the issues of home intervention vs. no home intervention, early vs. late home intervention, and preschool vs. no preschool. In addition, analyses of specific group differences were obtained for other comparisons of secondary interest which compared one of the two home intervention groups to one of the two no-home intervention groups. Multiple t-tests were performed on pairwise contrasts of all dependent variables that were statistically significant for the four research groups. These comparisons included Group 1 vs. 2, 3 vs. 4, 1 vs. 3, 2 vs. 3, 1 vs. 4, and 2 vs. 4. In addition, planned orthogonal contrasts were used to compare combinations of group means with other group means. These comparisons included Groups 1 and 2 vs. 3, 1 and 2 vs. 3 and 4, and 1 and 2 vs. 4. It is best if the comparisons in planned orthogonal contrasts are orthogonal to each other (independent of each other). It was determined that the Group 1 and 2 vs. 3 and 4 comparisons were orthogonal. However, the other two contrasts were not orthogonal. This was not considered serious, however, since the alpha level was raised only slightly (increased chance of Type I error).

Finally, it was determined if all primary and secondary comparisons were educational significant. The technique used to obtain this information was effect size analysis. In this analysis, treatment groups were pitted against control groups such as the early home intervention group (treatment) vs. the late home intervention group (control). Effect sizes, or differences between these groups in terms of standard deviation units, were then determined.

Results and Discussion

Covariate Selection and Analyses of Overall Group Mean Differences

Multiple R analysis. Since covarying was to be done on the six covariates collectively, multiple R tests were conducted to determine the relationship of the covariates to each dependent variable. The results of this analysis are in Table 17. Since the covariates account for over 50% of the variance of 11 dependent variables and over 50% of the variance of 24 dependent variables, the need to covary on the six factors is obvious.

Analysis of covariance. In order to determine group mean differences, two analyses of covariance were performed: univariate analysis of covariance with multiple covariates and multivariate analysis of covariance. Results of these analyses are shown in Table 18. Mean scores are listed from highest to lowest, with the group number in parentheses next to the mean. The possible number of points for each dependent variable is also given.

It is important to note on variables 20-23, a larger mean represents a small percent of understood communication. Higher scores on three other variables also indicate poorer performance: variable 16 (where higher scores indicate more consonant errors), variable 30 (where higher percent scores indicate more solitary vs. group play), and variable 35 (where higher scores reveal poorer child attitudes toward school).

Mean differences that are statistically significant at a .1 level are noted with asterisks. This alpha level is not considered too liberal (increased chance of Type I error) because of the following:

1. Higher power values:

	Medium-Sized Difference	Large-Sized Difference
A. Power at .05:	62%	99%
B. Power at .1:	89%	99%

(4 groups 23 subjects/group, 6 covariates)

Table 17

Multiple Rs for Dependent Variable

Dependent Variable	Multiple R
Woodcock Johnson Raw Scores	
Letter/Word Identification	.60
Word Attack	.58
Passage Comprehension	.66
Calculation	.73
Applied Problems	.64
Dictation	.61
Proofing	.61
Peabody Picture Vocabulary Test	
Raw Score	.57
Standard Equivalent Score	.55
Age Equivalent Score	.58
Test of Auditory Comprehension of Language Raw Score	.63
Communication Inventory Raw Score	.59
Communication Rating by Teacher	.49
Expressive One-Word Picture Vocabulary Test Raw Score	.61
Arizona Articulation Proficiency Test	
Raw Score	.74
Consonant Score	.71
Lee's Developmental Sentence Scoring	
Raw	.52
MLU	.56
Parent Attitude Scale	
Total Raw Score	.47
Reactions to Outside Help Raw Score	.44
Anxiety/Guilt Raw Score	.46
Acceptance Raw Score	.41
Meadow-Kendall Social-Emotional Assessment	
Social Adjustment Raw Score	.55
Self-Image Raw Score	.51
Emotional Adjustment Raw Score	.51
Parent Questionnaire Raw Scores	
Time Hearing Aid Worn	52
Time Spent Reading	.21
% Solitary vs. Group Play	.38
Number of Friends	.24
Child's Attitude Toward School	.45
% of Child's Communication Understood by Family	.59
% of Child's Communication Understood by Non-Family	.60
% of Family Communication Understood by Child	.41
% of Non-Family Communication Understood by Child	.40
Time Spent Communicating with Child	.55
Child Behavioral Rating	.23

Results of Univariate Analysis of Covariance with Multiple Covariates
and Multivariate Analysis of Covariance

	Dependent Variables	F-test p value	Mean adjusted scores by group				Possible points for each dependent variable	Group S.D.	MS_e
ACADEMIC ACHIEVEMENT Wilks F = 2.42 p = [illegible]	1) Woodcock-Johnson: Letter/Word Identification	.007*	23.71 (1)	23.60 (2)	17.54 (3)	16.28 (4)	54	9.08	50.40
	2) Word attack	.001*	8.27 (2)	6.36 (1)	4.28 (3)	1.44 (4)	26	5.96	21.72
	3) Passage comprehension	.062*	7.68 (2)	6.47 (1)	5.97 (3)	4.23 (4)	26	4.58	11.92
	4) Calculation	.000*	13.43 (2)	12.90 (1)	8.13 (3)	6.45 (4)	42	6.42	15.14
	5) Applied problems	.076*	17.80 (1)	15.94 (2)	13.18 (3)	11.18 (4)	49	8.61	44.84
	6) Dictation	.065*	11.95 (2)	10.35 (1)	8.32 (3)	6.86 (4)	40	6.82	29.72
	7) Proofing	.083*	4.08 (2)	3.93 (1)	2.76 (3)	1.18 (4)	29	3.97	10.09
LANGUAGE/COMMUNICATION Wilks F = 1.73 p = .006	8) Peabody Picture Vocabulary Test: Raw score	.115*	63.34 (1)	54.76 (2)	49.65 (4)	48.47 (3)	175	21.97	335.68
	9) Standard score equivalent	.073*	52.07 (1)	44.62 (2)	37.11 (3)	35.81 (4)	160	21.35	321.82
	10) Mental age	.120*	67.71 (1)	62.56 (2)	54.02 (4)	53.88 (3)	175	21.99	333.83
	11) TACL	.146*	80.60 (1)	77.69 (2)	72.62 (3)	71.36 (4)	101	14.84	137.08
	12) EOWPVT (Gardner)	.012*	62.91 (2)	61.24 (1)	51.54 (4)	50.71 (3)	110	16.51	167.06
	13) Communication Inventory	.055*	36.24 (1)	33.78 (3)	33.19 (2)	29.91 (4)	40	6.79	30.24
	14) Teacher Rating of Communication Skills	.149*	3.97 (1)	3.48 (2)	3.46 (3)	3.06 (4)	5	1.07	.91
	15) Arizona Articulation: Raw Score	.018*	66.12 (1)	65.07 (2)	58.46 (3)	43.44 (4)	100	29.19	381.30
	16) Consonant Error	.309	32.43 (4)	26.24 (1)	25.23 (3)	24.24 (2)	54.5	17.99	170.82
	17) DSS: Mean Length of Utterance	.485	6.81 (2)	6.23 (4)	6.21 (1)	5.45 (3)	Unlimited[a]	3.23	7.66
	18) Raw Score	.010*	8.14 (1)	5.83 (2)	4.93 (3)	3.53 (4)	14	3.99	11.15
	19) Time Hearing Aid Worn	.000*	3.86 (2)	3.72 (1)	2.98 (3)	2.42 (4)	4 (76-100% of child's waking hours)	1.00	.58
	20) % of Child's Communication Understood by Family	.064*	2.69 (4)	2.22 (3)	2.04 (2)	1.83 (1)	4 (0-24%)	.94	.58
	21) % of Child's Communication Understood by Non-Family	.062*	3.96 (4)	3.38 (3)	2.88 (2)	2.52 (1)	4 (0-24%)	1.17	.91
	22) % of Family Communication Understood by Child	.138	2.53 (4)	2.52 (3)	2.09 (2)	1.91 (1)	4 (0-24%)	.92	.73
	23) % of Non-Family Communication Understood by Child	.002*	3.83 (4)	3.82 (3)	2.95 (2)	2.43 (1)	4 (0-24%)	1.26	1.27
	24) % Time Communicating with Child	.844	3.55 (3)	3.52 (4)	3.45 (2)	3.28 (1)	4 (more than 2 hours)	.86	.66
PARENT ATTITUDE Wilks F = 1.0 p = .167	25) Parent Attitude Scale: Total	.468	105.63 (1)	98.23 (3)	98.01 (4)	96.35 (2)	128	[illegible]	[illegible]
	26) Reactions to Outside Help	.440	22.81 (1)	21.84 (4)	21.06 (2)	20.44 (3)	28	4.83	20.33
	27) Anxiety/Guilt	.402	23.02 (1)	22.91 (4)	21.37 (2)	21.22 (3)	28	4.26	17.15
	28) Acceptance	.319	59.75 (1)	56.58 (3)	53.81 (2)	53.29 (4)	72	9.43	83.39
PSYCHO-SOCIAL BEHAVIORS Wilks F = [illegible] p = .122	29) Meadow-Kendall: Social Adjustment	.031*	75.58 (1)	66.67 (2)	65.65 (3)	62.97 (4)	92	13.37	124.49
	30) Self-Image	.135*	71.74 (1)	63.93 (3)	66.25 (2)	62.99 (4)	92	11.83	106.80
	31) Emotional Adjustment	.313	44.36 (1)	42.88 (2)	[illegible].97 (3)	39.66 (4)	52	7.53	44.44
	32) Time Spent Reading	.409	2.17 (2)	2.07 (4)	2.06 (3)	1.67 (1)	4 (more than 2 hours)	.83	.73
	33) Solitary Play	.223	41.92 (4)	28.85 (2)	28.12 (3)	13.23 (1)	100% (solitary play vs. group play)	27.37	699.69
	34) Number of Friends	.265	8.48 (4)	6.46 (1)	5.05 (3)	2.21 (2)	Unlimited	7.91	63.42
	35) Attitude Towards School	.695	1.42 (2)	1.31 (3)	1.30 (4)	1.22 (1)	1 (worse than most)	.49	.21
	36) Rating of Child's Behavior	.093*	2.46 (1)	2.19 (3)	2.11 (4)	1.91 (2)	1 (better than average)	.57	.32

[a]Unlimited but sentence length of 12 considered very long for child of this age.

2. Effect sizes consistently favoring the treatment groups.

3. Consistency of p values of variables that are highly correlated.

Total group standard deviations and within-group errors (MS_e) are given in the table. Also given are F values for the Wilk's Multivariate Test, along with the significance levels of these F values.

Discussion. As revealed in Table 19, statistically significant differences among groups exist for the majority (67%) of the dependent variables. In addition, when dependent variables are categorized into logical groups, statistically significant group differences exist for three of the four dependent variable categories. These dependent variable category differences confirm the existence of overall group mean differences and the fact that individual dependent variable differences exist within a category.

Analyses of Specific Group Mean Differences

Multiple comparison procedures. In order to determine which specific group mean differences contributed to the overall group mean differences, multiple t-tests were performed on all pairwise contrasts, and planned orthogonal contrasts were performed on group combination contrasts. Results of these analyses are in Table 19 on the following page. All f and t values that are statistically significant at the .10 level are noted with asterisks. Negative t-values for variables 18-21 indicate better performance for the first group in the pairwise comparison since higher scores on these variables are indicative of poorer performance. Negative values for any other t-scores indicate better performance by the second group in the pairwise contrast.

Discussion. Eighty percent of all f and t values favor Groups 1 and/or 2 when compared to groups 3 and/or 4 at levels of statistical significance. This can be seen more specifically in Table 20, which summarizes the percent of f and t values that favor the home intervention children in Groups 1 and 2.

Table 19

Results of Multiple Comparison Procedures

Dependent Variables	Multiple - T Tests: (critical t = 1.68) Group 1 vs. 2	Group 1 vs. 3	Group 1 vs. 4	Group 2 vs. 3	Group 2 vs. 4	Group 3 vs. 4	Planned Orthogonal Contrasts: (critical F = 2.77) Group 1 & 2 vs. 3 & 4	Group 1 & 2 vs. 3	Group 1 & 2 vs. 4
Woodcock-Johnson: Letter/Word Identification	.06	2.95*	3.56*	2.90*	3.50*	.60	21.24*	17.46*	25.29*
Woodcock-Johnson: Word attack	-1.39	1.52	3.59*	2.91*	4.99*	2.07*	21.50*	9.90*	37.39*
Woodcock-Johnson: Passage comprehension	-1.19	.49	2.20*	1.68*	3.30*	1.71*	7.70*	2.41	15.90*
Woodcock-Johnson: Calculation	-.46	4.14*	5.61*	4.6*	6.07*	1.47	53.55*	39.24*	70.00*
Woodcock-Johnson: Applied problems	.94	2.35*	3.36*	1.40	2.42*	1.02	11.54*	7.14*	16.99*
Woodcock-Johnson: Dictation	-.99	1.26	2.17*	2.25*	3.16*	.91	10.03*	6.34*	14.57*
Woodcock-Johnson: Proofing	-.16	1.24	2.93*	1.40	3.09*	1.68*	9.56*	3.61*	18.61*
Peabody Picture Vocabulary Test: Raw score	1.59	2.75*	2.54*	1.16	.95	-.22	7.00*	7.85*	6.19*
Peabody Picture Vocabulary Test: Standard score equivalent	1.41	2.83*	3.07*	1.42	1.68*	.25	10.33*	9.23*	11.49*
Peabody Picture Vocabulary Test: Mental age	.96	2.57*	2.54*	1.61	1.58	-.03	8.81*	8.93*	8.69*
TACL	.84	2.31*	2.60*	1.47	1.63*	.29	8.79*	7.31*	10.40*
NAPVT	-.44	2.76*	2.55*	3.20*	2.90*	-.22	16.89*	18.19*	15.63*
Arizona: Raw	.18	1.33	3.91*	1.15	3.76*	2.61*	13.23*	3.14*	30.29*
Communication Inventory	1.80*	1.52	3.91*	-.36	2.02*	2.39*	6.41*	.68	17.96*
Teacher Rating of Communication Skills	1.75*	1.82*	3.29*	.07	1.54	1.46	5.71*	1.82	11.78*
TAS - Raw	2.36*	3.26*	4.70*	.92	2.35*	1.43	16.02*	8.92*	25.19*
Time Hearing Aid Worn	-.59	3.36*	5.91*	3.95*	6.5*	2.56*	47.76*	26.29*	75.59*
% Child Communication Understood by Family	-.95	-1.77*	-3.91*	-1.82*	-2.95*	-2.14*	10.97*	3.30*	23.12*
% Child Communication Understood by Non-Family	-1.29	-3.07*	-5.14*	-1.79*	-3.86*	-2.07*	24.33*	11.96*	41.05*
% Child Understands Family	-.72	-2.44*	-2.40*	-1.72*	-1.70*	-.04	8.88*	8.72*	9.05*
% Child Understands Non Family	-1.50	-4.21*	-4.24*	-2.64*	-2.67*	-.03	23.87*	23.56*	24.08*
Rating of Child Behavior	3.24*	1.59	2.06*	-1.66	-1.18	.47	.09	.002	.41
Meadow Kendall - Social Adjustment	2.71*	3.02*	3.83*	.31	1.12	.81	8.70*	5.67*	12.57*
Meadow Kendall - Self-Image	2.56*	1.00*	2.07*	-.76	.37	1.07	2.20	.47	4.44*

Table 20

Percent of f and t Values Favoring Groups 1 and/or 2

Comparisons	% of Statistically Significant f and t Values Favoring Groups 1 and/or 2	% of all f and t Values Favoring Groups 1 and/or 2
1 vs. 4	100%	100%
1 and 2 vs. 4	95%	100%
1 and 2 vs. 3 and 4	92%	100%
1 and 2 vs. 3	79%	100%
2 vs. 4	75%	96%
1 vs. 3	71%	100%
2 vs. 3	46%	92%

These results indicate that the hearing impaired children in this study who received home intervention perform better on the majority of dependent variables than children who did not receive home intervention.

When performing multiple comparison procedures on Group 1 vs. 2 (early vs. late) and Group 3 vs. 4 (no home intervention/preschool vs. no home intervention/no preschool), the majority of t-values were not statistically significant. This indicates that early vs. late and preschool vs. no preschool effects are largely nondifferential for children in this study.

When considering the percent of statistically significant group differences in dependent variable categories, early intervention children perform better on communication/language skills, while late intervention children performed better on achievement tests. This may suggest that early home intervention more directly affects later language.

Analysis of Educational Significance

Effect size analysis. In order to determine the existence of educationally significant differences of specific group means, effect sizes were obtained for all comparisons discussed in the previous section. These results are shown in Table 21. Effect sizes of larger than .5 are considered to be important from an educational standpoint and are noted with asterisks.

It should be noted that effect sizes for all 36 dependent variables are given below, since even though statistically significant overall group differences do not exist for 12 of these dependent variables, there is still possibility for there to be educationally significant group differences for these variables (and vice versa).

For variables 18-21, 33, and 35, negative effect sizes still indicate better performance by the first group in the comparison, since higher scores on these variables are indicative of poorer performance.

Discussion. When considering the number of educationally significant effect sizes compared to statistically significant f and t values, it is apparent that there are more statistically significant f and t values for the 24 dependent variables that show statistically significant overall group differences. However, educationally significant effect sizes exist for the majority of the 24 dependent variables in the majority of group comparisons. These observations are summarized in Table 22.

There are more educationally significant differences for dependent variables that were previously determined to be statistically significantly different among groups than for those dependent variables that were not. For those dependent variables that were previously determined not to be statistically significantly different among groups, 14% (15 out of 108) of the effect sizes are educationally significant. Twelve of these 15 effect sizes favor the following groups: home intervention over no home intervention, early over late home intervention, and preschool over no preschool.

Table 21

Effect Sizes for Research Group Comparisons

Dependent Variables	1 vs. 2	3 vs. 4	1 vs. 3	2 vs. 3	1 vs. 4	2 vs. 4	1 & 2 vs. 3	1 & 2 vs. 3 & 4	1 & 2 vs. 4
Variables that are statistically significantly different among groups:									
1) Woodcock-Johnson: Letter/Word Identification	.21	.14	.60*	.67*	.82*	[illegible]	.67*	.74*	[illegible]
2) Woodcock-Johnson: Word Attack	-.32	.40	.35	.67*	.83*	1.15*	.51*	.75*	.97*
3) Woodcock-Johnson: Passage Comprehension	-.26	.38	.11	.37	.49	.75*	.28	.43	.62*
4) Woodcock-Johnson: Calculation	-.08	.26	.74*	.82*	1.00*	1.09*	.70*	.91*	1.05*
5) Woodcock-Johnson: Applied Problems	.22	.23	.54*	.25	.77*	.56*	.43	.54*	.66*
6) Woodcock-Johnson: Dictation	-.23	.21	.30	.53*	.51*	.75*	.42	.52*	.63*
7) Woodcock-Johnson: Proofing	-.04	.40	.29	.33	.69*	.73*	.38	.51*	.71*
8) Peabody Picture Vocabulary Test: Raw Score	.39	-.05	.60*	.29	.62*	.23	.40	.45	.43
9) Peabody Picture Vocabulary Test: Standard Score Equivalent	.35	.06	.70*	.35	.76*	.41	.53*	.56*	.59*
10) Peabody Picture Vocabulary Test: Mental Age	.21	-.01	.63*	.39	.62*	.39	.51*	.51*	.51*
11) TACL	.20	.00	.54*	.34	.62*	.43	.44	.40	.52*
12) ROWPVT	-.10	-.06	.64*	.74*	.59*	.69*	.69*	.66*	.64*
13) Arizona: Raw	.04	.51*	.26	.23	.70*	.74*	.24	.50*	.71*
14) Communication Inventory	.45	.57*	.36	-.09	.93*	.40	.14	.42	.64*
15) Teacher Rating of Communication Skills	.46	.38	.40	.02	.70*	.40	.25	.44	.76*
16) OIS - Raw	.50*	.35	.80*	.23	1.10*	.50*	.52*	.69*	.87*
17) Time Hearing Aid Worn	-.13	.56*	.74*	.87*	1.30*	1.43*	.81*	1.03*	1.32*
18) % Communication Understood by Family	-.22	-.50*	-.41	-.19	-.91*	[illegible]	-.30	-.56*	-.80*
19) % Communication Understood by Non-Family	-.31	-.50*	-.74*	-.43	-1.23*	-.92*	-.50*	-.83*	-1.00*
20) % Child Understands Family	-.20	-.01	-.66*	-.47	-.67*	-.40	-.57*	-.50*	-.50*
21) % Child Understands Non Family	-.41	-.01	-1.10*	[illegible]	-1.11*	-.70*	-.90*	-.90*	-.90*
22) Rating of Child Behavior	.96*	.14	.47	-.49	.61*	-.35	0	.07	.61*
23) Meadow Kendall - Social	.67*	.20	.74*	.08	.94*	.28	.41	.51*	.41
24) Meadow Kendall - Self-Image	.66*	.28	.46	-.20	.74*	.08	.13	.27	.14
Variables that are not statistically significantly different among groups:									
25) Arizona: Consonant Error	.11	-.41	.06	-.06	-.34	-.46	0	-.20	-.40
26) USS: MLU	-.19	-.24	.24	.42	-.01	.18	.33	.21	.09
27) Parent Attitude: Total	.56*	.01	.45	-.11	.46	-.10	.17	.17	.18
28) Parent Attitude: Reactions to Outside Help	.36	-.29	.49	.13	.20	-.16	.31	.16	.02
29) Parent Attitude: Anxiety/Guilt	.39	-.41	.42	.04	.03	-.36	.23	.03	-.17
30) Parent Attitude: Acceptance	.63*	.35	.34	-.29	.69*	.06	.02	.20	.37
31) Meadow Kendall: Emotional Adjustment	.21	.17	.45	.25	.62*	.43	.35	-.44	.53*
32) Time Spent Reading	[illegible]	-.02	-.46	.14	-.41	.12	-.16	-.17	-.10
33) Solitary vs. Group Play	-.57*	-.91*	-.54*	.03	-1.05*	-.40	-.26	-.51*	-.70*
34) Number of Friends	.51*	-.41	.10	-.36	-.26	-.79*	-.09	-.31	-.52*
35) Attitude towards School	-.41	.02	-.10	.22	-.16	.24	.02	.02	.04
36) Time % Communication	-.20	.01	-.31	-.12	-.20	[illegible]	-.21	.20	-.17

Table 22

Number of Statistically Significant f and t Values vs. Educationally Significant ES Values

Comparisons	No. of Statistically Significant f or t Values	No. of Educationally Significant ES Values
1 vs. 2	6 (t)	4
Group 3 vs. 4	8 (t)	4
Group 1 and 2 vs. 3	19 (f)	11
Group 1 and 2 vs. 3 and 4	22 (f)	17
Group 1 and 2 vs. 4	23 (f)	21
Group 1 vs. 3	17 (t)	14
Group 2 vs. 3	11 (t)	7
Group 1 vs. 4	24 (t)	33
Group 2 vs. 4	18 (t)	14

Comparisons that were previously determined to be statistically significantly different among groups consistently favor the first group in the comparison at a level of educational significance. In the 1 vs. 4, 1 vs. 3, 2 vs. 4, 1 and 2 vs. 4, and 1 and 2 vs. 3 and 4 comparisons, the majority of the effect sizes favor the children in home intervention groups.

Home intervention children are again favored in the Group 1 and 2 vs. 3 contrast, where all groups are equated on preschool. The positive long-term effects of home intervention vs. no home intervention on hearing impaired children are suggested in these results.

Implications and Recommendations for Future Research

In this section, research findings will be presented, and then implications of each finding will be listed. In a statistical sense, these implications are true only to the extent that external validity exists in the study. Measures taken to control threats to external validity were outlined in Table 15.

Finding

Hearing impaired children in this study who receive home intervention earlier in their lives performed better than children who did not receive home intervention on the majority of dependent variables in the areas of language, academic achievement and psycho-social behavior.

Implications

1. Home intervention promotes the development of basic skills that enhance later language, academic, and psycho-social functioning. Hearing impaired children who receive home intervention services may be able to function better at home and at school than children who do not receive home intervention services. Home intervention children may be better able to interact with family, peers, and teachers as evidenced by their superior communication and psycho-social competencies. They may also be able to function better academically in school since academic achievement skills are improved.

2. Parents and siblings who receive home intervention are apparently able to communicate more effectively with the hearing impaired child than parents and siblings who do not receive home intervention, since the child's communication and interactive skills are improved.

3. Teachers and professionals who deal with home intervention children may also be able to more effectively interact with these children because of improved communication, academic, and psycho-social skills. In addition, they may spend more time on the promotion of subject matter skills instead of language-related skills (contrary to the typical educational programming of hearing impaired children who enter school without strong language bases). Also, teacher time spent on management of hearing aid, management of problem behavior, and explanation of school tasks and protocol may be reduced with home intervention children.

Finding

Children in this study who received early home intervention performed better than late intervention children on some of the dependent variables. Early intervention children performed better on communication/language skills in relationship to academic skills.

Implications

1. The success of home intervention is dependent on many factors, including: timing, duration, and intensity of intervention efforts.

Finding

Hearing impaired children in this study who received preschool, but not home intervention services, performed better than children who did not receive preschool or home intervention on certain dependent variables.

Implications

1. The impact of home intervention may be strengthened by the provision of other services, such as preschool. The provision of home intervention and preschool may result in greater communication, academic, and psycho-social benefits for the child than the provision of either service (especially preschool) alone.

Finding

Many factors, particularly child and parent characteristics, account for the majority of the variance of the dependent variables if not controlled in the analyses.

Implications

1. Effectiveness of home intervention is dependent on many factors. In this study, it was determined that some of the most important factors contributing to the long-term effects of home intervention were: child age, hearing loss, parental index of social position, existence and severity of other handicaps, age of mother, and number of middle ear infections. Of course, there are others. Since it was not within the scope of this research project to specifically study what and how child and parent characteristics contribute to later success, a complete description of their efforts is not possible.

   However, from this study and others (Gage & Berlinger, 1979), one important implication that emerges is the necessity of optimizing factors that might contribute to later child success, such as (a) reducing middle ear infections, (b) mitigating effects of other handicaps, (c) improving SES, and (d) improving such parental characteristics as time interacting with child, aspiration for child achievement, emphasis of language development, provision of learning opportunities in the home, and acceptance of the child. It should be noted that the long-term impact of home intervention may be dependent on the nature of the intervention. The SKI*HI model (which directly habilitates communication in the hearing impaired child) may have more direct effects on later language than academic skills. Or perhaps the nature of any home intervention program is such that impact will be greater on later language vs. academic skills, since many skills requiring habilitation are age specific (langauge skills precede academic skills). In either case, the nature of the intervention may have an effect on later child performance.

Finding

Many dependent variables did not reveal statistically significant differences between research groups.

Implications

1. It is impossible to determine with precision why group differences did not exist for a few of the dependent variables. However, some possible reasons for the no-difference findings are:

   a. Problems with validity and reliability could have existed for measures used in this study, particularly the parent attitude scale, some items in the parent questionnaire, and the DSS-MLU.

   b. Intervention could have provided effective services during its tenure, but could not completely buffer families from the adverse effects of hearing impairment during later periods of developmental crisis.

   c. The advances made by intervention children during treatment could have been reduced or reversed when these children were later grouped in schools with no intervention children.

Whatever the reasons for no difference, it becomes apparent that further research is needed to see if some dependent variable differences really do not exist, and, if so, for what reasons, and to develop intervention strategies that more successfully remediates the dependent variable skills. While this study has resulted in some useful findings in regards to the long-term effects of intervention on hearing impaired children, much research remaines to be done:

1. Continued longitudinal data collection on the intervention children involved in this study are needed.

2. Studies are needed on the impact of home intervention on child and parent competencies not included in this study.

3. Studies are also needed on the impact of home intervention on areas other than child and parent competencies, such as sibling attitudes, family/marital structures, extended family involvement, and community awareness.

4. Further studies are needed to replicate the Clark and Covert (1979) study on short-term effects of home intervention and the effect of early vs. late home intervention on hearing impaired children.

5. Cost-effectiveness studies are needed which would involve:

   a. Identification of all treatment alternatives,

   b. Description of all components necessary for administration of treatment alternatives.

   c. Assignment of cost values to all resources.

   d. Analysis of cost outlay in terms of child and parent progress.

6. Studies need to be done isolating parent, child, and environmental factors that are highly related to later child success and that are remediable, such as parent-child interaction styles, parent motivation and aspiration for child's achievement, home environments arranged for learning, parent encouragement of child's autonomy, and parent acceptance of the handicapped child.

Products

Number in Product List	Title
117	Intervention with Hearing Impaired Children

INVESTIGATION 4:

A COST-EFFECTIVENESS ANALYSIS OF PARAPROFESSIONALS VS. PROFESSIONALS IN EARLY INTERVENTION FOR THE HANDICAPPED

Although hundreds of studies have been conducted on various aspects of early intervention for preschool children, there is still considerable disagreement as to the relative worth of early intervention services. Determining the relative worth of programs requires that both effects and costs of program alternatives be jointly analyzed. For example, the most effective program may be the most expensive, and, as a result, may not be the most cost-effective. Conversely, a program may be relatively inexpensive, but not particularly effective and, therefore, not cost-effective either. At a time when resources are limited, purchasing the best outcome per dollar is particularly sensible and mandates that comparative cost and outcome data be considered simultaneously if intelligent decisions are to be reached.

An important cost-effectiveness question facing preschool service providers is how to reduce costs while maintaining or increasing benefits. One technique which is often proposed as a way to reduce costs is to increase the use of paraprofessionals in providing intervention services to children. Of course, hiring lower-salaried staff raises questions about whether paraprofessionals are as cost-effective as professionals

The literature mentions many instances in which the use of paraprofessionals and volunteers as direct intervenors appears feasible and effective (Barbarack, Horton, & Karnes, 1973; Shortinghuis & Frohman, 1974). However, in no "professional vs. paraprofessional" study are the analyses of effects accompanied with accurately computed cost data. Other prevalent methodological problems found in "professional/paraprofessional" studies included the confounding of treatments, inadequate assignments of costs to providers, poor experimental designs, inadequate definitions of treatment, and no indication of quality or degree of treatment implementation.

The problem then is that there is little "high-quality" research which directly compares the cost-effectiveness of paraprofessionals versus professionals in delivering preschool intervention services to handicapped children. The present study was designed as a cost-effectiveness study where "costs" are defined as an, expended resources necessary to implement the program as defined. Program outcomes were measured on a broader scale than what has been typically done and included child and parent variables.

Objectives

This research had the following major objectives:

1. To evaluate the relative cost-effectiveness of two intervention service delivery strategies.

2. To field test and further refine a data collection system for recording and summarizing detailed information on the cost of implementing various components of early intervention programs.

The specific research questions to be answered are:

1. Are there statistically significant differences in child and family outcome measures between intervention (professional vs. paraprofessional) modalities.

2. Are costs for serving children with professionals significantly different from costs for serving children with paraprofessionals?

3. Is it more cost-effective to use paraprofessionals or professionals to provide speech, physical, and occupational therapy to moderately handicapped children?

Approach

Subjects

Subjects were 46 children ranging in age from 30 to 60 months from a community-based program for handicapped preschoolers. (Cost data were collected on approximately 76 subjects. This includes the 46 "professional vs. paraprofessional treatment" subjects, subjects, as well as 30 additional children who attended the same classes as these children, but who were not matched.) All the subjects were moderately or severely handicapped, and all were currently receiving occupational therapy and/or physical therapy and speech therapy.

Matching

The 46 subjects were matched within their respective classes according to severity of handicap, type of handicap, chronological age, and then randomly assigned to one of two experimental groups: "professional intervention group" or "paraprofessional intervention group." Matching according to severity of handicap was accomplished using pretest data on standardized IQ, motor and speech and language tests. Matching data based on these pretest results are contained in Tables 23 and 24.

Table 23

Pretest Matching Data for Children in Paraprofessional and Professional Treatment Groups

	CA Para	CA Pro	IQ Para	IQ Pro	Motor Para	Motor Pro	Language Para	Language Pro
Class 1 Xs	33.0	36.0	61.33	59.0	47.67	49.83	60.67	51.83
Class 2 Xs	57.6	55.2	68.2	63.8	59.8	66.1	68.3	68.7
Class 3 Xs	44.67	42.0	54.0	55.0	63.33	52.0	54.67	58.83
Class 4 Xs	36.67	37.0	94.5	82.0	46.5	48.5	65.67	67.67
Class 5 Xs	55.6	56.6	43.8	45.5	53.4	46.4	55.8	51.3
Class 6 Xs	46.5	47.25	--	--	19.75	20.0	19.75	21.0
Total Xs	47.61	47.52	59.00	56.71	48.59	47.54	54.02	53.00
Total SDs	11.52	11.18	21.30	21.14	18.40	18.61	22.70	21.87

N = 46

Table 24

Data for Individual Children on Matching Variables

	CA - PARA	CA - PRO	IQ - PARA	IQ - PRO	MOTOR - PARA	MOTOR - PRO	LANGUAGE - PARA	LANGUAGE - PRO
	31	32	61	61	48.5	53	52	45.5
	30	31	72	71	56.5	54.5	76	65
	38	45	51	45	38	42	54	45
CLASS 1 $\overline{X}$s	33.0	36.0	61.33	59.0	47.67	49.83	60.67	51.83
	57	54	62	60	68.5	82.5	64	55
	60	60	57	51	47.5	42.5	52.5	54
	60	61	77	73	68.5	61.5	72.5	78.5
	62	52	45	49	53.5	62.5	52.5	60
	49	49	100	86	61	81.5	100	96
CLASS 2 $\overline{X}$s	57.6	55.2	68.2	63.8	59.8	66.1	68.3	68.7
	50	50	44	50	49	45	50	50
	39	37	60	57	74	57	58	66.5
	45	39	58	58	67	54	56	60
CLASS 3 $\overline{X}$s	44.67	42.0	54	55	63.33	52	54.67	58.83
	35	33	112	95	55.5	60.5	94	85.0
	28	30	77	106	64	62	79	88
	47	48	—	45	20	23	24	30
CLASS 4 $\overline{X}$s	36.67	37	94.5	82	46.5	48.5	65.67	67.67
	63	63	37	40	62	46	50.5	58
	60	60	30	32	63.5	44.5	50	46
	55	54	68	58	56	58.5	63	58
	52	55	53	53	39.5	42	58.5	49
	48	51	31	44	46	41	57	45.5
CLASS 5 $\overline{X}$s	55.6	56.6	43.8	45.4	53.4	46.4	55.8	51.3
	33	33	47	46	21	24	31.5	38
	37	37	—	11	7	4.5	7.5	12.5
	58	61	—	—	33	28	21	13.5
	58	58	38	—	18	23.5	19	20
CLASS 6 $\overline{X}$s	46.5	47.25			19.75	20	19.75	21
TOTAL $\overline{X}$s	N=23 47.61	N=23 47.52	N=20 59.00	N=21 56.71	N=23 48.59	N=23 47.54	N=23 54.02	N=23 53.00

Table 25

Program Checklist

ECC ____________________

Date ____________________

Responsible Professional(s) ____________________

Student ____________________

Program ____________________

PROGRAM CHARACTERISTICS	CHECKLIST
1. Objective Target Behavior	1. Target Behavior
A. Behavior is observable (detectable by senses)	A. yes no
B. Behavior is quantifiable	B. yes no
C. Behavior is verifiable	C. yes no
2. Behavior Task Analyzed	2. Task Analyzed
A. Behavior broken into sufficiently small components?	A. yes no
3. Instructional Procedures Specific	3. Instructional Procedures
A. Prompts	A. Prompts
1. Is setting described?	1. yes no
2. Are materials specified?	2. yes no
3. Are intervenor behaviors (prompts) specified?	3. yes no
4. Are learner behaviors specified?	4. yes no
5. Are correction procedures specified?	5. yes no
4. Measuring Student Progress	4. Measuring Progress
A. Person(s) responsible specified?	A. yes no
B. How frequently specified?	B. yes no
C. When its measured specified?	C. yes no
D. How its recorded specified? (Is record form developed?)	D. yes no
5. Determine Next Step	5. Next Step
A. Data considered specified?	A. yes no
B. Criteria specified?	B. yes no
C. Who is responsible?	

scores. The Peabody Motor Scales were administered to obtain motor scores. (The Peabody Motor Scales measure gross and fine motor skills for children 0-83 months of age. It was standardized with 617 normal children who were representative to the U.S. population.) The Sequenced Inventory of Communication Development (SIDC) was administered to obtain speech and language scores. (The SICD measures the receptive and expressive language of normal and retarded children who are functioning between 4 months and ^ years of age. It was standardized on a sample of 252 children.)

Pretesting was accomplished during October, November, and December of 1983. Posttesting, using these same measures, was accomplished during May 1984. All posttesting was conducted by qualified personnel who do not know to which group children have been assigned.

Other Effects Measures

Parent Satisfaction. Parents were surveyed to obtain information regarding their degree of satisfaction with the treatment program. Results from the parent satisfaction surveys are shown in Table 26. These data do not indicate cross group differences.

Table 26

Summary of Parent Satisfaction Survey*

Area of Satisfaction	Paraprofessional Treatment (N = 19)	Professional Treatment (N = 22)	Combined Treatment (N = 41)	Non-Treatment (N = 19)
1. Quality of Services	4.55	4.54	4.54	4.73
2. Communication with Staff	4.42	4.72	4.58	4.63
3. Educational Goals and Objectives	4.63	4.59	4.60	4.63
4. General Program Satisfaction	4.65	4.45	4.54	4.47

*These data are based on a 5-point scale, with 5 being "very satisfied" and 1 being "very dissatisfied"

<u>Staff Satisfaction</u>. Professional and paraprofessional staff were surveyed to obtain information regarding their degree of satisfaction with several components of the treatment program. Results from these surveys are shown in Table 27.

Table 27

<u>Summary of Staff Satisfaction Survey*</u>

Area of Satisfaction	Paraprofessional Staff (N = 7)	Professional Staff (N = 20)	All Staff (N = 17)
1. Training of Paraprofessionals	3.3	2.8	3.3
2. Children's Motor Programs	4.1	3.5	3.8
3. Children's Language Programs	3.9	3.4	3.7

*These data are based on a 5-point scale, with 5 being "very satisfied" and 1 being "very dissatisfied"

Several comments made by professional staff, which may relate to the lower professional satisfaction scores on items 1 and 2, follow:

-- "More professional staff are needed."

-- "Professionals are expected to train paras in subtleties of therapy that has taken a college degree and years of experience."

-- "Paras do well, but not enough time to train them to be effective and consistent."

-- "[Paras] lack of understanding of theory is a weak point leading to a decrease in follow-through and quality."

(It should be noted that there were many other positive comments regarding the project in general from both paraprofessionals and professionals.)

<u>Educational Objectives</u>. For each child in the study, written Individual Program Plans containing educational activities and objectives were developed (prior to assignment to groups). A representative example of such a plan is shown in Figure 4.

FIGURE 4 SAMPLE INDIVIDUAL PROGRAM PLAN

Learner
Date St.
Date Mast.
Supervisor Pam Miller
Implementor

Program Name Opposites
Term. Obj. To receptively identify pairs of opposites spontaneously

Reinforcement Verbal/social praise, stickers, play time
Correction (1) Say "no" and provide correct response. (2) Manipulate through correct response.

STEP	MATERIAL	WHAT YOU DO WITH THE MATERIAL	WHAT YOU DO AND SAY	WHAT THE LEARNER DOES	TRIALS	CRIT.
1.	Pictures or Objects (ex: big book and a little book, etc.)	Present a pair of opposites.	While pointing to the picture or object, say "This is big. This is little. Now you point to little."	Child correctly points to opposite, with physical assistance from trainer.		100% 3 Consecutive Sessions
2.	Same	Same	Same	Child correctly points to opposite, following trainer's model.		
3.	Same	Same	Say "This is big. Point to little."	Child correctly points to opposite upon verbal request only.		
4.	Same	Same	Point to one of the two opposites and say "This is big If it is not big it must be ___?___." "Show me."	Child correctly points to opposite.		
5.	Same	Present three cards or objects, including 1 pair of opposites.	Point to one of the two opposites and say "This is big. If it is not big it must be ___?___."	Child correctly points to opposite spontaneously.		

The number of educational objectives achieved by each child were obtained upon completion of the study.

Cos' Measures

Three categories of cost data were collected: personnel costs (salaries and benefits), nonpersonnel costs (equipment, facilities, transportation), and contributed resources (parent time, volunteer time, consultants, etc.).

Cost data were collected from several sources. First, all intervenors tracked their time daily according to 15 activity categories (e.g., language, motor, preparation, inservice, etc.). They began time tracking approximately one month prior to treatment onset and continued through the end of the study. As can be seen in Figure 2, which is a sample time tracking form, each intervenor indicated (1) the time interval in which he/she is engaged in a particular activity, (2) the child(ren) who are involved, and (3) the type of activity. For example, Figure 5 shows that on Monday in class #3, the intervenor (Betty Smith) spent from 1:00 to 1:30 teaching speech program to 3 children (ID#s 127 136, and 170). She spent from 1:30 to 2:00 teaching motor programs to child #139; from 2:00 to 2:15 conducting other direct intervention activities with children #s 127, 136, and 166; from 2:15 to 3:00 preparing lessons for all the children in class #3; and so on.

These time tracking forms were completed on an ongoing basis to reduce suspected error associated with prior time accountability estimation methods and also to help define an appropriate and feasible time tracking sample schedule for future research and future use as a management tool.

In order to more effectively analyze, summarize, and manage the time tracking data a microcomputer software program were developed. Using this software program, raw data from the time tracking forms were entered directly into the computer, and the computer produced weekly tabular printouts which contained cost and time data for each student x activity x intervenor. Tables 28 and 29 illustrate these data. For example, Table 28 provides information for classroom DDI01 for the week of 1/23/84 to

Figure 5

SAMPLE TIME TRACKING FORM

Initials	ID #
C.T.	127
JD	136
RB	137
MA	139
KN	166
PM	170

Weekly Schedule - Direct Service

DDI
Site

Betty Smith
Staff Name

Class #3
Room Number

Initials	ID #

DAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	OTHER
Time	1:00-1:30	1:00-1:30	1:00-1:30	1:10-1:30	1:00-1:30	
Room Exceptions						
Activity	127 (s) 136 (s) 170 (s)					
Time	1:30-2:10	1:30-1:45	1:30-1:45	1:30-1:50	1:30-1:50	
Room Exceptions						
Activity	139(m)*	139(m)*	139(m)*	139(m)*	139(m)*	
Time	2:00-2:15	1:45-2:00	1:45-2:05	1:50-2:00	1:50-2:10	
Room Exceptions						
Activity	127 136 166 } (o)	ALL (c)	ALL BUT 139 (c)	NON-PROGRAM	ALL BUT 139 (c)	
Time	2:15-3:00	2:00-2:20	2:05-2:30	2:00-2:30	2:05-2:30	
Room Exceptions						
Activity	PREP	127 170 } (s)				
Time		2:20-2:45	2:30-3:00	2:30-3:00	2:30-3:00	
Room Exceptions						
Activity		166 139 136 } (o)	PREP	137 139 166 } (m)	PREP	
Time		2:45-3:00				
Room Exceptions		Room #9				
Activity		137 139 166 } (m)				

[redacted]ALUE[redacted]R TH[redacted]K G[redacted]-23[redacted]0 O[redacted]84 [redacted] SAMPLE COST [redacted]INT[redacted] [redacted]
FOR INDIVIDUAL STUDENT

STUDENT: BOBBI JO -12124

Activity		Intervenor Names: BEARD	REED	MATTHE	PRESTW	LEYMAS	CRAWFO	JONES	GOWDIE	MEDINA	PARENT	KELLER	TOTAL	YTD
SPEECH	Tot Hrs	0.9	0.9	1.2	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2	3.4	7.2
	Bill Hrs	0.5	0.3	0.3	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	1.2	
	Cost	6.0	2.1	2.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.3	11.1	24.0
MOTOR	Tot Hrs	0.4	0.8	0.2	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.2	1.9	4.3
	Bill Hrs	0.2	0.3	0.1	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.9	
	Cost	2.0	2.1	0.4	0.0	2.9	0.0	0.0	0.0	0.0	0.0	0.2	7.6	13.1
OTHER	Tot Hrs	0.2	0.1	0.2	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.3	1.0	7.2
	Bill Hrs	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.3	
	Cost	0.7	0.2	0.3	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.4	2.6	16.9
CSUPERV	Tot Hrs	0.2	1.1	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	2.0	6.3
	Bill Hrs	0.1	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.6	
	Cost	0.9	2.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2	3.9	13.8
PREP	Tot Hrs	3.0	3.4	4.1	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.6	11.2	37.2
	Bill Hrs	0.6	0.7	0.6	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	2.1	
	Cost	8.2	6.0	4.0	0.0	0.0	0.0	0.7	0.0	0.0	0.0	0.7	19.6	78.8
INSERV	Tot Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6
	Bill Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
	Cost	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.8
ASSESS	Tot Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
	Bill Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
	Cost	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.9	0.0	8.2
PARENT	Tot Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
	Bill Hrs	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
	Cost	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.6
TOTALS:	Tot Hrs	4.6	6.3	6.3	0.0	0.2	0.0	0.7	0.0	0.0	0.0	1.4	19.5	
	Bill Hrs	1.4	1.5	1.3	0.0	0.3	0.0	0.2	0.0	0.0	0.0	0.4	5.0	
	Cost	17.7	12.3	7.6	0.0	2.9	0.0	2.4	0.0	0.0	0.0	1.9	44.9	
	YTD Hrs	17.6	15.4	21.5	0.1	0.9	0.6	0.7	0.9	4.1	0.0	1.4	63.1	
	YTD Cost	56.4	30.1	28.6	1.7	3.7	1.1	2.4	22.8	8.4	0.0	1.9	157.2	

Table 29

Class Room Code = 00101

SAMPLE COST PRINTOUT
WEEKLY SUMMARY FOR CLASS (DDI01)

VALUES FOR THE WEEK OF 01-23-84 TO 01-27-84

	Intervenor Names														
	BEARO	REEO	MATTHE	PRESTW	LEYMAS	CRAWFO	JONES	GOWOIE	MEOINA	PARENT	KELLER	Tot Hrs	Cost	YD Hrs	YTOCost
SPEECH	2.8	3.1	3.8	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.5	10.86	96.72	35.97	327.51
MOTOR	1.7	2.5	1.1	0.5	0.8	0.0	0.3	0.0	0.0	0.0	0.4	7.26	68.01	17.98	171.14
OTHER	0.7	0.7	1.1	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.7	4.12	36.65	25.08	227.43
CSUPERV	0.8	2.4	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	5.25	43.13	20.88	172.04
PREP	7.5	8.5	7.8	0.0	0.0	0.0	0.8	0.0	3.5	0.0	1.5	29.51	261.90	94.69	862.85
INSERV	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	1.08	9.54
ASSESS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.00	0.00	2.91	31.37
PARENT	0.0	0.0	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.17	7.33	2.08	19.26
LUNCH	0.0	0.3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.33	2.79	0.33	2.79
Tot Hrs	13.5	17.5	16.5	0.5	0.8	0.0	2.8	0.0	3.5	0.0	3.5	58.50			
Cost	174.0	146.2	103.7	6.7	6.6	0.0	32.7	0.0	28.3	0.0	18.3		516.53		
YTO Hrs	54.3	48.0	62.2	2.3	3.6	1.0	3.2	2.7	20.3	0.0	3.5			201.00	
YTO Cost	702.6	401.1	390.0	31.3	31.6	12.8	37.6	34.8	163.8	0.0	18.3				1823.93

B>

1/27/84 for one student (Bobbi Jo). It shows the amount of time and the associated cost for each intervenor for each activity. Additionally, year-to-date (YTD) figures for hours and cost are tabulated. For example, in Table 28, intervenor Beard spent 9 hours during the week working on speech programs with Bobbi Jo. The cost for Beard's speech time was $6. Note that Tot Hrs (total hours) and Bill Hrs (billing hours) are different. This is because an intervenor may have worked with more than one child at a time. The software program takes this into account and only charges proportionate costs of an intervenor's time to each child. Table 29 presents the same type of data as Table 28, only summarizes for all children in classroom DDI01.

Second, program personnel were interviewed in person, and accounting records were used to obtain costs for personnel (e.g., salary and fringe benefits), equipment, facilities, and transportation.

Third, program records and parent interviews were used to collect demographic information. Demographic data will be used to further illustrate the conditions, types of subjects, and generalizability of findings of the study.

Results and Discussion

Effects of paraprofessional and professional treatment programs. Means and standard deviations of posttest effects measures are presented in Table 30. IQ, fine motor, gross motor, receptive language, and expressive language posttest scores are presented for the paraprofessional and professional treatment groups. There were no statistically significant differences on any of these posttest measures across groups. That is, children who receive motor and language therapy from paraprofessionals did not perform significantly different than children who received their language and motor therapy from professionals.

Costs of paraprofessional and professional treatment group programs. The costs for paraprofessional and professional programs are shown in Table 31. In this table, as with other cost tables, data are presented by mean cost per child. Operational

Table 30

Mean Posttest Scores (Standard Deviation)

	Paraprofessional Treatment Group 1 (N = 23)	Professional Treatment Group 2 (N = 23)	Treatment Groups 1 and 2 (N = 46)
IQ Test	60.3 (20.0)	61.0 (22.7)	60.6 (21.1)
Fine Motor Test (Raw Score)	163.1 (50.3)	146.9 (40.2)	149.9 (45.2)
Gross Motor Test (Raw Score)	186.8 (61.6)	174.0 (53.9)	180.4 (57.6)
Gross/Fine Motor Test (Raw Score)	170.2 (54.8)	160.7 (45.7)	165.4 (50.1)
Receptive Language Test	56.0 (22.1)	59.7 (22.1)	57.9 (21.9)
Expressive Language Test	51.0 (24.7)	55.2 (19.3)	53.1 (22.0)
Receptive/Expressive Language Mean Test	53.9 (22.1)	57.4 (20.5)	55.7 (21.1)

Table 31

Mean Cost of Service Per Child by Treatment Condition

COST VARIABLE	All Children (N = 95)	Paraprofessional Treatment Group 1 (N = 23)	Professional Treatment Group 2 (N = 23)
OPERATIONAL COSTS			
Personnel			
--Instructional Staff	1,338	1,381	1,419
--Administrative Staff	465	562	464
--Consultants	33	33	33
Subtotal Personnel	1,836	1,876	1,916
Equipment Subtotal	93	93	95
Facility Subtotal	488	530	550
Other			
--Travel	6	4	4
--Communications	58	58	58
--Miscellaneous	89	89	89
Subtotal Other	153	151	151
SUBTOTAL OPERATIONAL COSTS	2,570	2,650	2,712
CONTRIBUTED COSTS			
--Parent Volunteers in Class	26	33	27
--Child Transportation by Parents	525	633	640
--Parent Travel Time	277	334	338
--Home Instruction Time	241	86	218
--Home Space	11	4	10
Subtotal Contributed Costs	1,080	1,090	1,233
TOTAL COSTS	3,650	3,740	3,945

costs have been subcategorized by personnel, equipment, facility, and other. Contributed costs have been subdivided by parent volunteers in class, child transportation by parents, parent travel time, home instruction time, and home space.

As expected, the total costs for the paraprofessional treatment group are slightly less than the professional treatment group. This is true when operational costs are considered separately as well as when contributed costs are included.

Mean hours and costs of direct instruction are presented in Table 32. As one might expect, costs for langauge and motor therapy in the paraprofessional treatment group are lower than costs in the professional treatment group. Professionals in this study made approximately 40% more per hour than the paraprofessionals.

Cost-effectiveness. With the effectiveness and the cost data, it is now possible to examine the question of cost-effectiveness for paraprofessional and professional treatment programs for these handicapped preschoolers. As will be shown below, the final answers to the questions of cost-effectiveness depend to some degree on how decision makers value various types of resources (e.g., operational vs. contributed resources) and outcomes (e.g., IQ score gains vs. receptive language score gains). Table 33 provides cost-effectiveness ratios for paraprofessional and professional treatment groups. CE ratios are calculated as unit of effect divided by cost. This means the greater the size of the ratio, the more value per dollar. Considering only operational costs, the paraprofessional and professional treatment groups attained the cost-effectiveness ratios for IQ. On the three motor test measures, the paraprofessional treatment group attained slightly better cost-effectiveness ratios than the professional treatment group. Conversely, for the three language test measures, the professional treatment group attained slightly better cost-effectiveness ratios. Cost-effectiveness ratios based on the percent of language, motor, and educational objectives achieved were virtually indistinguishable between the two treatment groups. Considering the operational and contributed costs, the same pattern was found. That is, the paraprofessional treatment group attained

Table 32

Mean Hours and Costs of Direct Instruction By Treatment Condition

COST VARIABLE	All Children (N = 95)			Paraprofessional Treatment Group (N = 23)			Professional Treatment Group 2 (N = 23)		
	HOURS		COST	HOURS		COST	HOURS		COST
Language/Speech Therapy	24	(4.96/hr)	119	28	(3.71/hr)	104	21	(5.71/hr)	120
Motor Therapy	22	(6.59/hr)	145	21	(5.33/hr)	112	21	(6.92/hr)	145
All Other Direct Instruction	124	(4.42/hr)	425	142	(3.03/hr)	430	164	(2.80/hr)	459

Table 33

Cost-Effectiveness Ratios for Paraprofessional and Professional Treatment Groups (Unit of Effect - Cost)

Cost Category	Paraprofessional Treatment Group 1	Professional Treatment Group 2
OPERATIONAL COSTS		
IQ Posttest	.023	.023
Fine Motor Posttest	.062	.054
Gross Motor Posttest	.071	.064
Mean Gross/Fine Posttest	.064	.059
Receptive Language Posttest	.021	.022
Expressive Language Posttest	.019	.020
Mean Receptive/Expressive Posttest	.020	.021
% Language Objectives Achieved	.020	.020
% Motor Objectives Achieved	.017	.016
% Educational Objectives Achieved	.015	.016
% Total Objectives Achieved	.018	.017
TOTAL COSTS (Operational and Contributed)		
IQ Posttest	.016	.015
Fine Motor Posttest	.044	.037
Gross Motor Posttest	.050	.044
Mean Gross/Fine Posttest	.046	.041
Receptive Language Posttest	.015	.015
Expressive Language Posttest	.014	.014
Mean Receptive/Expressive Posttest	.014	.015
% Language Objectives Achieved	.014	.013
% Motor Objectives Achieved	.012	.011
% Educational Objectives Achieved	.011	.011
% Total Objectives Achieved	.012	.012

slightly better cost-effectiveness ratios on the motor tests, while they achieved slightly lower cost-effectiveness ratios on the language tests.

Table 34 presents cost-effectiveness ratios based solely on the cost of providing speech and motor therapy. The cost-effectiveness ratios developed from these data more clearly demonstrate in Table 33 that the paraprofessional group achieved better motor related CE ratios; however, when the data are presented in this form, the slightly higher speech related CE ratios attained by the professional group are not apparent. The paraprofessional treatment group achieved very similar cost-effectiveness ratios with respect to the speech therapy cost areas as well.

These analyses indicate that overall, paraprofessionals have the potential to provide a high-quality service at a reduced price.

Table 34

Cost-Effectiveness Ratios for Paraprofessional And Professional Treatment Groups: Speech and Motor Therapy Costs (Unit of Effect - Cost)

Cost Category	Paraprofessional Treatment Group 1	Professional Treatment Group 2
SPEECH THERAPY COST		
--Receptive Language Posttest	.54	.50
--Expressive Language Posttest	.49	.46
--Mean Receptive/Expressive Posttest	.52	.47
MOTOR THERAPY COST		
--Fine Motor Posttest	1.46	1.01
--Gross Motor Posttest	1.67	1.20
--Mean Gross/Fine Posttest	1.52	1.11

Products

Number in Product List	Title
96	Efficacy and Cost-Effectiveness

INVESTIGATION 5:

A LONGITUDINAL COST-BENEFIT ANALYSIS OF THE ABECEDARIAN PROGRAM

Cost Analysis

A cost analysis of the Abcedarian program was conducted to determine the value of resources used by the program in each of five prototypical years. Resources were identified by listing program ingredients as discussed in Levin (1983). Each year resources included professional staff, volunteers, and other non-personnel resources (facilities, equipment, materials, etc.). For comparison purposes, costs were estimated in three ways: One was based upon actual resource use and cost as reported by the Abecedarian program; another upon the estimated cost of providing the same model in the public school system (Kakalik, Furry, Thomas, & Carney, 1981); and a third based upon private preschool costs (Kagan & Neugebauer, 1983; NAEYC, 1984; Ruopp, Travers, Glantz, & Coelen, 1979). All costs were adjusted for inflation using the Implicit Price Deflator (Economic Report to the President, 1986) and are reported in 1986 dollars.

For personnel in the actual Abecedarian program, costs were estimated based upon the mean of salaries reported by the program. The cost of volunteers, who assisted with transportation and worked as aides, was estimated by assigning the minimum wage rate, $3.35 per hour (U.S. Department of Labor, 1986), to their time contributed to the program. Supplies and miscellaneous items were valued according to the program's reported use of these items. Transportation costs were estimated by assigning a rate of $.21 per mile, and driving time was captured in the cost of volunteer time. Administration costs were estimated by dividing total administration costs for the Frank Porter Graham Child Development Center among its 150 staff members to determine cost/staff. Cost/staff was then multiplied by the number of staff members directly involved in each program year. Facility costs were unavailable for the Center and had to be estimated from the average cost of facilities in daycare programs across the country (Ruopp et al., 1979).

For the cost of providing the same model in the public school system, the cost of a director, aides, social worker, transportation clerk, substitute teacher, secretary, and non-personnel resources was estimated based upon data presented in Kakalik et al. (1981). Teacher salaries were estimated from National Education Association data (U.S. Department of Education, 1985). The Addendum to this investigation (pp. 125-128)n gives a breakdown of FTE salaries used to estimate personnel costs.

Private daycare costs were estimated from the National Daycare Study (Ruopp et al., 1979), the NAEYC (Public Policy Survey, 1986), and the Child Care Information Exchange Center Director's Survey. Tables 35 through 37 summarize the results of this effort, reporting resource usage and costs for each of the 5 years that children participated in the program.

Table 35 shows that the prototypic Abecedarian program enrolled 14 newborns in Year 1. Salaried personnel consisted of 1 full-time supervisor, 3 full-time "cottage parents," 1 part-time substitute, and 2 part-time volunteers. In addition, consultants provided in-service and pre-service training. Table 35 indicates that total cost per child for personnel was $5,786. The estimated cost of providing the same personnel in the public school system would be $4,952 and $3,600 in a private daycare setting. Personnel expenditures represent 78%, 64%, and 69% of total costs for each of these estimates, respectively. Non-personnel costs in the nursery included supplies, miscellaneous items (equipment repair, insurance, etc.), transportation, and center administration. Total non-personnel expenditures were $1,645 for the Abecedarian program. Non-personnel costs could not be broken out into cost per child for each category in the same way for public and private nursery programs. The Addendum shows the breakdown of annual non-personnel expenditures for Abecedarian, public and private nursery schools. Non-personnel expenditures were

Table 35

Personnel and Non-Personnel Resources in Year One (Nursery n=14) (1986 dollars)

Personnel	FTE	Abecedarian	Public Preschool	Day Care
Supervisor	1.0	$22,408	$30,917	$13,717
Cottage Parent	3.0	51,687	32,427	30,942
Substitute	0.5	6,328	5,405	5,157
Consultant		483	483	483
Volunteer		97	97	97
		$81,003	$69,329	$50,396
Total Personnel (per child)		$5,786(78%)	$4,952(64%)	$3,600(69%)
Non-Personnel (per child)				
Supplies		$ 561		
Miscellaneous		$ 184		
Transportation		$ 82		
Administration		$ 330		
Facilities		$ 488		
Total Non-Personnel		$1,645(22%)	$2,767(36%)	$1,600(31%)
Total Cost (per child)		$7,433	$7,719	$5,200

$2,767 for a comparable public, and $1,600 for a comparable private program. These expenditures account for 22%, 36%, and 31% of total program costs, respectively. Total cost per child was $7,433 for Abecedarian, $7,719 for public preschool, and $5,200 for private day care in Year 1.

In Year 2, half of the expenditures were accounted for by the nursery program and half by the preschool program because children entered the preschool at 18 months. To calculate total cost, we simply divided resources and children from the nursery and preschool in half and added these 2 values together:

$$\text{TOTAL COST}_{(\text{year } 2)} = 1/2(\text{PRESCHOOL COST} + \text{NURSERY COST})$$

For this reason, some of the resources are not presented as whole numbers. Table 36 provides a summary of the resource requirements for year two and costs per child. Salaried personnel consisted of 1 director, 2.5 teachers, 4 aides, 1 part-time substitute teacher, 2 volunteers, 1 transportation clerk, and 1 secretary. A part-time social worker and consultants were also employed. Total personnel costs per child in Year 2 (n=29) were $6,157. For the same service in a public school, we estimated $6,323, and $3,475 for a private nursery/daycare program. For each program, respectively, personnel expenditures account for 79%, 70%, and 69% of total program costs. Non-personnel costs--transportation, administration, facilities, supplies, and miscellaneous--were $1,645 for the Abecedarian program, $2,767 in a public program, and $1,600 in a private program with percentages of 21%, 30% and 31%, respectively. Total cost per child was $7,802 for Abecedarian, $9,090 for public preschool, and $5,075 for private daycare.

Resource expenditures in Years 3 through 5--preschool years--are summarized in Table 37. The following personnel resources were prorated across 44 children: 1 director, 5 teachers, 5 aides, 1 secretary, 1 transportation clerk, and 1 part-time substitute. As in Years 1 and 2, a social worker, volunteers, and consultants were utilized on a part-time basis. Per child personnel expenditures were $6,275 for

Table 36

Costs for Personnel and Non-Personnel Resources in Year Two (n=29) (1986 dollars)

Personnel	FTE	Abecedarian	Public Pre School	Day Care
Director	1.0	$23,838	$35,082	$17,164
Teacher	2.5	56,020	77,293	34,293
Aides/Cottage Parents	4.0	68,916	43,236	41,256
Substitute	0.5	6,328	5,405	5,157
Transportation Clerk	0.5	9,586	8,315	
Secretary	0.5	9,380	8,399	
Social Worker	0.125	3,286	4,438	1,715
Volunteers		203	203	203
Consultants		1,000	1,000	1,000
Total Personnel Costs		$178,557	$183,371	$100,788
Personnel Costs (per child)		$6,157(79%)	$6,323(70%)	$3,475(69%)
Non-Personnel (per child)				
Supplies		561		
Miscellaneous		184		
Transportation		82		
Administration		330		
Facilities		488		
Total Non-Personnel		$1,645(21%)	$2,767(30%)	$1,600(31%)
Total Cost (per child)		$7,802	$9,090	$5,075

Table 37

Costs for Personnel and Non-Personnel Resources in Years 3 to 5 (Preschool n = 44, 1986 dollars)

Personnel	FTE	Abecedarian	Public Pre School	Day Care
Director	1.0	$25,268	$39,247	$20,611
Teacher	5.0	112,040	154,585	68,585
Aides	5.0	86,145	54,045	51,570
Substitute	0.5	6,328	5,405	5,157
Transportation Clerk	1.0	19,172	16,629	
Secretary	1.0	18,760	16,798	
Social Worker	0.25	6,572	8,876	3,429
Volunteers		308	308	308
Consultants		1,517	1,517	1,517
Total Personnel Costs		$276,110	$298,230	$151,177
Personnel Costs (per child)		$ 6,275	$ 6,778	$ 3,436
Non-Personnel (per child)				
Supplies		561		
Miscellaneous		184		
.ransportation		82		
Administration		330		
Facilities		488		
Total Non-Personnel		$1,688(21%)	$2,767(30%)	$1,600(31%)
Total Cost (per child)		$7,943	$9,545	$5,036

Abecedarian, $6,778 for a public program, and $3,436 for a private program. These figures represented 79%, 71%, and 69%, respectively, of total program costs. Non-personnel expenditures in the preschool years were $1,640, $2,767, and $1,600, accounting for 21%, 29%, and 31% of total program costs, respectively. Total cost per child was $7,943 for Abecedarian, $9,545 for public preschool, and $5,036 for private daycare during the preschool years.

Table 38 summarizes the cost of the Abecedarian program in each year and total cost under a range of discount rate assumptions and compares it to a public and private daycare setting. Implementing the Abecedarian model in the public school would be the most expensive option, followed by actual costs of the program at the Frank Porter Graham Child Development Center. The least expensive option would be to incorporate this model in a private daycare setting.

Benefit Analysis

Benefits of the Abecedarian program were estimated based upon the Perry Preschool program and the cost of private daycare in the United States. Although it appears that the Abecedarian program may result in greater long-term benefits--it was a more comprehensive program than the Perry Preschool program--we assumed that these benefits would represent the minimum that may be expected. Benefits include: childcare, the reduced cost of special education and/or grade retention, earnings increase, and welfare and crime cost reduction. Childcare benefits were estimated based upon a national survey of working mothers, which indicated that the average mother spends $2,878 (1986 dollars) per year on daycare. We used this figure for Years 2 through 5, and a slightly lower figure, $2533, for Year 1 because children were not actually enrolled for their entire first year--they began the program in early infancy (Ramey & Bryant 1983). Table 39 shows the value of daycare in each year at 0%, 3%, 7%, and 0% interest. All other benefits were based upon a benefit-cost analysis of the Perry Preschool program (Barnett, 1985).

Table 38

Cost of Abecedarian Program compared to public and private daycare settings (1986 dollars)

Year	Abecedarian	Public Pre-School	Daycare
1	$7,433	$7,719	$5,200
2	$7,802	$9,090	$5,075
3	$7,943	$9,545	$5,036
4	$7,943	$9,545	$5,036
5	$7,943	$9,545	$5,036
Total (undiscounted)	$39,064	$45,444	$25,383
Discounted Total			
3%	$36,281	$42,130	$23,606
5%	$34,609	$40,140	$22,538
7%	$33,068	$38,307	$21,552
10%	$30,972	$35,814	$20,213

Table 39

Abecedarian Childcare Benefits (1986 dollars)

Year	Undiscounted	3%	5%	7%	10%
1	$2,533	$2,496	$2,472	$2,449	$2,415
2	$2,878	$2,753	$2,675	$2,600	$2,495
3	$2,878	$2,673	$2,548	$2,430	$2,268
4	$2,878	$2,595	$2,426	$2,271	$2,062
5	$2,878	$2,520	$2,311	$2,123	$1,874
Total	$14,045	$13,037	$12,432	$11,873	$11,114

Table 40 indicates the 1986 dollar value of these benefits under a range of discount rate assumptions. Net benefits--benefits minus costs--vary according to the discount rate. It appears, however, that the Abecedarian program is a sound economic investment for the range of 0% to 5% (see Table 39) if the program were replicated in a private daycare setting and 0% to 3% in a university (Abecedarian program setting) or public school setting. There is considerable debate among economists concerning the proper discount rate to apply for social investments. Typically, it is assumed that the relavant range is 3% to 7%. However, taking a real rate of 3% or 5% may be most realistic considering the current rate of inflation. Since the Abecedarian program was an experiment, taking place at a university, costs were not meant to be constrained and may well be much lower under more competitive circumstances. Therefore, these estimates give us a good indication of the potential of this model to yield a reasonable return to society and, more importantly, change the outlook for poor and disadvantaged children.

Table 40

Benefits of the Abecedarian Program (1986 dollars)

Benefit	Undiscounted	3%	5%	7%	10%
Childcare	$14,045	$13,037	$12,432	$11,873	$11,114
Educational cost savings	9,107	5,929	4,496	3,436	2,325
Earnings increase	1,337	744	508	350	202
Welfare reduction	105	60	41	28	16
Crime reduction	2,377	1,430	1,030	748	404
Total (to age 19)	$26,971	$21,200	$18,507	$16,435	$14,061
College Costs	-1,502	-817	-550	-372	-219
Earnings increase	100,092	27,616	12,742	6,250	1,829
Welfare reduction	4,042	1,561	879	511	250
Crime reduction	6,842	2,169	1,226	716	369
Total (age 19+)	$109,474	$30,529	$14,297	$ 7,105	$ 2,229
Total Benefits	$136,445	$51,729	$32,804	$23,540	$16,290

ADDENDUM

Annual FTE Salaries (1986 dollars)

Personnel	Abecedarian (actual)[1]	Public Pre School[1]	Day Care[1]
Supervisor	$22,408	$26,201	$12,470
Director	25,268	33,261	18,737
Teacher	22,408	26,201[2]	12,470[3]
Aide/Cottage Parent	17,229	8,379	9,376[3]
Substitute	12,655	8,379	9,376
Transportation Clerk	19,172	13,947	none
Secretary	18,760	14,116	none
Social Worker	26,286	29,841	12,470

[1] Abecedarian figures include fringe benefits:

Public school fringe benefits (% of salary):

Director, Teacher - 19%

Aide - 29%

Transporataion, Secretary & Social Worker - 19%

Private Daycare - 10%

[2] Estimated by the National Education Association

[3] Estimated from National Daycare Study (Ruopp, et al., 1979), the Child Care Information Exchange Director's Survey (1983), and the National Association for the Education of Young Children Survey, (1984)

Annual Non-personnel Resource Expenditures for the

Abecedarian Program (1986 dollars)

Facilities	$ 488	(per child)[1]
Administration	$ 154,083	
Transportation		
Staff	381	
Children	4,389	
Supplies	7,855	(nursery)
	26,082	(preschool)
Miscellaneous	10,680	

[1] estimated from Ruopp, 1979 p. 119

Annual Non-personnel Expenditures for

Public Preschool (1986 dollars)

Resource	Cost
Transportation	$ 279
Food	154
Facility Maintenance	760
Debt	428
Miscellaneous	44
Special Education Administration	199
School Administration	366
General District Administration	350
Other	187
Total	$2,767

Annual Non-personnel Resource Expenditures in

Private Daycare Setting (1986 dollars)

Resource	Cost (per child)
Occupancy	$ 488
Supplies	418
Administration	474
Other	209
Donated Equipment*	11
Total	$1,600

estimated from

*NDCS p. 226

1976-77 (dollars)	1986 (dollars)
$330/center/year	$619/center/year
	or
	$\frac{\$619}{56}$ = $11/child

Products

Number in Product List	Title
5	Economic Costs and Benefits
35	Economics of Early Intervention
51	Methodlogical Issues

INVESTIGATION 6:

A QUANTITATIVE SYNTHESIS OF SINGLE-SUBJECT RESEARCH: METHODOLOGY AND VALIDATION

The systematic and objective evaluation of a body of research can be as important, or more important, than the presentation of one single experimental outcome (Pillemer & Light, 1980). Over the past decade, progress has been made in developing and refining procedures for more effective reviews of past literature. Given the recent explosion of educational and psychological research literature in the past years (Glass, 1976), as well as improved technology for obtaining a wide variety of research reports, the search for more sophisticated review procedures takes on new meaning. Pillemer and Light (1980, p. 178) describe the problem faced by typical reviewers:

> Faced with tens or even hundreds of studies on a single topic, a reviewer unarmed with systematic procedures is forced to utilize subjective criteria for deciding how to synthesize. He may choose several favorite studies, relatively well-done from a classical experimental-design standpoint. Or he may favor studies carried out by investigators he respects. In either case, his impressionistic conclusions will often differ from those of the next well-intentioned reviewer.

The limitations of such subjective review procedures have been described by others (e.g., Glass, 1976; 1977, Light & Smith, 1971; Rosenthal, 1978). In an effort to address such limitations, Jackson (1980) proposed several criteria to be used in the writing of an effective "integrative review," summarized below:

1. The topic for review should be specifically <u>defined</u> and <u>delimited</u>. This means the reviewer should define precisely the topic to be reviewed, and the aspects of the topic which will <u>not</u> be reviewed.

2. The reviewer should cite and review <u>previous review efforts</u>. Just as an experimenter should cite relevant previous experiments and describe how his/her particular experiment will be different, so should a reviewer cite previous reviews and describe how the present review will provide additional information.

3. The reviewer should <u>cite procedures</u> for obtaining research articles. Articles reviewed should not be selected simply because they were readily available. Rather, articles should be selected for specific, objective reasons. If certain articles are <u>not</u> included in a review, the criteria by which these articles were excluded should be explicitly stated.

4. The reviewer should describe common independent and dependent variables of related studies. Such a procedure shows the reader how different studies are interrelated.

5. The reviewer should examine covariance of study outcomes with study characteristics. Such an evaluation can potentially extend the review beyond the conclusions of specific individual studies.

6. Conclusions of the review should be supported with empirical data. When a reviewer makes certain conclusions at the end of a review, it should be clearly stated how and from what data sources the conclusions have been derived.

7. Finally, and most germane to the present article, reviewers should state criteria by which study outcomes are evaluated. Often, reviewers simply cite the conclusions of the individual researchers, or fail to describe how particular research efforts were determined to be successful or unsuccessful.

The issue of objectively evaluating study outcomes by some common metric has been addressed repeatedly in the evaluation of research employing experimental "group" designs (see Pillemer & Light, 1980, for a summary of such procedures). Initially, studies were evaluated with respect to reported "statistical significance" of individual studies. If several treatments, for example, were being compared, the treatment which had resulted in the highest proportion of "statistically significant" outcomes could be considered the treatment of preference. Such a procedure is problematic, however, because of the relation statistical significance bears to sample size, and because significance indices say little about the absolute magnitude of particular treatments. An improvement on this method has been thought to be the calculation of a standardized mean-difference "effect size" (Glass, McGaw, & Smith, 1981). Such an effect size is computed by subtracting the control from the experimental group mean and dividing the difference by the control group standard deviation. Such a computation results essentially in a Z score, and by such a standardized metric, quantitative comparisons can be made across studies in order to determine relative effectiveness of different treatments on different study outcomes, and allow for analysis of covariation between study outcomes and study characteristics. Recently, such procedures have been employed to synthesize findings from the ever-increasing volume of research studies in special education (Carlberg &

Kavale, 1980; Casto & Mastropieri, 1986; Cook, Scruggs, Mastropieri, & Casto, 1985-1986; Kavale, 1980; 1981; 1982; Kavale & Forness, 1983; 1984; Kavale & Nye, 1985-1986; Mastropieri, Scruggs, & Casto, 1985). Such syntheses have done much to integrate diverse findings from the many different perspectives and approaches to special education, and to provide important summary statements which can become the basis for future research or conceptual efforts (see Kavale & Forness, 1985).

Although such procedures have done much to integrate a field of diverse approaches, methodologies, and theoretical orientations, one major quantitative problem remains which has only recently received attention. Current meta-analysis techniques necessarily exclude single-case research because the differences in quantitative properties between the two types of research design. In many cases, however, single-subject research could have been excluded for other reasons. In a recent meta-analysis, Casto and Mastropieri (1986) evaluated the overall effectiveness of early intervention programs for handicapped children. Such an evaluation necessarily excluded single-subject research efforts for quantitative reasons. However, since single-subject methodology is rarely employed to evaluate the effectiveness of educational <u>programs</u> (as opposed to specific interventions), such research would also have been excluded for conceptual reasons.

Nevertheless, the problems of synthesizing single-subject research literature remain. It can easily be argued, furthermore, that synthesis of single-subject research is as important, or perhaps more important, than the synthesis of group research efforts, since single-subject research efforts depend to a large degree on replication efforts for demonstration of external validity (Tawney & Gast, 1984), and also because of the volume of single subject research which has appeared in the special education literature (Sindelar & Wilson, 1984). Finally, synthesis of single-subject research literature is important because individual studies often do not directly compare the relative effectiveness of particular treatments. Synthesis procedures, appropriately employed, could help determine whether a specific treatment is consistently effective in inducing behavior change. Such information can be potentially helpful to the field of special education.

The Search for an Appropriate Outcome Metric

The major obstacle to the synthesis of single-case research is the failure to develop a single, generally-agreed-upon outcome metric (Center, Skiba, & Casey, 1985-1986). Although most of Jackson's (1980) criteria can be met in a qualitative review (see Rutherford & Nelson, in press; Rutherford & Nelson, 1982, for some positive examples), a quantitative evaluation of the covariation of study outcomes with study characteristics cannot be made without such a metric.

Typically, single-subject research has involved the plotting of operationalized behaviors across various treatment phases in a time-series fashion. Unlike time-series analyses in such areas as econometrics and meteorology (Box & Jenkins, 1976), however, outcomes of single-subject research have generally been evaluated by "visual inspection" methods (Parsonson & Baer, 1978). Visual inspection of graphed data involves judgmental evaluations of such phenomena as baseline "trends," overlapping data between phases, and changes in variability across phases. These judgments are considered simultaneously to determine the overall effectiveness of particular treatments. Parsonson and Baer (1978) caution that findings should be sufficiently tangible that no reasonable person would dispute the outcome.

In practice, however, visual analysis procedures have often been found to be unreliable. Kazdin (1978) argued, "The problem with visual inspection is that individuals who peruse the data may not see eye to eye" (p. 638). Gottman and Glass (1978) agreed; "Clearly, the 'eyeball' test gives results that can vary from judge to judge and can conflict sharply with the results of statistical tests" (p. 199). Furthermore, when interrater reliability of those 'expert' in visual inspection procedures have been calculated, the results have been discouraging: DeProspero and Cohen (1979) reported an overall reliability of .61, while Jones, Weinrott, and Vaught (1978) reported figures as low as .39. Mastropieri and Scruggs (1985-1986) obtained reliabilities near unity, but only when rating scales had been collapsed to only three alternatives: "effective," "partially effective," and "ineffective." Such a small number provides little practical utility in the attempt to discriminate between a variety of possible study outcomes.

Even though these arguments caution against the use of visual inspection procedures as an overall outcome metric, some alternative statistical procedures have been reported by several authors. These are described briefly below.

Alternatives to Visual Inspection

Gentile, Roden, and Klein (1974) recommended the use of parametric statistical tests (such as the t-test) with single-subject data. They suggest that violation of the assumption of independence could be somewhat ameliorated by combining related treatment phases which had been separated by baseline phases. Likewise, Huitema (1985) has argued that autocorrelations in single-subject data are less common than previously suggested. According to Kazdin (1976), however, such arguments fall short of unequivocal demonstration that: (a) single-subject data are never correlated, or (b) that any obtained autocorrelation does not represent a serious violation of statistical assumptions. Finally, regardless of the exact nature of autocorrelation, the fact remains that within-subject variability is typically much smaller in single-subject research than that obtained by research employing across subject variability as an error term. Such differences result in effect sizes which are on a different scale than those computed in group research and therefore are not easily comparable. Also, the small number of observations employed in many baseline and treatment phases may result in inaccurate variability estimates.

An alternative to the analysis of variance model, a time series approach (e.g., McCain & McCleary, 1979) has been proposed for use in evaluating single-subject data (Kazdin, 1976). Rather than presuming data independence, such a model theoretically allows for the evaluation of auto-regressive components (e.g., linear trends, seasonal effects). It is then possible to compute statistics based upon this model, which could be transformed into effect sizes. Although such an approach seems promising, conditions necessary for the implementation of this model are rarely met in single-subject research. Single-subject designs do evaluate individual performance over time, a necessary component for time series analysis; however, in order to evaluate fully all auto-regressive components in the data, a relatively large number of data points are required. The exact number necessary is uncertain; however, Box and Jenkins (1976) recommended 100 data points, while Gottman and Glass (1978) recommended 50. Even Kazdin's (1976) "rule of thumb," that ten data points per phase are minimal, specifies a condition rarely met in single-subject research.

In addition to the above parametric alternatives, two nonparametric procedures have been suggested for evaluating single-subject data. The Randomization test (Levin, Marascuilo, & Hubert, 1978) employs the ranking of phase means across treatment replications in reversal designs. The "R_n" test (Revsky, 1967) can be used to evaluate multiple baseline designs by ranking performances across subjects or settings as different treatments are implemented. Corresponding probability ratios from these tests could be theoretically converted into effect sizes. Unfortunately, as with the time series model, both nonparametric procedures specify conditions rarely met in single-subject research. The Randomization test presumes several replications of treatments, while the "R_n" test requires that each of several individual subjects be assigned at random to treatment orders in a multiple baseline design.

More recently, Russell Skiba and his colleagues at the University of Minnesota (Center, Skiba, & Casey, 1985-1986; Skiba, Casey, & Center, 1985-1986) developed statistical procedures for the computation of single-subject "effect sizes" and applied them to the quantitative synthesis of single-subject literature describing nonaversive treatments of behavior problems. These researchers, following the argument of Huitema (1985) that autocorrelation is not a major problem in single-subject data, applied a piecewise regression technique (Green, 1978) which allowed the evaluation of slope as well as treatment effects. Center et al. (1985-1986) mentioned as limitations to the model the fact that: (a) three separate but interdependent effects rather than the one effect used for meta-analysis of group research, are employed, (b) single-subject "effect sizes" so computed are not comparable and may not be as meaningful as those of group research effect sizes; and (c) fitting regression models to data which "typically provide only five data points in the baseline phase" (Center et al., p. 398), can result in inaccurate conclusions. Nevertheless, the particular method of research synthesis bears further investigation. The reader is referred to Center et al. (1985-1986) and Skiba et al. (1985-1986) for further description of these procedures.

A Non-Parametric Approach

It has been suggested (Scruggs, Mastropieri, & Casto, in press) that potential problems of data independence, sample size, normality, and homogeneity can best be resolved through non-parametric considerations, i.e., by evaluation of the ordinal relationships single-case data bear with each other. And although exact transcription of specific data points in published displays is often difficult if not impossible, an evaluation of ordinal relationships is much simpler.

Mastropieri and Scruggs (1985-1986) applied such procedures to the evaluation of single-subject research in the area of preschool treatments for social withdrawal. In this investigation, three interrelated outcome measures were employed. One was a 3-point overall rating of intervention effectiveness, while the other two were measures of non-overlapping data: a "treatment" effect, which involved the computation of non-overlapping data between treatment and immediately preceding baseline phases; and an "experimental control" effect, which involved the computation of non-overlapping data between all adjacent phases. Mastropieri and Scruggs (1985-1986) reported that synthesis procedures successfully discriminated between different study characteristics, and revealed findings equivalent to those reported by the authors of the original studies. In addition, information was provided on overall study characteristics and their quantitative relation to study outcomes.

The use of three outcome metrics proved to be somewhat cumbersome, however. Analysis of results indicated that the "treatment effect" measure of overlapping data was essentially veridical to the "experimental control" effect as well as the 3-point overall rating of treatment outcome. Only the "treatment effect" measure, therefore, was employed in the second and third synthesis efforts (Scruggs, Mastropieri, Cook, & Escobar, in press; Scruggs, Mastropieri, & Cook, 1986), with no apparent loss of information. The remainder of this paper will describe the computation of the outcome metric, its justification, and specific conventions associated with its use.

Percent of Non-Overlapping Data

Systematic analysis of components of single-subject graphic displays revealed that a most important evaluative criterion of an effective outcome is the proportion of overlapping data displayed between treatment and baseline. Kazdin (1978, p. 637) suggested,

> If performance during an intervention phase does not overlap with performance during the baseline phase when these data points are plotted over time, the effects usually are regarded as reliable. The replication of nonoverlapping distributions during different treatment phases strongly argues for the effects of treatment.

Although it must be acknowledged that data overlap is not the <u>only</u> evaluative criterion which can be applied to single-case data, it is the only major evaluative criterion which can consistently be applied in the largest number of cases. (It should be remembered that the standardized effect size of group research reports is not the only possible evaluative criterion for outcomes, nor can it be confidently applied in every case.)

Proportion of overlapping data can easily be computed in the great majority of cases, and provides a good measure of treatment effectiveness in most cases (Kazdin, 1978; Tawney & Gast, 1984; exceptions are described below). In addition, it can be shown that data overlap is not insensitive to other evaluative considerations, such as baseline variability and slope changes. Figure 6, (a) and (b) show an example of how a more variable baseline results in a lower percent of non-overlapping data (PND) score; while Figure 7, (a) and (b) shows how a steeper acquisition rate results in a higher PND score than a less steep acquisition rate. One area in which measures of

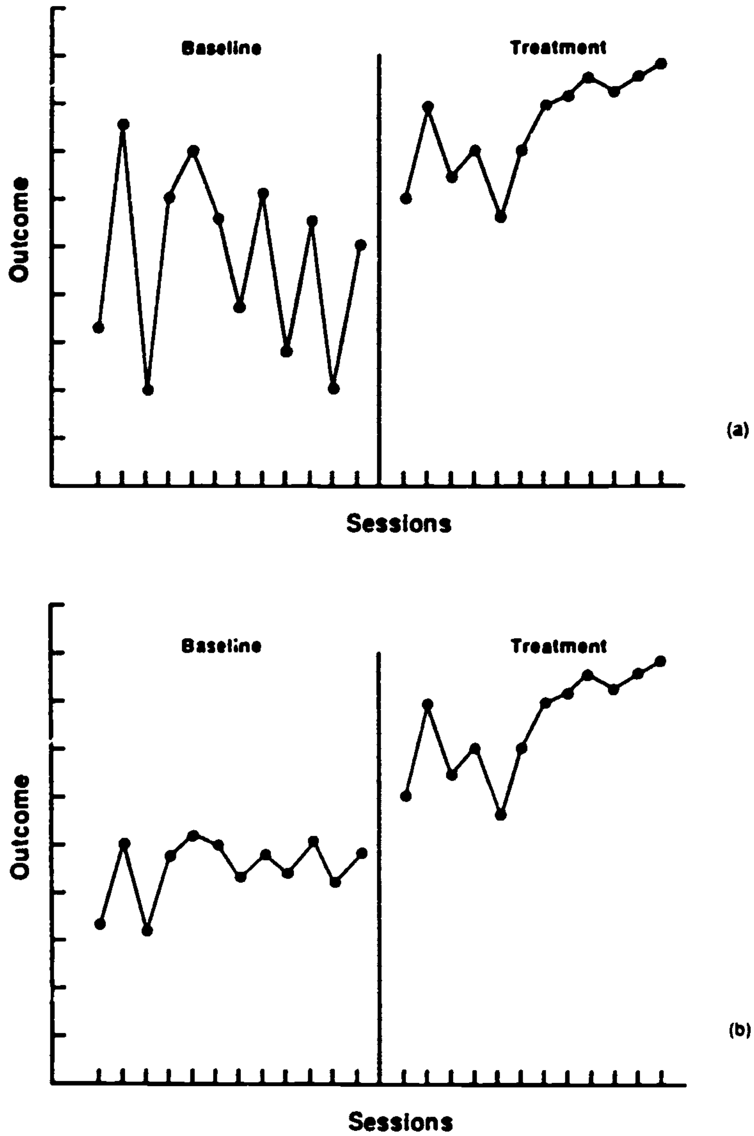


Figure 6 Examples of more (a) and less (b) baseline variability affecting proportion of overlapping data.

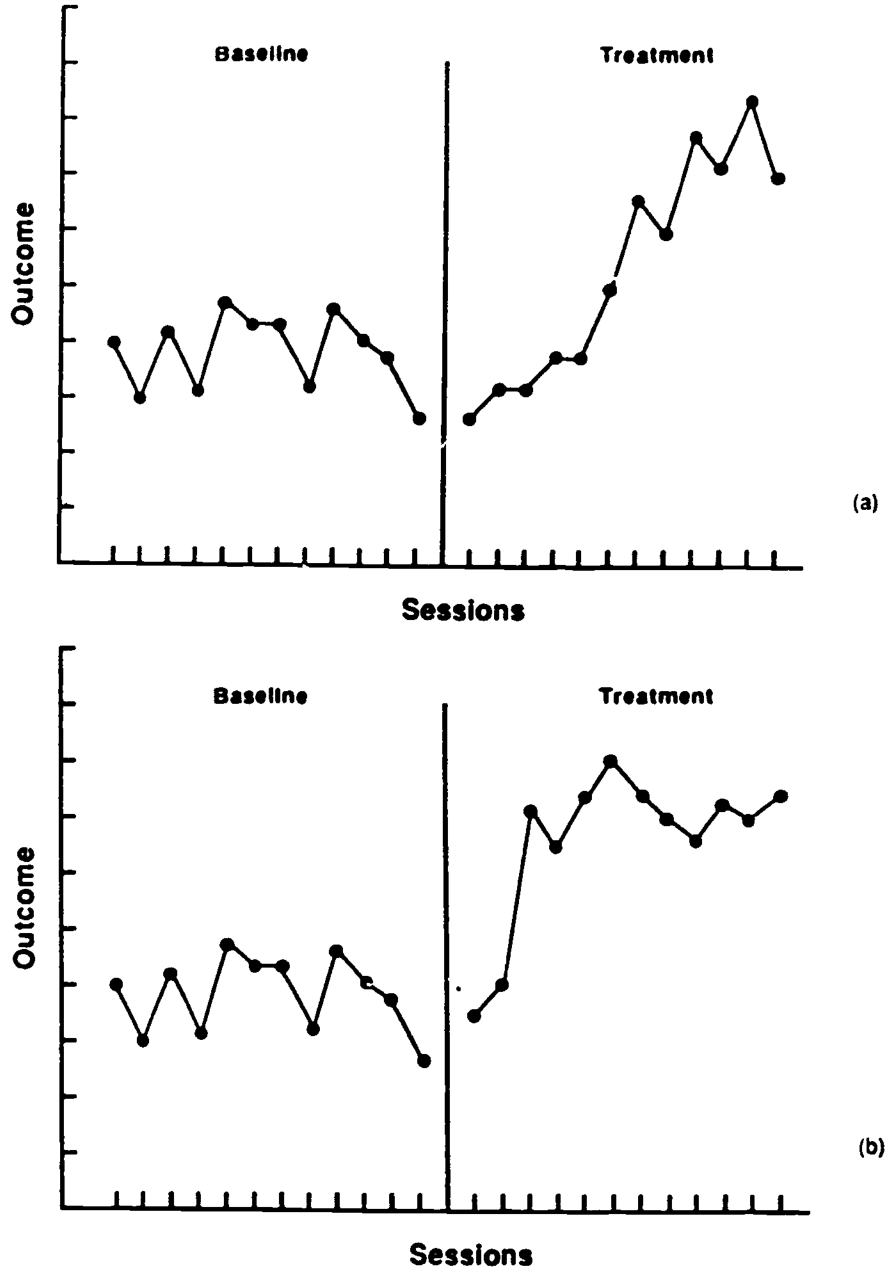


Figure 7. Examples of more steep (b) and less steep (a) treatment phase acquisition rates affecting proportion of overlapping data.

data overlap do not discriminate is among different cases in which baseline and treatment phases are all non-overlapping. In this instance, however, all treatments appear highly effective,and the inability to discriminate between different instances of this case is analogous to lack of specific discriminability of extreme cases found in non-paramatric tests of ranks (Siegel, 1956).

Proportion of non-overlapping data is easily computed. Figure 8(a) provides an example of computation of an AB phase (as in a multiple baseline design) while Figure 8(b) demonstrates the computation of data overlap in a reversal (ABAB) design. The coder simply indicates the number of treatment data points which exceed the highest

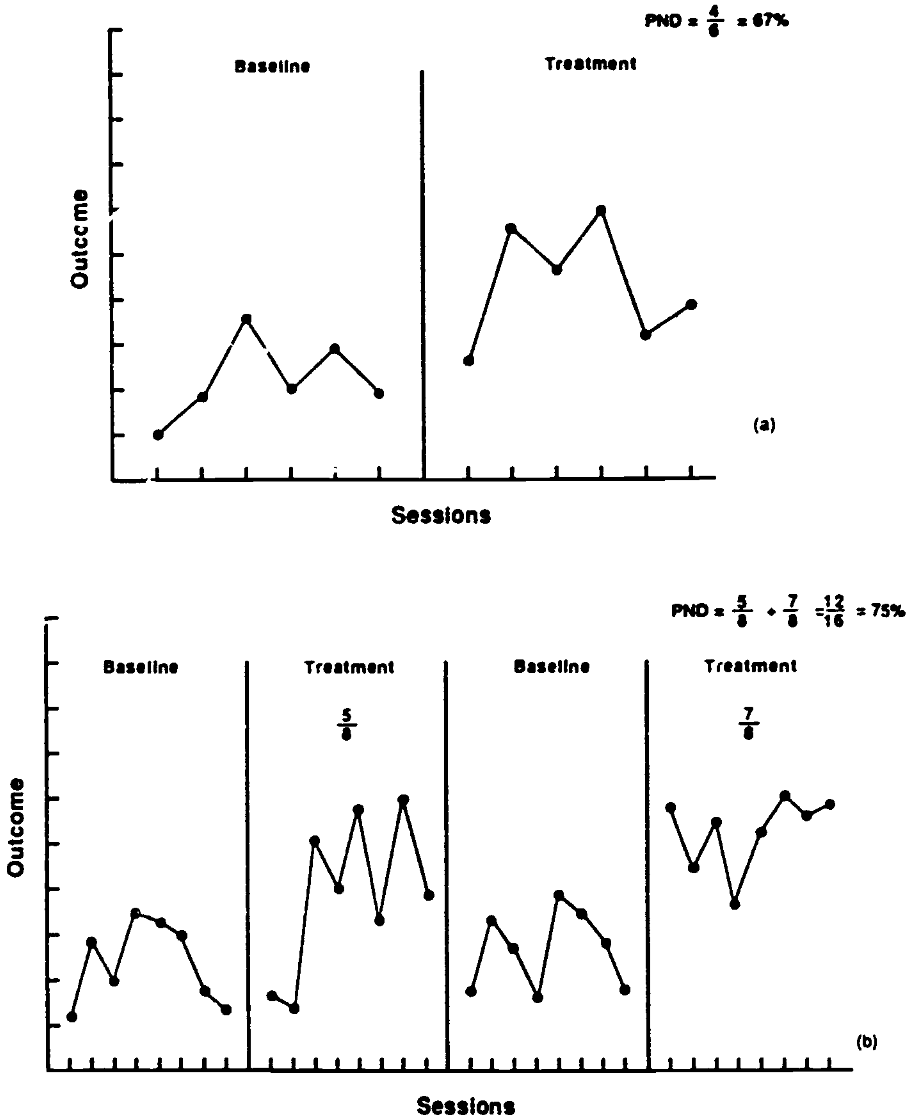


Figure 8. Computation of nonoverlapping data in an AB design (a) and an ABAB design (b).

baseline data point in an expected direction and divides by the total number of data points in the treatment phase. When phases are exactly duplicated, as in the ABAB design, measures of overlap are combined by dividing the total number of non-overlapping treatment data points by the total number of treatment data points in the two phases. When computation is completed, these outcome measures can be combined across studies to determine relative effectiveness of particular treatments. Outcomes associated with different treatments can be compared by means of non-parametric tests, such as the Mann-Whitney U test (Siegel, 1956), which make no assumptions regarding normality, homogeneity, or reference to hypothetical population 'parameters.' For example, Mastropieri and Scruggs (1985-1986) found that outcomes associated with prompted and reinforced modeling statistically exceeded those associated with modeling alone according to a Mann-Whitney U test ($p = .019$), in influencing social interaction of withdrawn preschoolers.

Although overlapping data has been argued to represent the most powerful overall outcome metric for evaluation of study outcomes, there are instances in which measures of data overlap are not appropriate. These instances, and alternative coding procedures associated with each instance, are described below.

Specific Conventions Associated With Computation of Overlapping Data

Orthogonal slope changes. Several hypothetical instances can be enumerated in which a direct measure of data overlap may not be an appropriate measure of treatment effectiveness. One more obvious case involves an "extinction" effect present in the second baseline of a reversal design orthogonical to the acquisition slope of the second treatment phase, as in Figure 9. As can be seen in Figure 9, computation of overlapping data would, in the second treatment phase, contradict a "visual analysis" interpretation. And, indeed, in many such designs, some maintenance of treatment effects may be expected during the first few observations of a "return to baseline" phase. In these cases, we have chosen to compute the proportion of overlapping data

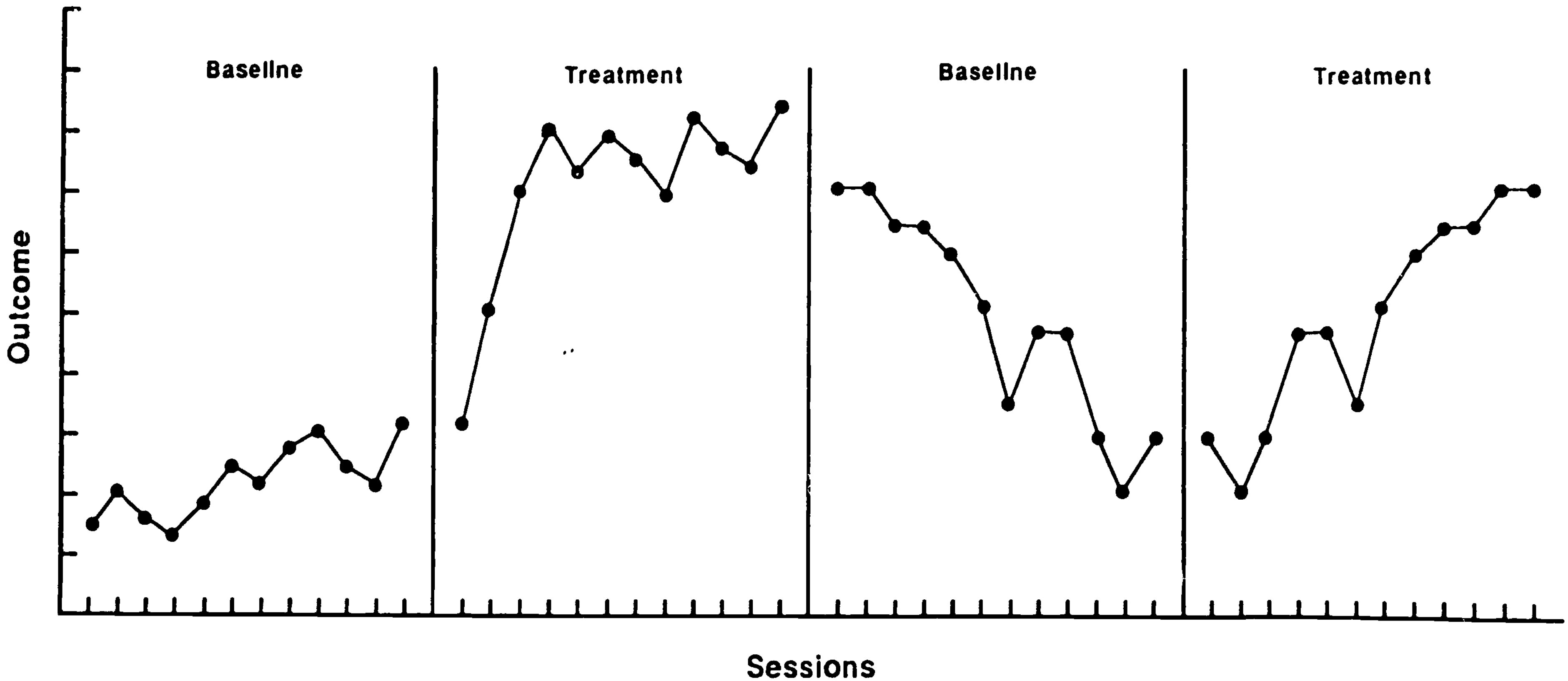


Figure 9. Example of an orthogonal slope change, in the second baseline and treatment phases.

between the second treatment and the first baseline, given the second baseline does return to the first baseline level of performance. Such a procedure allows for computation of overlapping data for the second phase of treatment without contamination by extinction patterns in a second baseline phase. It should be acknowledged here, however, that such effects in reversal designs may be less frequent than expected. Scruggs and Mastropieri (1985-1986) found no obvious instances of this phenomenon in a set of 16 studies investigating treatments for preschool conduct disorders. In fact, the computed correlation of data overlap between the first and second AB phase was $r = .74$, indicating the overall similarity of reversal phases When first and second phases did not agree, these cases were more likely to be due to failures to return to previous baseline levels of performance. In such cases, outcomes appear less reliable, and measures of non-overlapping data seem more appropriate.

<u>Inappropriate baseline trends</u>. Measures of data overlap may also yield inaccurate indices of treatment outcomes in the presence of baseline data which exhibit trends in the expected direction of treatment, such as the data displayed in Figure 10. In such cases, measures of non-overlapping data will overestimate

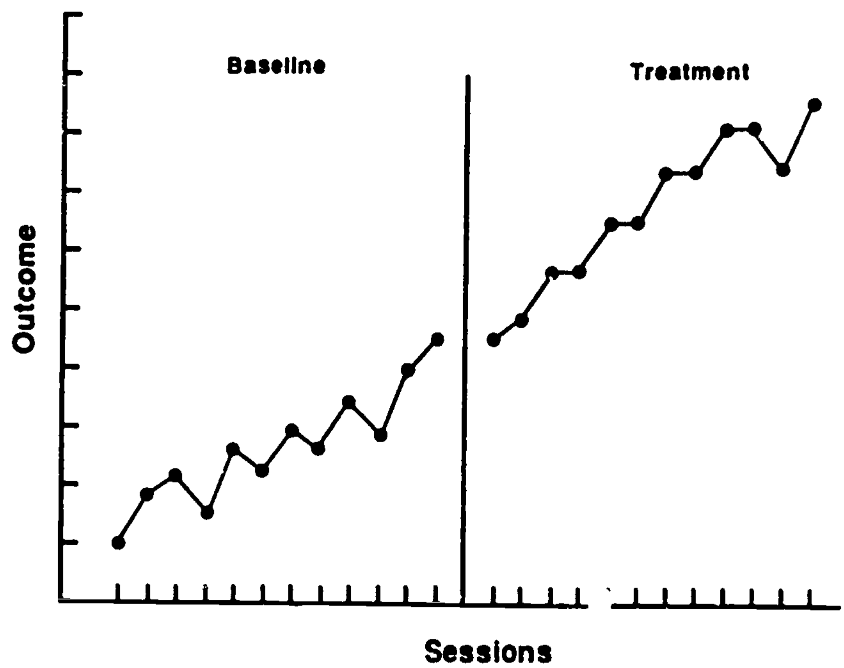


Figure 10 Inappropriate baseline trend.

treatment effects. In cases in which inappropriate baseline trends are obvious, and seriously compromise interpretability of study outcomes, there is little alternative but to exclude such data from further analysis. Calculation of relative phase slopes, one possible alternative, is problematic for three reasons. One is that the number of within-phase data points is often too low for confident estimation of slopes. Huitema (1985) reported that fully half of the phases reported in Journal of Applied Behavior Analysis over a ten-year period contained less than six observations. Such phases are easily influenced by "outliers," or random fluctuation, which can seriously compromise any attempt to calculate a slope. Computation of such "slope changes" is also problematic because of lack of empirical criteria for determining when two such slopes are truly "different." Finally, even if a treatment phase slope does appear to exceed a baseline slope, the hypothesis that the observed effect is simply an artifact of fitting two linear models to one curvilinear effect cannot be ruled out (see Figure 11). Kazdin (1978) stated,

> Despite the desirability of intervening in many situations in which baseline trends move in the direction of therapeutic change, evaluating the effect of the intervention in these situations is extremely difficult. (p. 632)

Although excluding relevant data from a research synthesis may seem in itself to be a problem, it is less so than including misleading information (see Slavin, 1984). It is also true that sucn research can be mentioned in the review in a qualitative manner. Finally, it has been seen that although inappropriate baseline trends can be found, they do not appear to be common in published single-subject research literature; in the three synthesis efforts reported by the present authors, such problems have rarely been observed.

"Floor" or "ceiling" effects. Another instance in which data overlap may not be an appropriate measure involves "floor" or "ceiling" effects in graphed data (see Figure 12). In Figure 12(a), clear treatment effect is evident despite the fact that

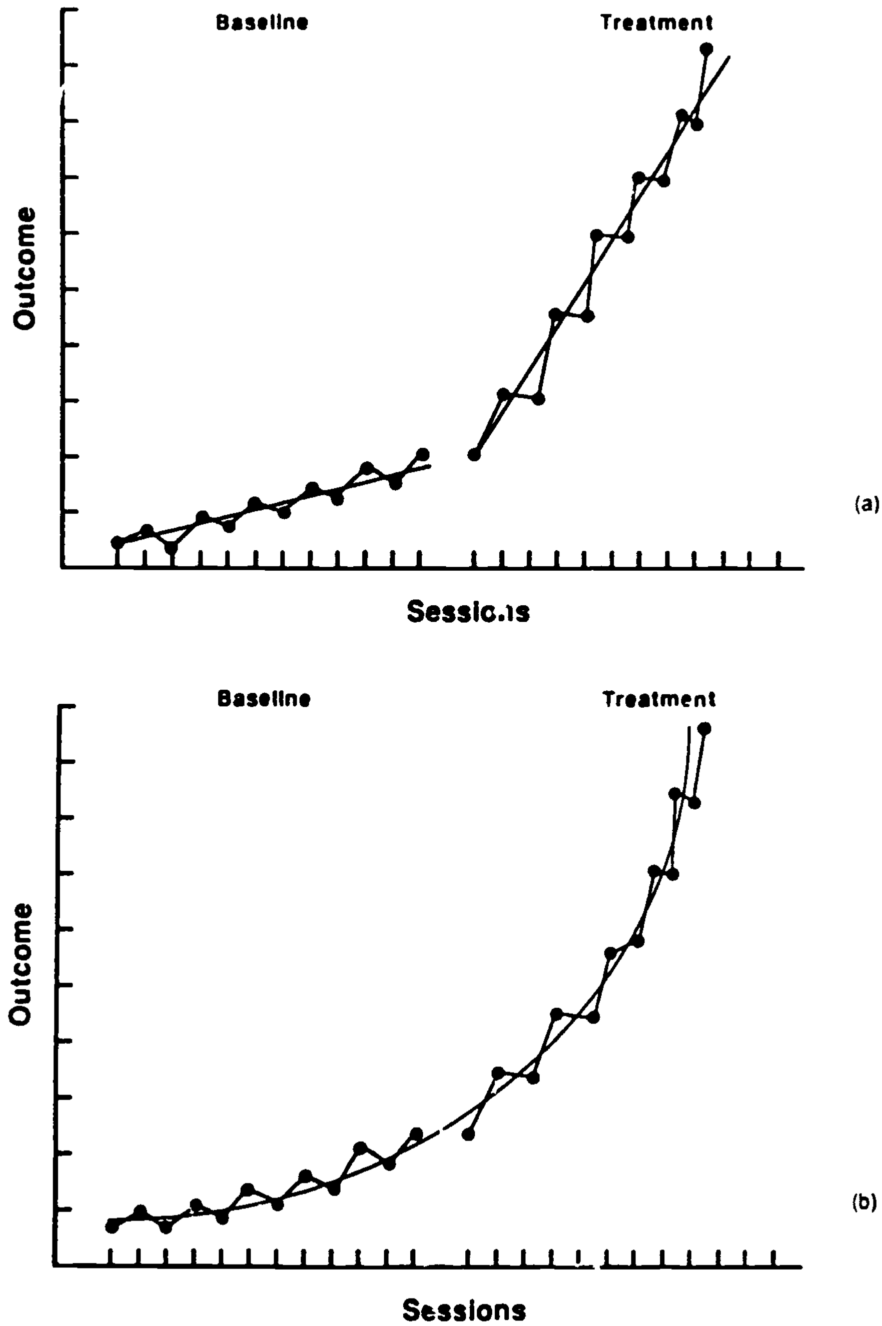


Figure 11 Example of two linear trends (a) and one curvilinear trend (b) fitted to the same AB data.

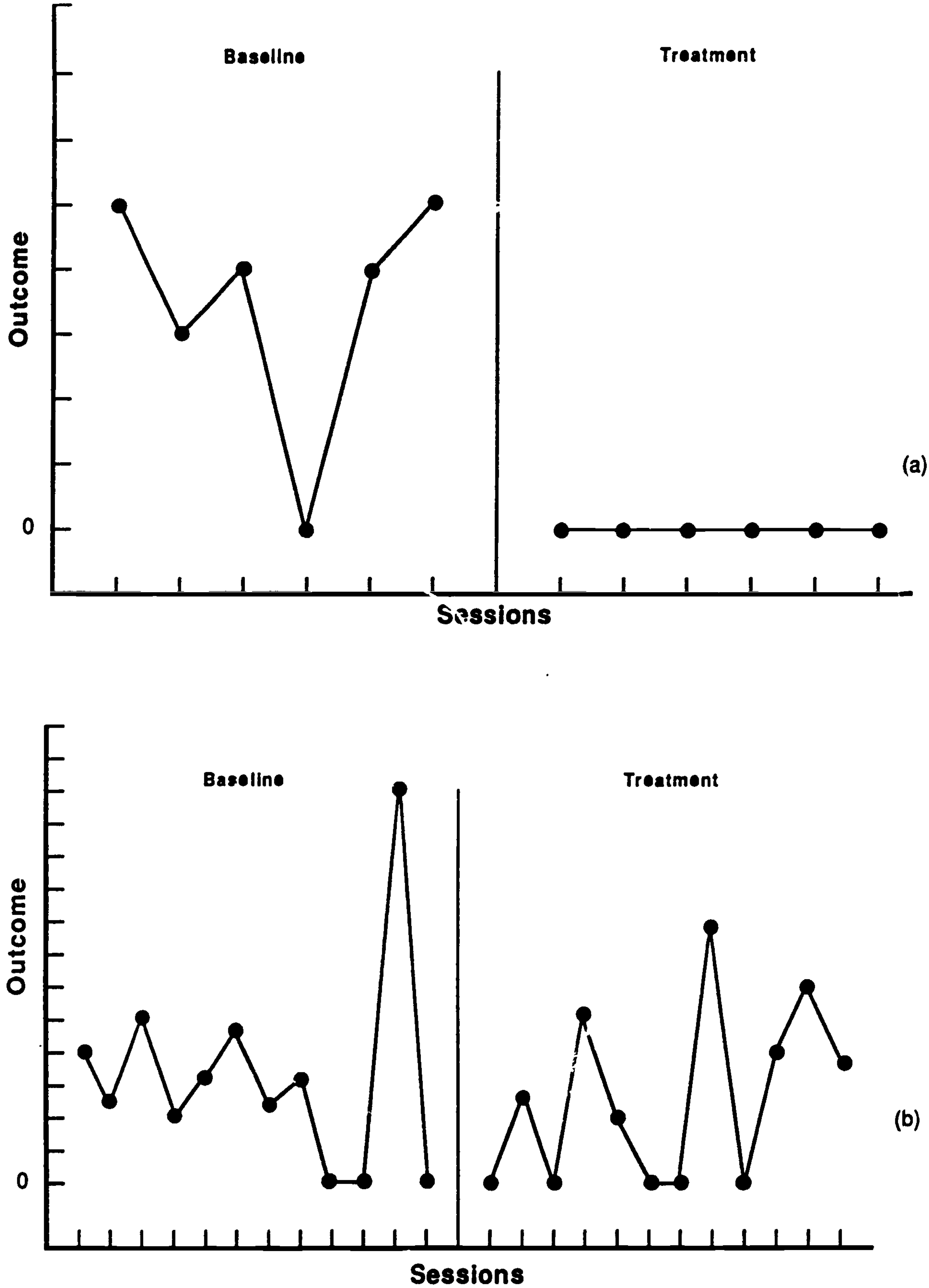


Figure 12. Examples of problematic (a) and nonproblematic (b) "floor" Performances.

all treatment data overlap at least one baseline data point. In Figure 7(b), although some restriction in range is evident, variability in both phases is similar, and it would nonetheless probably be concluded that the treatment had had little effect on observable behavior. A calculation of completely overlapping data would then be more representative of treatment effectiveness than that of the data presented in Figure 7(a). In order to discriminate between appropriate and inappropriate calculations of overlapping data in the presence of "floor" or "ceiling" effects, we developed the following rule: A measure of data overlap cannot be confidently calculated when treatment data reflecting floor levels of performance are compromised by no more than three, nor less than 33 1/3% of zero baseline data points, and baseline and treatment levels of variability are markedly different. By these conventions, a measure of data overlap would not be calculated for Figure 7(a), but would be calculated in Figure 7(b), as the baseline phase contains only 25% "floor" data points, and across-phase variability is similar. Again, although data overlap cannot be calculated for the data presented in Figure 7(a), it is still possible to describe this treatment in the review paper, and make qualitative comparisons with other, similar treatments. It should also be noted that reduced variability due to "ceiling" or "floor" effects also compromises the accuracy of standardized effect sizes in meta-analyses of group research.

A related problem, for which we have adopted no specific scoring convention, is concerned with "zero baseline," by which is meant all baseline data are equal to zero. In such cases, even a minor effect could result in relatively high levels of non-overlapping data. Such data are problematic from more than the point of view of quantitative synthesis. Such "zero baselines" seem to the present authors conceptually similar to administering college board exams to a 4-year-old. The fact that the subject will most probably score at or below "chance" levels does not mean the child has no academic knowledge or skills--it simply means the measure employed was not sensitive to the academic behaviors the child <u>was</u> able to exhibit. Likewise, in the case of "zero baselines," it is often difficult to believe that the subject was exhibiting no task-relevant behavior at all: it often seems that that observational measure was not sensitive to relevant levels of behavior which were being exhibited. In the case of "zero baselines," we have continued to calculate measures of data overlap. However, we also have coded the intervention as one in which a "zero baseline" was observed. At the time of final data analysis, calculations can be made to determine whether, for example, a specific treatment was effective only in the presence of "zero baseline" data. In such a case, conclusions regarding the ultimate effectiveness of a particular treatment could be qualified.

Unusual or complex cases. In some instances, single-subject research designs are sufficiently complex to potentially compromise a "simple" computation of data overlap, as suggested in preceding figures. Although it is not feasible to include examples of all possible research designs encountered in these single-subject literature, some general rules have been helpful. In cases in which designs have been extremely complex. it is helpful to consider the data which bear the most direct relation to the central purpose of the study. It is also helpful to carefully determine which particular treatments are being evaluated, and the relative basis for comparison of each treatment. In such cases it is important to assess reliability of coding procedures. That is, different coders should agree independently on the measures of data overlap which are to be computed. As in many of the other conventions, it has been helpful to apply "common sense" in evaluating outcomes most approp ate to the study purpose.

Summary. In this section, coding considerations have been described which can be employed in three cases for which a measure of data overlap may not be appropriate: orthogonal slope changes, inappropriate baseline trends, and ceiling or floor effects. In addition, some considerations were provided for coding complex designs. In some cases, conventions can allow for more appropriate calculations. In other cases, data overlap is simply not an appropriate measure of treatment effects.

It is thought, however, that in many of these cases, there are also problems with "visual analysis" or other methods of evaluation. Finally, past experience has suggested that measures of data overlap are highly appropriate outcome metrics in the vast majority of data displays.

Summary and Conclusions

In this report we have described a method for synthesizing single subject literature. Our initial applications of this methodology to single-subject literature have suggested that the method is versatile and systematic in uncovering covariation of study outcomes and study characteristics. The reader is referred to Mastropieri and Scruggs (1985-1986), Scruggs, Mastropieri, Cook, and Escobar (in press), and Scruggs, Mastropieri, and Cook (1986), for synthesis reports, and Scruggs, Mastropieri, and Casto (in press) for a review of findings and procedures.

Our use of overlapping data across phases as an outcome metric has freed us from the constraints of making assumptions regarding normality, homogeneity, and independence of single-case data, and allowed us to focus on one meaningful outcome which can be computed easily and applied to a large body of individual studies. We also have found that measures of overlapping data, when appropriately computed, are easy to interpret meaningfully. For example, when we report that median percent of non-overlapping data associated with generalization efforts was 33.3% (Mastropieri & Scruggs, 1985-1986), it should be obvious to anyone familiar with applied behavior analysis that such an outcome represents very poor experimental control. If, on the other hand, a mean effect size of .29 is reported (as in Scruggs, White, & Bennion, in press), such a metric is not as easily interpretable without additional qualifying information. Although proportion of data overlap is not the only consideration used in the visual analysis of single-subject data, neither is "effect size" the only consideration used when interpreting results of "group" research efforts. The PND score, however, like the effect size, is a metric which can reasonably be computed across a wide variety of studies, and, with appropriate controls, reveals a consistent, meaningful outcome. Unfortunately, no means have yet been found for the quantitative integration of single-subject and group-oriented research outcomes. It is possible, however, to compare such results in a qualitative manner (see Scruggs, Mastropieri, & Casto, in press). In many cases, however, single-subject and group research efforts address different questions which would not lend themselves to simple comparison at any rate.

Positive aspects of the present methodology notwithstanding, some caveats should be mentioned. First, there is no quantitative procedure which can replace thoughtful, careful analysis and evaluation. When conducting such an analysis, it is important to remember that methodology should serve researchers; researchers should not subordinate themselves to particular methodologies. If an individual reviewer feels that synthesis methodology has taken him or her away from the central purposes of the review, such procedures are clearly inappropriate.

Second, it must be borne in mind that all synthesis efforts are basically reviews in scope and purpose, and should not be elevated beyond this worthwhile and necessary role. Cook and Leviton (1980) appropriately cautioned against the over interpretation of quantitative outcome metrics:

> While qualitative reviews may be equally prone to bias, the descriptive accuracy of a point estimate in meta-analysis can have mischievous consequences because of its apparent `objectivity,' `precision,' and `scientism.' To naive readers, these lend a social credibility that may be based on procedural invalidity. (p. 455)

Such consequences can be avoided if the reviewer (a) acknowledges the quasi-experimental nature of such synthesis efforts, and (b) supports the overall conclusions of the synthesis with data from original research reports. It should be acknowledged, however, that the presently-described synthesis methodology is likely to attract criticism similar to that directed against previous attempts at quantitative synthesis; for example, the "mixing of apples and oranges" argument or the debate over inclusion of methodologically weaker studies. These and other criticisms have been responded to in the literature (e.g., Carlberg et al., 1984). Such objections do demonstrate that research sythesis is potentially subject to error, as is any qualitative review procedure. However, it should be acknowledged that there is nothing in a quantitative synthesis that precludes any aspect of the finest qualitative review of relevant research literature. Research syntheses can be viewed as supplements to rather than replacements for traditional reviews, and in this capacity can make an important contribution to the complicated field of research integration. As Rosenthal (1984) noted:

> The alternative to the systematic, explicit, quantitative procedures [of meta-analysis] is even less perfect, even more likely to be applied inappropriately, and even more likely to lead us to error. There is nothing in this set of meta-analytic procedures that makes us less able to engage in creative thought. All the thoughtful and intuitive procedures of the traditional review of the literature can also be employed in a meta-analytic review. However, meta-analytic reviews go beyond the traditional reviews in the degree to which they are more systematic, more explicit, more exhaustive, and more quantitative. Because of these features, meta-analytic reviews are more likely to lead to summary statements of greater thoroughness, greater precision, and greater intersubjectivity or objectivity. (p. 17)

If properly conducted, such synthesis procedures as the ones described here have a great potential for advancing knowledge and understanding of special education research.

Products

Number in Product List	Title
41	Single Subject Methodology
42	Reply to Owen White
43	Response to Salzberg, Strain, and Baer

INVESTIGATION 7:

IMPROVING PRE-ACADEMIC SKILLS IN DEVELOPMENTALLY DELAYED PRESCHOOLERS THROUGH THE USE OF A HIGHLY STRUCTURED COGNITIVE INTERVENTION PROGRAM

Introduction

The purpose of this investigation was to compare the effectiveness of a highly structured preschool intervention program to a program of lower structure. Forty-four delayed preschoolers enrolled in a Head Start Program were matched and randomly assigned to the two treatment groups. The intervention period lasted four months. Utilizing a pretest/posttest design, the children in the highly structured group achieved significantly greater gains on cognitive measures than the children in the program of low structure. Further, degree of delay and gender did not interact significantly with the degree of program structure. When age was examined, younger children scored significantly higher than older children on one of the two cognitive measures. However, this was deemed to be artifact of the scoring procedures for that test. Relevance of the findings are discussed in relation to the large body of literature on disadvantaged preschoolers which supports the use of highly structured intervention programs. This study adds to the small number of intervention studies on the delayed population by supporting the use of highly structured intervention with delayed preschool children.

Over the last 20 years, the number of early intervention programs for handicapped, at-risk, and disadvantaged children has increased dramatically. Over half of the states now mandate services for preschool handicapped children, and similar legislation is pending in several more. Head Start serves more than 400,000 children each year. The federal Handicapped Children's Early Education Program (HCEEP) has funded more than 300 demonstration projects over the last 16 years, 22 of which have been approved for national dissemination by the Joint Dissemination Review

Panel of the U.S. Department of Education, resulting in more than 2,000 replications in other sites (Roy Littlejohn Associates, 1982).

The largest federally funded program for disadvantaged children, Project Head Start, has been in operation on a nationwide basis since 1965 focusing on cognitive, language, and social-emotional development, as well as health and nutrition, and family and community involvement for children ages three to five. In addition, Head Start programs identify and provide services to handicapped children to equal 10% of each program's total enrollment (Hubbell, 1983). The national Head Start administration provides inservice training to local Head Start personnel and has developed packaged training materials related to working with delayed preschoolers (Alonzo, Morr, & Raynor, 1979; Hayden & Smith, 1979). However, the responsibility of choosing and implementing a remediation program for an individual child is left frequently with the classroom teacher.

A critical question that arises is whether Head Start teachers have the time and expertise necessary to develop individualized programs for delayed preschoolers. Apparently the opinion that they do not has predominated, and this has given rise to many different types of Head Start curricula.

An important early study of the effectiveness of preschool programs for disadvantaged children was reported in Darlington, Royce, Snipper, Murray, & Lazar (1980). This article summarized the research undertaken by members of the Consortium for Longitudinal Studies. The Consortium conducted eight separate preschool programs at a total of eleven sites. Children were re-evaluated for follow-up at ages ranging from nine through 19 years. Follow-up data included an individualized intelligence test, school record information, achievement tests and interviews with participants and parents. Although IQ scores were higher for three to four years following preschool, these effects tapered down over the next several years. However, school success was determined by decreased grade retention and less frequent placement in special education classes of treatment children as compared to the control group.

A review by Hubbell (1983) focused on the efficacy data reported on 124 separate studies utilizing Head Start children exclusively. Several of these studies investigated the effects of different curricula on cognitive development. Miller and Dyer (1975; 1983; 1984), for example, compared four different curriculum approaches and control group, and found significant differences among the four educational programs. The children who had attended the two most structured programs, the Bereiter-Englemann Program (Bereiter & Englemann, 1966) and the DARCEE Program (Gray, Klaus, Miller, & Forrester, 1966), scored higher on the Stanford-Binet Intelligence Test than children using the other models were tested through Grade 10.

On the other hand, the Huron Institute (1973) evaluated the use of the Planned Variation Program which was introduced into Head Start in 1969 to add specific curricula to existing programs. Eleven curriculum options were available for children in 29 Head Start locations, including three sites that had control groups attending no preschool. Although the Head Start experience substantially improved performance on five cognitive measures, no significant differences was found between the Planned Variation Program and the regular Head Start curriculum. In addition, no one curriculum approach stood out as being more or less effective.

The results of these two studies are somewhat equivocal. Whereas the Miller and Dyer (1975) study indicated greater gains for children using the more structured curricula, the Huron Institute (1973) evaluation did not find a relationship between the curriculum used and gains made by the children.

Another major evaluation of the effects of Head Start programs was summarized by McKey et al. (1985). This report, commonly referred to as the Systhesis Project, conducted a meta-analysis on the impact of Head Start on children's cognitive and social-emotional development and health status, as well as its impact on families and communities. The 22 page Executive Summary of this report summed up its findings on cognitive gains made by Head Start children in the following manner: Children enrolled in Head Start enjoy significant immediate gains in cognitive test scores, socioemotional test scores of former Head Start students do not remain superior to those of disadvantaged children who did not attend Head Start. However, a small subset of studies finds that former Head Starters are more likely to be promoted to the next grade and are less likely to be assigned to special education classes (McKey et al., 1985, pg. 1).

Some controversy has arisen from its results related to long term outcomes of Head Start graduates. Although the synthesis implies that Head Start benefits are short-lived, several cautions have been raised by Schweinhart and Weikart (1986). The synthesis included studies of both high- and low-quality design and certain analyses contained only few studies, making it difficult if not tenuous, to draw firm conclusions.

More recently, further support for the notion that greater program effectiveness is attained when using more highly structured curricula comes from the initial findings of a meta-analysis of the efficacy research done in the area of early intervention (Casto, White, & Taylor, 1983). Beginning with analysis of 64 review articles published between 1966 and 1982, the Early Intervention Research Institute at Utah State University coded over 264 primary research studies concerning early intervention with both disadvantaged and handicapped preschoolers. The major portion of the studies conducted were conducted with the disadvantaged group. One of the most consistent findings of the meta-analysis was that for disadvantaged preschoolers, more highly structured programs were associated with more favorable scores on outcome measures (McGee 1972; Miller & Dyer, 1975; Wooden, 1976.)

In the Casto et al. (1983) meta-analysis, as in the present study, the degree of intervention structure variable was defined as:

Very structured: 50% or more of the intervention must be based on a detailed set of outcome objectives supported by a task analysis with scripted presentation of activities and procedures and criteria for progressing to material.

Somewhat structured: 50% or more of the intervention must be organized around preconceived activities which is based on explicit scope and sequence of learning. The relation of the various parts of the curriculum should be specified, and there should be the intention for the interventionist to follow a preconceived, organized plan of instruction.

Not structured: Any intervention not meeting the criteria for 1 and 2.

Although the importance of the degree of intervention structure variable has been documented in the disadvantaged population, the limited number of studies using true experimental designs with delayed children warrant further investigation. The current study was designed with two purposes. The first was to determine the relative efficacy of a highly structured pre-academic intervention program with delayed preschool children as compared with a traditional preschool intervention program with a low level of structure. It was hypothesized that there would be no significant differences between the gains achieved on cognitive measures for children enrolled in a highly structured intervention program as compared to children in a program or low structure. The second purpose of this investigation was to determine if such child characteristics as degree of delay, age, and gender covary with the degree of program structure. In this regard it was hypothesized that no significant differences between gains achieved by high and low structure groups on cognitive measures would be observed between: a) children with greater degrees of delay as compared to children delayed to a lesser extent; b) boys in comparison with girls; and c) younger and older children.

Method

Subjects and Design

Fifty developmentally delayed preschoolers (31 boys and 19 girls) between the ages of four and six years (X = 54.7 months) who were enrolled in a county Head Start Program in a western state served as subjects for the study. During the four-month intervention period, all of the subjects participated in the Head Start program 4-1/2 hours per day for 4 days a week. Prior to the onset of the study, all children (n =243) enrolled in the county Head Start Program had been administered the Boyd Developmental Progress Scale (Boyd, 1974). Those children exhibiting a developmental lag of at least 6 months in the cognitive area (n = 50) were identified for inclusion in the study. These children were then pretested using the Goodenough-Harris Draw-A-Person Test (Harris, 1963), and the CAMS Pre-Academic Test (Casto, 1979). To provide systematic control, children were then matched on chronological age (months) and CAMS pretest scores before being randomly assigned to the experimental (n = 25) group or the comparison group. The experimental group consisted of 14 boys and 8 girls, while the comparison group was composed of 15 boys and 7 girls. Following the 16-week intervention period, 22 of the 25 matched pairs were posttested. Attrition was due to one member of each of the three pairs having moved from the area. The mean CAMS score was 78.5 (SD = 11.43) for experimental group participants and 78) (SD = 11.01) for comparison group subjects. Both groups were scheduled to receive intervention procedures twice weekly for 1/2 hour as part of their Head Start experience as described in greater detail below.

Measures

The subjects in this study were administered both criterion referenced and standardized tests on a pretest/posttest basis. Pretesting was conducted by two trained examiners, while posttesting was conducted by four different trained individuals.

The criterion-referenced test used was the CAMS Pre-Academic Placement Test. It served as an indicator of the cognitive skills attained by the subjects during the period of intervention as well as a means to place the experimental children into the high structure group at the correct entry point.

The Goodenough-Harris Draw-A-Person Te t (Harris, 1963) was administered as a standardized pretest/posttest measure. It is a brief, non-verbal test of intelligence that was administered individually to all children. Its purpose is to measure intellectual maturity which Harris defined as the ability to form concepts of an abstract character. Examinees are required to draw a picture of a man, woman, or themselves, which yields a score as a deviation quotient. Reliability is reported as follows: test-retest 0.68; split-half 0.89 (Sattler, 1982).

The Slosson Intelligence Test (Slosson, 1984) was used as a posttest only, serving as an additional measure for assessing gains. It is an age scale that provides mental age or a ratio IQ. Test-retest and split-half reliability are both reported as .80 - .90 for children between four and six years of age (Sattler, 1982).

High Structure Curriculum

For this study, the CAMS Pre-Academic Program was utilized as the highly structured curriculum. The CAMS (Curriculum And Monitoring System) Program (Casto, 1979) has been shown to be effective in increasing the standard scores of delayed preschoolers (Casto, 1980). The CAMS materials were developed initially as a method of assessing and providing curricula for delayed preschoolers from birth to five years of age in rural-remote areas. The program was designed so that after the child was assessed using the CAMS placement tests, s/he could be placed directly into the curriculum at the level indicated by the placement test. Because the curriculum is developmentally sequenced and task analyzed, personnel can be trained to carry out the intervention procedures and require only minimal supervision.

There are six CAMS programs in the areas of: (a) Pre-Academic Skills, (b) Receptive Language, (c) Expressive Language, (d) Motor Skills, (e) Self-Help Skills, and (f) Social-Emotional Development (Casto, 1979).

High Structure Intervention Group

Three itinerant teachers were trained in the use of the CAMS curriculum during a 3-hour training session and provided the intervention. The CAMS pretest was used to place each child at the appropriate levels of the CAMS curriculum. The program

167

offered the children in the experimental group 1/2 hour of direct instruction two times weekly by an itinerant teacher with a teacher-child ratio of 1:1 or 1:2. In actuality, the children averaged 1.5 sessions per week with a range of 9-22 sessions during the 4-1/2 month intervention period. As the child met criteria for mastery for one CAMS objective, s/he advanced to the next objective as indicated by the placement test results.

Low Structure Intervention Group

This group received the traditional Head Start intervention for delayed children provided by their regular Head Start teacher. The choice of intervention procedures was selected by the classroom teacher and was not specified by the researcher. The procedure consisted of identifying deficit areas from each child's test results and providing extra practice on those skills on the average of one to two 20-minute sessions per week. This form of intervention corresponded to Casto's (1983) low structure definition.

Both Groups: Posttest

Posttesting was conducted using four trained examiners who were "blind" to the child's treatment condition and pretest results. Posttest instruments included the CAMS Pre-Academic Test, the Draw-a-Person Test, and the Slosson Intelligence Test (Slosson, 1984) (Slosson-inflated).

Results

Several statistical procedures were utilized to analyze the pretest/posttest data collected in the study. Initially, t-tests for independent means were calculated. To account for pretest differences on certain measures, analyses of covariance were performed. In addition, the calculation of effect sizes, as referenced to earlier (Glass, 1976), was used to determine if the intervention procedures were educationally significant. Table 41 depicts all pre- and posttest means, standard deviation and independent group t-tests.

Table 41

Means, Standard Deviations, and T-Tests for Treatment Groups

	Pretest			Posttest		
	High Structure Group	Low Structure Group		High Structure Group	Low Structure Group	
Test	X (SD)	X (SD)	t	X (SD)	X (SD)	t
CAMS	78.50 (11.43)	68.9 (11.01)	-0.12	92.50 (3.94)	89.00 (7.12)	0.04
DAP	84.32 (11.87)	95.41 (15.90)	-2.62*	93.73 (11.37)	93.59 (11.90)	0.01
Slosson	-------	-------	------	107.4 (9.	105.77 (12.84)	0.631

CAMS = CAMS Pre-Academic Placement Test

DAP = Goodenough-Harris Draw-A-Person Test

Slosson = Slosson Intelligence Test

[a]n = 22 subjects

*$p < 0.05$

On the Draw-A-Person pretest, the children in the high structure group obtained a mean score of 84.32 (SD = 11.87), whereas the students in the low structure group scored a mean of 95.41 (SD = 15.90). An independent group t-test indicated that children in the low structure group scored significantly higher on the Draw-A-Person pretest than did children in the high structure group, $t(42) = -2.62$, $p = 0.012$. However, on the Draw-A-Person posttest, an independent t-test revealed no significant difference between children in the high ($\bar{X} = 93.73$, SD = 11.36) versus the low ($\bar{X} = 93.59$, SD = 11.95) structure groups, $t(42) = 0.01$, $p = 0.969$.

To account for these pretest differences on the Draw-A-Person Test, three additional analyses were undertaken. Gain scores were calculated to account for the pretest differences in the two treatment groups. Table 42 depicts the Draw-A-Person gain score data. Inspection of Table 42 indicates that the high structure group achieved a mean gain of 9.41 IQ points (SD = 8.86), whereas the traditional group showed a mean decline of 1.81 IQ points (SD = 10.34). An independent group t-test on these gain scores showed a significant difference which favored the high structure group, $t(42)=3.87$, $p<0.001$.

Table 42

Means, Standard Deviations, and T-Tests on Draw-A-Person Gain Scores

	$\bar{X}$	SD	t	p
High Structure[a]	9.41	8.86	3.87	<0.001*
Low Structure[a]	-1.81	10.34		

[a]n = 22 participants

*$p < 0.05$

Next, the magnitude of effect attributed to each intervention was estimated by using a standardized mean difference effect size, defined as $(X_C - X_E) \div SD$ (Glass, McGaw, & Smith, 1981). This "effect size (ES)" measure is essentially the difference between experimental and control groups measured in z-score units, and has been used in recent years to describe the impact of educational programs (Cohen, 1977; Tallmadge, 1977). Calculation of an effect size on the Draw-A-Person pre- and posttests produced as ES =0.80. Tallmadge (1977) has stated that an effect size of 0.25 or greater can be considered to be educationally significant.

An analysis of covariance was performed using the degree of structure as the independent variable, the Draw-A-Person posttest as the dependent variable, and the Draw-A-Person pretest as the covariate as indicated in Table 43. A main effect was found for the degree of structure, $F(1,41)=6.99$, $p=0.012$, suggesting that after adjusting for pretest differences, the high structure group achieved significantly greater scores than the low structure group.

Table 43

ANCOVAs Using DAP Pretest and Covariate

	DAP Posttest		Slosson Posttest	
	F	p	F	p
Degree of Structure	6.99	0.012*	6.69	0.013*
Degree of Delay	1.88	0.178	0.98	0.329
Structure and Delay	0.86	0.771	0.20	0.657
Degree of Structure	5.93	0.020*	6.78	0.013*
Gender	2.04	0.161	0.57	0.813
Structure and Gender	0.24	0.628	2.23	0.144
Degree of Structure	7.16	0.011*	6.32	0.016*
Age	5.21	0.028*	0.28	0.610
Structure and Age	0.67	0.797	0.26	0.969

*$p < 0.05$

The CAMS Placement Test revealed a mean gain of 11.05 percentage points (SD = 10.87) for the high structure group, whereas the low structure group scored a mean gain of 6.78 percentage points (SD = 9.36). While children in the high structure group showed a tendency towards greater gains on the CAMS Pre-Academic Test than did children in the low structure group, these differences were not significant, $t(42) = 1.40$, $\underline{p} = 0.170$. Calculation of an effect size, however, produced an ES=0.42, making this gain educationally significant (Tallmadge, 1977).

On the Slosson Intelligence Test, which was administered as a posttest only, the high structure group achieved a mean IQ of 107.41 (SD = 9.30), and the low structure group attained a mean IQ of 105.77 (SD = 12.84). An independent group t-test revealed no significant differences between the two groups $t(42) = 0.48$, $\underline{p} = 0.631$. Because of IQ's obtained for the two groups in the Draw-A-Person pretest were significantly different, an ANCOVA (degree of structure X Slosson Intelligence Test) using the Draw-A-Person pretest as the covariate, was performed to account for possible pretreatment differences on the Slosson.

The adjusted Slosson IQ mean for the high structure group was calculated to be 110.30, whereas the low structure group mean Slosson IQ was 102.88. This procedure depicted in Table 43, produced a significant difference between the two groups $F(1,41) = 6.69$, $\underline{p} = 0.013$, favoring the high structure treatment group.

Additional ANCOVAs included degree of delay, gender, and age as in dependent variables. The children were divided into two groups to evaluate the effects of delay. Upon using the total sample mean on the Draw-A-Person pretest to divide the group ($X = 89.84$, $SE = 2.25$), the degree of delay was defined as: $IQ \leq 88$ = greater delay; $IQ \geq 92$ = lesser delay. A 2 (Structure: high vs. low) X 2 (Delay: high vs. low) analysis of covariance, using the Draw-A-Person posttest as the dependent variable and the Draw-A-Person pretest as the covariate produced a main effect for degree of structure, $F(1,36) = 8.50$, $\underline{p} = 0.006$, but did not reveal a main effect for degree of delay, $F(1,36) = 1.88$, $p = 0.178$. No significant interaction for degree of structure by degree of delay was found, $F(1,40) = 0.86$, $p = 0.771$.

A similar analysis of covariance was conducted using the Slosson Intelligence Test (1984) as the dependent variable and the Draw-A-Person pretest as the covariate. No main effect was found for the degree of delay, $F(91,40) = 0.98$, $p = 0.329$, and there was no significant interaction between the degree of delay and the degree of structure, $F(1,36) = 0.201$, $p = 0.657$. By contrast, this analysis revealed a main effect for degree of program structure, $F(1,36) = 5.79$, $p = 0.021$.

A 2 (Structure: high vs. low) X 2 (Gender: male vs. female) analysis of covariance was conducted using the Draw-A-Person posttest as the dependent variable and the Draw-A-Person pretest as the covariate. No main effect was found for gender, $F(1,39) = 2.04$, $p = 0.161$, and no significant interaction was revealed between degree of structure and gender. The degree of structure, however, did produce a main effect in this analysis, $F(1,39) = 5.93$, $p = 0.020$.

The same procedure was repeated using the Slosson Intelligence Test as the dependent variable. No main effect was produced for gender, $F(1,39) = 0.57$, $p = 0.813$, and no interaction was yielded between structure and gender, $F(1,39) = 2.23$, $p = 0.144$. Yet a main effect was found for the degree of structure, $F(1,39) = 6.78$, $p = 0.013$.

Initially, the same method was used to divide groups into younger and older which was used to divide groups into greater and lesser degrees of delay. That is, the average age of the group, 54.7 months, was selected and the ages within the margin of standard error were eliminated. However, because the age range of the children was restricted to the preschool ages, this procedure reduced cell sized to such low numbers that an alternate method was utilized. A decision was then made to separate children into younger and older groups by splitting them at the median age which was 55 months of age. This essentially divided the group in half.

To evaluate the effects of age, a 2 (Structure: high vs. low) X 2 (Age: younger vs. older) analysis of covariance was conducted, using the Draw-A-Person posttest as the dependent variable and the Draw-A-Person pretest as the covariate. While significant main effects were found for both age, $F(1,39) = 5.21$, $p = 0.028$, and for degree for structure, $F(1,39) = 7.16$, $p = 0.011$, a significant interaction between degree of structure and age was not apparent, $F(1,39) = 0.67$, $p = 0.797$. The adjusted posttest means for the Draw-A-Person posttest were: younger--97.02, older--90.30. In other words, younger children made significantly greater gains on the Draw-A-Person test than the older children, irrespective of their group placement.

A second analysis of covariance was performed to evaluate the effect of age, this time using the Slosson Intelligence Test as the dependent variable. Unlike the previous analysis, this test did not produce a main effect for age, $F(1,39) = 0.28$, $p = 0.601$. No significant interaction was found between degree of structure and age, $F(1,39) = 0.260$, $p = 0.960$. Nonetheless, degree of structure again yielded a main effect, $F(1,39) = 6.32$, $p = 0.016$.

Discussion

The results of this study are consistent with the results of the Casto et al. (1983) meta-analysis findings for disadvantaged children and with the Head Start studies (Hubbell, 1983; Miller & Dyer, 1975). The findings reported here extend the results of previous research on degree of structure to the delayed population. More specifically, this study indicates that the experimental group, which was matched on age and CAMS pretest scores with a control group, made significant gains on the CAMS posttest.

These results are concomitant with the research on disadvantaged children which indicates that children in intervention programs of high structure have made significantly greater gains (Miller & Dyer, 1975; Mundy, 1973; Wooden, 1976).

Moreover, these findings provide support for the use of highly structured programs with the delayed population. Nonetheless, the body of research on delayed children is much smaller than that in the area of disadvantaged. A recent summary of this research by Casto and Mastropieri (1986) indicated that: "When the data is considered from all studies, there appears to be little difference in programs with various degrees of structure. When the effect sizes are adjusted, there is a trend which favors the more structured programs, but the data are inconclusive" (pp. 418-419). The present study adds to this smaller body of research on delayed children and clearly supports the effectiveness of even short-term highly structured interventions.

Analysis of the effects of the degree of delay, gender, and age revealed few significant findings. The results of this investigation support the use of highly structured interventions with the delayed population, and this does not appear to be affected significantly by the child's degree of delay. These findings are important to a preschool teacher who often has limited time and resources in providing appropriate intervention for delayed children. The results suggest that a highly structured intervention for a delayed child can result in significant improvement.

The variable of gender did not have a main effect for the posttest gains. Boys and girls made similar gains following the two types of intervention programs. Although the sample included a greater number of boys than girls, the boy:girl ratio of this investigation is consistent with the general finding that delayed populations are composed of greater numbers of boys (Blackman, 1981). Again, the degree of structure did yield a main effect on the posttest outcomes.

While age did not yield a main effect for the Slosson Test of Intelligence, there was a main effect for age on the Draw-A-Person posttest. One possible explanation for the higher scores on the Draw-A-Person Test for younger children could be the scoring criteria used. The test was scored using a point scale which gives credit for the number of body parts included by the child in his/her drawing.

For example, a head, eyes, nose, and mouth counts as a raw score of four points; similarly, a head, body, arms, and legs is also scored as four The age equivalent for these two drawings if 4 years 0 months. The quality of the drawings was not judged. Yet there was a definite trend as the children were old - to draw with improved quality, but not include a greater number of body parts. In other words, the scoring on initial drawings may have been somewhat inflated.

Several follow-up studies on various intervention programs have revealed that IQ gains made during the preschool years tend to attenuate after the first few years of elementary school (Goodstein, 1974; Miller & Bizzell, 1984). On the other hand, the Perry Preschool Project, a longitudinal study of the effects of preschool, followed a group of disadvantaged youth into their early 20's (Berrueta-Clement, Schweinhart, Barnett, Epstein, & Weikart, 1984). Results to age 19 indicated the children who attended preschool had improved cognitive performance during childhood, improved achievement levels, decreased delinquency and teenage pregnancy, and increased employment records.

Because most follow-up studies have focused on the disadvantaged population, a reassessment of the subjects of this study would provide valuable data for future preschool program development. And, although the present study assessed only cognitive changes, it would be imperative to measure other outcome variables as described by Berrueta-Clement et al. (1984), which are more representative of life span development.

An interesting aspect of the intervention program described here concerns the degree of implementation of the intervention program. The treatment goal was two 1/2 hour sessions weekly. In actuality, children averaged only 1.5 sessions per week. Although this minimum amount of intervention time resulted in significant experimental group gains over the control group, it was apparent that the intervention had not been implemented to the degree planned in the research design.

A consistent finding with disadvantaged populations is that program intensity/duration is not related to intervention effectiveness. The meta-analysis data available for handicapped populations suggest that it may very well be an important variable for handicapped populations. Table 44 presents these data.

Table 44

Average Effect Sizes for Interventions of Different Intensity

	UNADJUSTED				ADJUSTED for differences on age at start, quality of outcome measure, and time of measurement
Total hours of intervention	All studies				All studies
	ES	S_{es}	N_{es}	($n_{studies}$)	ES
Less than 50 hours	.56	.16	22	(8)	.45
50-10 hours	.62	.12	21	(10)	.63
More than 500 hours	.86	.12	39	(10)	.88

Intensity	Hours per week ES	S_{es}	N_{es}
Less than 2 hours weekly	.59	.77	149
2 - 10 hours weekly	.71	.77	59
Over 10 hours weekly	.80	.35	28

Twelve previous reviews concluded that "longer, more intense intervention is better." The EIRI meta-analysis results also suggest that duration/intensity is a critical variable for handicapped infants and young children but not for disadvantaged populations. This same conclusion was reached by Lazar et al. (1981) based on data from the Consortium for Longitudinal Studies.

As Casto and Salehi (1986) have noted, the issue of the optimum intensity and duration required for an intervention program to be effective is important. Cost-effectiveness considerations would suggest that if a program of two hours intensity per day produces the same gains as a program of 6 hours intensity, then the two-hour program would be preferred. The data presented in this study suggests tentatively that intensity and duration may not be important variables to be considered in designing programs for handicapped preschoolers since this intervention was of low intensity and for a briefer period of time.

Degree of structure. Another consistent finding with disadvantaged populations is that more highly structured programs are directly associated with more effective outcomes (Casto & White, 1985). This conclusion is not as well supported by the data from the EIRI meta-analysis from the handicapped population. Table 45 presents these data.

When the data are considered from all meta-analysis studies, there appears to be little difference in programs with various degrees of structure. When the effect sizes are adjusted, there is a trend which favors the more structured programs, but the data are inconclusive. In this study, however, degree of structure was associated with intervention gains.

Table 45

Average Effect Sizes for Different Levels of Structure in the Intervention Curriculum

	UNADJUSTED			ADJUSTED for differences on age at start, quality of outcome measure, and time of measurement
	All studies			All studies
Degree of structure	ES	N_{es}	($n_{studies}$)	ES
Very structured	.82	18	(19)	.88
Somewhat structured	.82	83	(32)	.79
Little or no structure	.83	6	(4)	.76

Summary and Conclusions

The results of this study indicate that a highly structured intervention program produced significantly higher scores on standardized cognitive measures that a program of lower structure. Other child characteristics did not significantly influence group gains. More specifically, neither degree of delay nor gender significantly affected the gains made by children in either group. Although a main effect for age was found on the Draw-A-Person Test, but not the Slosson Intelligence Test, a plausible explanation has been offered.

A preschool teacher frequently has limited resources for providing appropriate educational interventions for the delayed children in a classroom. Because this study suggests that degree of delay gender, and age did not significantly affect the gains made by the two treatment groups, the high structure variable is an important aspect for the teacher to consider when choosing an intervention curriculum. Finally, because research on the effects of preschool intervention with the delayed population represents a relatively small number of studies, this investigation makes an important contribution to the existing literature.

INVESTIGATION 8:

THE EFFECTS OF THREE LEVELS OF PARENT INVOLVEMENT WITH PRESCHOOLERS ON DEVELOPMENT AND BEHAVIOR PROBLEMS

Problem

A recent emphasis upon parent involvement in the education of children with special needs is the result of several factors. First, an increasing awareness of the critical importance of the early years of life in the intellectual and social development of the child (Caldwell, 1970). Second, disenchantment with traditional forms of child therapy in producing change that would endure outside the treatment environment or would generalize to new problems (Graziano, 1971; Mash, Handy, & Hamerlynck, 1976). Third, the belief that in many cases parents and other nonprofessionals are able to achieve educational and clinical outcomes equal to or better than those obtained by professionals (Boomer, 1982; Durlak, 1979). Fourth, the high potential for cost-effective intervention, which allow the unserved and underserved to be positively impacted (Pezzino, 1984; Shear, 1980).

Although several literature reviews have concluded that programs involving parents in early intervention are most effective than programs with no parent involvement (Goodson & Hess, 1975; Hewett, 1977; Weikart, Epstein, Schweinhart, & Bond, 1978), other reviewers report that support of parent training programs is "propelled more by a belief in the efficacy of parent education than by actual demonstrations of effectiveness" (Tramontana, Sherrets, & Authiers, 1980, p. 40). A recent comprehensive integrative review of early intervention studies (White & Casto, 1984) also casts some doubt on the popular belief that parent involvement is necessarily effective in enhancing early intervention programs for disadvantaged and handicapped populations. After examining studies of programs involving parents, nonparent involvement programs, and studies comparing different levels of parent involvement, the authors stated that "these data suggest that programs for disadvantaged and at-risk children which involve parents extensively can be effective, but they are no more effective than programs which do not involve parents" (p. 22).

The results of the White and Casto (1984) review introduce findings that appear at odds with previously accepted notions regarding the value of parent education and involvement in early childhood intervention. Unfortunately, problems associated with the available research on the subject severely limited their conclusions. Problems associated with the parent involvement research conducted to date include sample limitations, methodological flaws, and limited use of outcome measures. Importantly, the limited use of outcome child measures and the lack of parental outcome measures (White & Casto, 1984; Moreland, 1982) suggests the possibility that benefits of parent training for child intervention may be present but remain largely untested.

The problem then is, despite popular notions attesting to the validity and importance of parent-involvement interventions with special needs children, analyses of available literature have revealed a lack of consensus regarding the efficacy of parental involvement in early intervention programs. Further, available primary research studies are severely limited due to design and methodological flaws, restrictive and ambiguous definitions of parent involvement, and inadequate outcome child and parental measures.

The Importance of Early Education and Related Parent Education

For the past decade, a growing number of sources (Bronfenbrenner, 1975; Garland, 1981; Goodson & Hess, 1975) have reported that early intervention programs have enhanced the capacities of infants and young children who are handicapped or at risk for developmental delays. Additionally, programs which actively involve parents as major intervenors tend to be more successful than programs which do not require active parent participation (Bronfenbrenner, 1975; Comptroller General, 1979; Goodson & Hess, 1975).

Reeder and Casto (1984) reported that over 150 studies have been cited in recent reviews of early intervention literature (Bronfenbrenner, 1975; Comptroller General, 1979; Dudzinski & Peters, 1977; Garland et al., 1981; Goodson & Hess, 1975; Gordon, 1972; Heinz, 1979; Reisinger, Ora, & Frangia, 1976; Simensson, Cooper, & Scheiner, 1982; Weikart, 1975) which purportedly demonstrate that parents can be effective teachers of their disadvantaged, at-risk, or handicapped children.

Some of the reasons cited by Shearer (1980) and Parker and Mitchell (1980) as to why parents should be utilized as teachers of their disadvantaged and/or handicapped children follow:

1. Mothers are intuitive teachers of their children.
2. During the first TWO years of life, parents are the primary socializing agents. They are the primary models for language acquisition and development.
3. Families are already the child's natural reinforcing agents; therefore, they are particularly effective when provided with the skills necessary to teach new behaviors and rectify inappropriate behaviors.
4. The home environment provides parents natural opportunities to engage in teaching activities.
5. Parents usually have the affection and motivation to become involved with their children, especially when given training to implement intervention procedures.
6. Behaviors taught the child and reinforced by the parents tend to be maintained longer and generalize better to other people and settings.
7. Well-trained parents can help offset the shortage of professionally trained personnel.
8. It is cost-effective to train parents in intervention procedures.

In summary, parent involvement in intervention procedures has been looked upon as a critical component to the success of any early intervention program.

Other virtues of parental involvement extolled by reviewers are:

1. Training parents to be effective teachers of their own children can produce higher levels of self-esteem in both the target child and the mother and has led to a greater degree of self-confidence in mothers (Bronfenbrenner, 1975; Comptroller General, 1979; Dudzinski & Peters, 1977; Honig, 1980; Kysela et al., 1980)

2. Parental involvement in preschool programs appears crucial to formation of achievement motivation in children attending such programs (Heinz, 1979).

3. Center-based programs have been found to result in greater gains on child development measures when accompanied by a home-based, parent participation component, then center-based programs without this additional supportive element (Simeonsson et al., 1982).

4. Early intervention programs which involve a high degree of parental participation have been associated with reduced health, social, and educational problems in later years of a child's life. Fewer placements in special education have taken place, and fewer children have been involved in grade retentions than children not having received the benefits of early intervention procedures (Comptroller General, 1979).

The Feasibility of Intervening with Parents of Children with Behavior Deficits

There is an increasing trend by educators to involve parents in the education process (Boomer, 1982; Freeman & Ritvo, 1976; Norquist & Wahler, 1973). Parent involvement has been further promoted by educators who report that parents, family members, and volunteers (e.g., peers) should be used for more than babysitting and clerical duties, and that nonprofessionals often get better therapeutic results than professionals (Durlack, 1979). A primary emphasis in parent education programs for young disadvantaged or handicapped children has been on training parents as behavioral change agents. Researchers have reported several adaptations of parent education in behavioral modification training to be effective in dealing with problems such as aggression (Wiltz & Patterson, 1974) and noncompliance (Roberts & Forehand, 1978), as well as in reinforcing and shaping desirable cognitive or social behaviors (Freeman & Ritvo, 1976; Angney & Hanley, 1979), self-help skills (Marshall, 1966), and the elimination of inappropriate behaviors (Koegel & Cr ert, 1972).

A sample of the literature which attests to the teaching effectiveness of parents and other nonprofessionals (peers, siblings) even with children with severe behavioral handicaps follows.

Lovaas et al. (1973) compared autistic children that had been treated in a clinic without parental involvement to autistic children whose parents were trained to conduct therapy. They found that parent training and involvement resulted in more durable treatment gains. They further suggested that parental involvement aids generalization to nontraining conditions.

In Ludlow's (1979) study, two similar groups of parents of Down syndrome children were identified. One group received support and was actively involved in their children's training. Children whose parents were involved in training achieved higher IQ and DQ scores than similar children whose parents had not been involved.

Koegel et al. (1978) trained parents to use a discrete trial format to instruct their autistic children. Parents were taught by modeling, videotapes, and practice to (a) present discriminative stimuli, (b) use prompts, (c) use shaping, and (d) deliver consequences. The results supported the position that parents could reliably use these procedures to improve the behavior of autistic children after training, but not before.

Koegel et al. (1982) summarized the results of a more comprehensive parent training study:

1. Parent training produces better initial improvement and more durable improvement in significantly less time than direct treatment in the clinic.

2. Parent training is superior to clinic treatment because the parents are present in many different settings.

3. Parent training was associated with significant increases in daily recreation and leisure time activities.

4. Measures of psychological and marital adjustment were no different for parents who were trained compared with those who were not.

5. Naive judges rated the behavior of the autistic children training by their parents as "meaningful to community members."

According to Koegel et al. (1982), parents can be trained to be competent teachers for handicapped children and their involvement in the treatment process may be critical to initial and long-term behavioral improvements.

The above cited literature supports the feasibility of utilizing parents as intervenors with behaviorally handicapped children based on parent effectiveness issues. Additional support for parent intervention feasibility is provided in a recent review of literature on the training of parents as behavioral change agents (McLoughlin, 1982). In addition to teaching effectiveness, McLoughlin (1982) stresses that (a) problems of transferring gains from a clinic setting to the home are negated; (b) others in the family may benefit; (c) the number of change agents is increased and a minimum of professional staff can have a broad impact, and thus, treatment costs are reduced; (d) the principles and techniques for modifying child behavior are empirically grounded and are consistent with the parents' role as the major transmitter of cultural expectations; and (e) behavior modification techniques generally appeal to parents, for these seem to be based on common sense and do not assume pathological behavior of children.

Unanswered Questions About Efficacy

Despite the large number of favorable conclusions drawn about the importance, feasibility, and attractiveness of parental involvement, several reviewers have expressed concern over the lack of early intervention research studies which have employed sound methodological practices from which conclusions regarding efficacy have been drawn. As pointed out by Reeder and Casto (1984), many studies have not used random assignment of children to treatment groups. In many cases, no control groups were used. Several studies failed to use independent, "blind" data collectors, and a number of studies failed to obtain interrater reliability scores on subject variables (Ambron-Robinson, 1977; Parker & Mitchell, 1980; Simeonsson et al., 1982). Other reviewers have commented on the lack of program content description, making it difficult to analyze and replicate the research already done in this area. Most early intervention programs have been short-term and cross-sectional in nature, with few studies being conducted longitudinally to check for the maintenance of developmental gains. In several studies where an experimental/control group design was used, the participants assigned to each group came from the same geographical area, making possible intercommunications between the subjects in each group. Consequently, the reported results may well have been confounded (Ambron-Robinson, 1977; Bronfenbrenner, 1975; Parker & Mitchell, 1980). One very important observation has been made by Goodson and Hess (1975) who claim that studies which fail to confirm the positive results of the early intervention research go unreported. Consequently, those programs which have been assessed have come from a biased sample.

In looking at outcome measures, reviewers such as Simeonsson et al. (1983) and Goodson and Hess (1975) have found that a wide variety of dependent measures have been used and that many of these measures involved the use of nonstandardized instruments. Consequently, they claim that it has been difficult to make legitimate comparisons between the various studies. Casto and Lewis (1984) and Casto and White (1983) have stated that outcome measures have tended to be too narrowly focused. In their integrative review of the intervention literatures, over 40% of the studies examined used some sort of IQ measure as the dependent variable. Furthermore, Casto and Lewis (1984) found that most studies failed to document the amount of parental involvement. Consequently, they assert that the issue of how much parental involvement is required for optimal child developmental progress has not been adequately addressed.

Recently, other conclusions reached by early intervention research reviewers have been challenged. In a comprehensive review of over 2,000 early intervention articles, investigators at the Early Intervention Research Institute (EIRI) at Utah State University (Casto & White, 1983; White & Casto, 1984) have failed to confirm many of the earlier conclusions of previous reviewers. Of most relevance to this report is the fact that Casto and White (1983) found that children do not benefit more from programs that have had a high degree of parental involvement when compared with children enrolled in programs with little or no parental participation.

Purpose and Objectives

The purpose of the proposed study was to investigate the effectiveness of three levels of parent involvement with preschoolers on child outcome measures of developmental skills (i.e., cognitive, social, language, motor) and behavior and on parent measures of strengths and needs in rearing preschool children, marital adjustment, and family relationships. Specific objectives of the study were:

1. To determine whether "Parent Involvement A" (structured parent involvement in the classroom and at home) results in significantly different child and parent outcome scores than "Parent Involvement B" (limited parent involvement at home but no involvement in classroom).

2. To determine whether Parent Involvement A or B results in significantly different child and parent outcome scores than "Noninvolvement C" which consists of a no-treatment waiting list control group (i.e., children did not attend class nor were the parents "formally" involved).

Procedures

Population and Sample

Children, ages 33 to 60 months, who were living in the Cache Valley, Utah area, who were having behavior problems (as indicated by parent report), and/or who had been identified as delayed in one or more areas of developmental functioning (personal-social, adaptive, motor, communication, cognitive) served as subjects. Additionally, subjects had at least one parent who was willing and able to participate in any one of the three possible intervention conditions.

The sample was accessed through community advertisements and by mailing announcements to parents of approximately 260 preschoolers who were on existing Utah State University preschool waiting lists. All interested parents completed a parent survey form which provided demographic data as well as initial information about the child's problem or delay. Preschoolers who were eligible for the study were given a developmental screening test to further determine eligibility.

A total of 42 preschoolers and their parents were selected to participate in this study. Selection was based on the existence of behavior problems and/or developmental delays as indicated by both the parent survey and the developmental screening test. Subjects were matched prior to placement in treatment groups according to chronological age, presenting problem (behavior and/or developmental delay), and sex. Matching procedures and screening instruments are described in the design, data, and instrumentation sections below.

Limitations upon the generalizability of findings in this study from the sample to the target population reflect the makeup of the accessible population. First, this population is predominately white and of rural middle class. These population characteristics potentially limit the generalizability of the findings. Secondly, the accessible population consists of parents who are sufficiently motivated both to respond to advertisements by completing a Parent Survey form and to commit to the required time involvement of the study. This limits generalizability of results to parents with lesser motivation to participate in parent education. It should be noted that both the sample and the target population were limited to preschoolers with at least one parent who is willing and able to participate in a parent education program. This requirement will exclude from the population and sample single employed parents, two working parent households, or others who cannot meet the weekly participation schedule.

Design

Matching. Following a review of the Parent Survey responses and administration of the Battelle Developmental Screening Test, subjects were matched according to presenting problem (based on parent report and the results of the Battelle), chronological age, and sex. The end product of the matching was 15 trios of matched children. After matching was completed, the three children in each matched trio were randomly assigned into one of three treatment groups.

Description of Treatment Groups. The three treatment groups of the study represented three levels of parent involvement in preschool education. In "Parent Involvement Group A," parents and children participated in a center-based program that included a home involvement component. In "Parent Involvement Group B," children participated in a center-based program but parents were not formally involved except through limited home involvement. In "Nontreatment Group C," neither parents nor children participated in classroom or home involvement, but remained on a waiting list. Table 46 illustrates the experimental conditions associated with each treatment group. The three groups (A, B, C) are described more fully below.

"Parent Involvement A" - High Parent Involvement. The 13 children in this group participated in a 2-day-per-week, 2-1/2 hours-per-day class located on the Utah State University campus. The preschool class began in February following screening and selection procedures and continued through May, at which time posttesting began. This class was taught by one teacher who was assisted by two parent aides.

Table 46

Treatment Groups

Settings	A High Parent Involvement (N = 13)	B Low Parent Involvement (N = 14)	C No Treatment Control (N = 15)
CLASS	Child and Parent Involvement	Child Involvement	No child or Parent Involvement
HOME	Limited Parent Involvement	Limited Parent Involvement	No Involvement

The general curriculum for the classroom was similar to that found in Head Start classes and included typical preschool topics such as color naming and identification, sizes and shapes, number concepts, and telling time.

"Parent Involvement in Group A" took place both in the classroom and at home. Parent involvement included the following:

1. <u>Parent Training</u>. Parents in Group A received structured training in behavior management, developmental processes, and implementing intervention strategies. Training was based on <u>Teaching Handicapped Children: A Guide for the Trainers of Parents</u> (Fezzino & Lauritzen, 1984). Training occurred in small parent groups (e.g., eight parents per group) for 1-1/2 hours each week. Training activities were coordinated with classroom observations and home assignments.

2. <u>Parents' Classroom Activities</u>. As part of the parents' training program, parents made periodic observations and recordings of the children's classroom behaviors. Additionally, parents under the supervision of the teacher and in conjunction with their training program provided a limited amount of one-on-one instruction to their own child during class time.

3. <u>Home Involvement</u>. Formal home involvement by parents was conducted for an average of 10 to 15 minutes daily and was structured through the parent training component described above. In addition, weekly Home-School Learning suggestions were sent home with the children as part of the regular curriculum. These were suggested weekly activities that parents and children could participate in together at home.

<u>"Parent Involvement B" - Low Parent Involvement</u>. The 14 children in the Low Parent Involvement Group B also participated in a 2-day-per-week, 2-1/2 hours-per-day class. This class followed the same time schedule as the High Parent Involvement Class (Group A), except it was held on alternate days. The physical classroom and the class curriculum were the same. The class was taught by the same teacher, who was assisted by two nonparent aides. Nonparent aides also received the same structured training as the parent aides in Group A.

Parent Involvement in Group B was minimal. Parents did not participate in parent training or classroom interactions. Parents were welcome to observe the classroom through observation booths, however, observation time was not structured or otherwise required. Parent involvement in the Low Parent Involvement Group B, then, consisted of <u>Home Involvement</u>: Parent involvement consisted of the weekly Home-School Learning Sheets that were sent home with the children. The same sheets and procedures were followed as with Group A, namely, sheets were sent home with the children with suggestions of activities that parents may participate in with their child. Parent responses in terms of description of completed activities were requested on this form.

"Noninvolvement Group C". The 15 subjects in Group C did not participate in the classroom, nor did they receive structured intervention from their parents as a requirement of participation in this study. Children in Group C remained on a preschool waiting list until openings become available. The only formal involvement of Group C children and parents involved participation in screening and posttesting. All subjects were tested during the same time periods, and results of the testing were discussed with the parents following posttesting.

Data and Instrumentation

Testing. All subjects were pretested with the Battelle Developmental Inventory Screening Test prior to assignment to groups. Screening was completed by qualified testers who were "blind" with respect to the study design and the requirements. Screening was accomplished during January of 1985.

Additional pretesting included measures of a child's relationship to his/her parent's (IOWA Parent Behavior Inventory) and children's problem behaviors (Burks' Behavior Rating Scales).

Posttesting included the same parent report measures administered during pretesting and an additional assessment of parent's knowledge of behavioral principles (Knowledge of Behavioral Principles as Applied to Children). Subjects were tested with the full Battelle Developmental Inventory. Posttesting was also conducted by qualified testers who were "blind" to subjects group assignments. Posttesting was accomplished during June of 1985.

Independent Variables. The independent variable is the level of parent involvement which varied across the three groups: Group A - High Parent Involvement, Group B - Low Parent Involvement, and Group C - Nontreatment.

Actual involvement of parents and children in treatment groups was assessed with child attendance records, parent sign-in sheets, and records of returned Home-School Learning Sheets reporting weekly activities. Records of involvement were useful in determining the extent to which defined treatment group levels of involvement (i.e., high vs. low) actually occurred as planned.

Dependent Variables. The dependent measures of the study included both child and parent outcome measures. These measures were administered on a pre- and posttest basis as described above. Outcome measures and testing schedules are listed in Table 47.

Table 47

Outcome Measures and Testing Schedules

Test	Person Assessed	Schedule
Battelle Screening Test	Child	Pre
Battelle Developmental Inventory	Child	Post
Burks' Behavior Rating Scales (Preschool & Kindergarten Edition)	Child	Pre and Post
IOWA Parent Behavior Inventory	Mother & Father	Pre and Post
Knowledge of Behavior Principles as Applied to Children	Mother & Father	Post

Child Outcome Measures. The Battelle Developmental Inventory Screening Test was administered as a pretest prior to assignment to groups. The Battelle Screening Test consists of items selected from the five domains of the Battelle Developmental Inventory based on high item-domain score correlations.

The Battelle Developmental Inventory (BDI) was utilized as the major child (posttest) assessment instrument. The Battelle is a standardized assessment battery of developmental skills of children aged 0-8 and of special education populations across five broad domains: personal-social, adaptive, motor, communication, and cognitive. The inventory was normed on 671 Whites and 129 minorities from the major geographical regions of the United States. Forty-nine percent of the sample was male and 51% was female. A handicapped norm sample of 160 children was also tested. Standard Errors of Measurement (SEMs) for the subdomains of the test across age groups range from 0 to 5.47, with the majority of SEMs under 1.00. Overall test/retest and interrater reliability coefficients for the BDI across ages and domains are both reported as .99. Content validity was addressed during the development of the BDI and construct intercorrelations for the subdomains of the test range from .56 to .99. Initial criterion referenced validity scores have been established with a variety of standardized tests.

The Burks' Behavior Rating Scale, preschool and kindergarten edition, is designed to identify particular behavior problems and patterns of problems shown by children, ages 3 through 6 years. It is a questionnaire of 185 items that may be completed by individuals who know the child in question well (e.g., parents and teachers). Item test/retest reliability coefficients are reported as ranging from .60 to .96. Information on content, criterion, and factorial validity is available.

Parent Outcome Measures

The IOWA Parent Behavior Inventory (IPBI) was administered as a pre and posttest to all parents. The IPBI is designed to assess parents' behaviors and related to the parent-child relationship. Ratings are based on each parent's perception of his/her own behavior.

The Knowledge of Behavior Principles as Applied to Children Inventory (KBPAC) was administered as a posttest to all parents. The KBPAC is a 50 item self-report questionnaire which is designed to assess parents' understanding of the probable causes of and appropriate techniques for altering their child's behavior.

Results and Discussion

Effects of the high-parent-involvement, low-parent-involvement, and nontreatment conditions on children's developmental progress are reported in Table 48 which contains means, standard deviations, and F values of Battelle posttest scores.

Neither the Battelle subtest scores (personal-social, adaptive, motor, communication, and cognitive) nor the Battelle total scores between the three groups were statistically significantly different.

Several reasons may be considered as to why no significant group differences occurred. The most obvious is that since no subjects from any group voluntarily withdrew from the study (with the exception of 2 Ss who moved out-of-state) and since all parents volunteered to participate recognizing that they had an equal chance of assignment to groups, one can argue that the subject pool contained parents that were highly motivated and interested in their child's education. It is, therefore, plausible that parents in the low-treatment and nontreatment groups may have been providing a sufficiently enriched environment for their children so as to mask any high- versus low- or nontreatment differences. This is sometimes described as a "value-added" effect.

Other plausible reasons for no group differences include sensitivity and stability of the Battelle and the possibility that treatment durations were insufficient to cause differences.

Means, standard deviations, and F values of posttest IOWA and Burks' scores from fathers and mothers are contained in Table 49. No statistically significant group differences were found on any of these measures. Once again, treatment duration and instrument sensitivity are plausible explanations for this lack of group difference.

Table 48

F Values for Group Comparisons on Battelle Posttest Measures (Reported in Z-Scores)

Subtests	Low Parent Group X*	Low Parent Group SD	Low Parent Group n	High Parent Group X*	High Parent Group SD	High Parent Group n	Nontreatment Control Group X*	Nontreatment Control Group SD	Nontreatment Control Group n	F	Sig. of F
Personal-Social	-.031	1.058	14	-.045	.827	13	.194	.773	14	.41	.66
Adaptive	.416	1.230	14	.362	1.007	13	.170	1.137	14	.25	.78
Motor	.433	1.075	14	.003	.[illegible]74	13	.183	.885	14	.77	.47
Communication	.072	.680	14	-.412	1.090	13	-.123	.794	14	1.18	.32
Cognitive	.210	.716	14	-.279	.859	13	-.346	.925	14	2.02	.15
Total	.190	.998	14	-.206	.866	13	-.069	.842	14	1.03	.367

*Adjusted Mean - All variable covaried on the total age adjusted Battelle Screening (pretest) and on subjects' age.

Table 49

F Values for Group Comparisons on IOWA and Burks' Posttest Measures (Reported in Z-Scores)

Variable	Low Parent Group X*	SD	n	High Parent Group X*	SD	n	Nontreatment Control Group X*	SD	n	F	Sig. of F
IOWA-Father	.361	1.223	14	-.262	1.019	13	-.019	.845	14	2.038	.145
IOWA-Mother	.062	.889	14	.117	.863	13	-.286	1.276	14	.707	.500
Burks'-Father	.012	1.157	14	.049	.678	13	-.003	1.173	14	.024	.976
Burks'-Mother	.320	1.224	14	-.117	.673	13	-.209	.947	14	1.633	.209

*Adjusted Mean - All variables were covered on their respective pretest scores.

Products

Number in Product List	Title
41	Early Intervention for Behavior Disordered Children

INVESTIGATION 9:

THE EFFECTS OF AN EARLY SENSORIMOTOR INTERVENTION PROGRAM ON THE DEVELOPMENT OF INFANTS WITH PERINATAL INTRAVENTRICULAR HEMORRHAGE

In the past few years, there has been a dramatic increase in the availability and quality of services for handicapped infants and children (Mulliken & Buckley, 1983). This increase has been accompanied by a heightened public awareness of the importance of treating the individual once a handicap has been identified, and of directing efforts toward earlier identification, prediction, and prevention of such conditions (Hunt, 1980). With Public Law 99-457 mandating early preschool services, it is anticipated that public and professional interest will continue to grow.

Our current ability to identify and appropriately treat children who are at risk for developing various handicapping conditions is limited (Mulliken & Buckley, 1983). Thus, research aimed at developing early diagnostic techniques and differential intervention programs for infants at risk for handicaps needs further attention.

One little explored, yet potentially important, indicator of later handicapping conditions is the occurrence of cerebral intraventricular hemorrhage (IVH) during the first few days of life in low birth weight (LBW) and, on rare occasions, full-term infants. Simply described, an intraventricular hemorrhage is the development of a lesion in the infant's brain which produces an abnormal bleeding from cranial capillaries which may extend into the ventricular system. The bleeding is believed to result in different degrees of neurological damage based on the severity of the hemorrhage (Volpe, 1981).

Brain-imaging procedures such as real-time ultrasonography and computed tomography (CT) scan are used to make a positive identification of IVH and to classify the hemorrhage into one of four stages of severity. Stage One IVH is the most mild form of hemorrhage, whereas Stage Four IVH is the most severe (Papile, Burstein, Burstein, & Koffler, 1978). Stage One IVH occurs in the subependyma at either the germinal matrix or the choroid plexus. Stage Two hemorrhage is a subependymal hemorrhage with extension into the ventricles, but with normal ventricular size. Stage Three IVH is a subependymal hemorrhage, with extension to the ventricles, which is accompanied by moderate to severe ventricular dilatation. Stage Four, the most severe form of IVH, is a subependymal hemorrhage with ventricular extension, with or without dilatation, plus a parenchymal lesion. Dramatic clinical symptoms such as seizures, loss of muscle tonus, cessation of breathing, and unreactive pupils may mark the onset of IVH; however, at times IVH is clinically silent (Tarby & Volpe, 1982).

Approximately 10% of all infants born in the U.S. are premature with low birth weights, and 31-55% of these infants suffer IVH (Ahmann, Lazzara, Dykes, Brann, & Schwartz, 1980; Bowerman, Donn, Silver, & Jaffe, 1984). As noted previously, IVH also has been observed on rare occasion in full-term normal birth weight (NBW) infants (Fenichel, Webster, & Wong, 1984) as well as in utero (Hill & Rozdilsky, 1984). Thus, it is readily apparent that IVH has come to be known as one of the major health problems in the newborn intensive care unit (Pasternak, Groothuis, Fischer, & Fischer, 1983).

Of infants who suffer IVH, an estimated 50-60% survive (Volpe, 1981). However, information on the future developmental progress in this population is limited and controversial (Hynd, Hartlage, & Noonan, 1984). For example, Williamson, Desmond, Wilson, Andrew, and Garcia-Prats (1982) found that 29% of IVH Stage One and Two LBW infants exhibited moderate handicapping conditions by the age of 3, whereas Papile, Munsick-Bruno, and Schaefer (1983) found that only 15% of such children could be diagnosed as having these handicaps. Both Papile et al. (1983) and Williamson et al. (1982) found that up to 80% of premature LBW survivors who experienced Stage Three or Four IVH demonstrated moderate to severe handicapping conditions, such as cerebral palsy, by the third year of life.

The problem addressed in this study is the dearth of information which presently exists about the relationship between IVH in infants and handicapping conditions exhibited later on in childhood (Hynd et al., 1984; Stewart, 1983). More importantly, there is no published research on the differential effects of early treatment programs designed to minimize or ameliorate the effects of IVH. The specific purpose of this study was to compare the developmental status of a control group of IVH infants who received only routine medical care between the ages of 3 and 12 months with that of an experimental group of similar infants with IVH exposed to an early sensorimotor intervention program based on individual need during the same period of time. This study served as the initial phase of a projected seven-year follow-up study. Data collected throughout the later phases of the longitudinal study will be used to determine the degree to which the later incidence of handicapping conditions (especially mild handicaps such as learning disabilities, behavioral disorders, educable mental retardation, language impairment, and hyperactivity) is associated with IVH during the neonatal period, and whether a sensorimotor intervention program begun in the first year of life prevents, or decreases the intensity of any of these conditions.

The specific research hypothesis tested was: **Infants who suffered IVH immediately following birth, and who were exposed to routine medical care and early sensorimotor intervention between the ages of 3 and 12 months, have developmental scores on the Battelle Developmental Inventory (BDI) that are the same as those of similar infants who received only routine medical care without intervention.**

Setting and Population

The setting for this study was a moderately large metropolitan city with a population of 300,000 persons. The study sample consisted of infants with perinatal IVH who were patients in neonatal intensive care (NICU) at the University of Utah

Medical Center (UUMC) or Primary Children's Medical Center (PCMC) between January 1985 and June 1986.

The UUMC had an average annual population of 513 infants with a 15% incidence of IVH during the study period, whereas PCMC had an average annual population of 336 infants with a 12% incidence of IVH. The incidence of IVH in both populations was much lower than nationally reported incidence rates (Volpe, 1987).

All infants in UUMC and PCMC neonatal intensive care units with a gestational age less than or equal to 40 weeks, and who were diagnosed as having IVH by ultrasonography, were eligible for the study. Severity of hemorrhage was classified by a radiologist utilizing Papile's four stage system (Papile et al., 1983).

Parents of infants eligible for the study were contacted about participation in the project via a letter from the respective NICU medical director. This letter contained general information on the nature and purpose of the study, as well as a return postcard on which the parent indicated whether or not they were willing to be contacted by the investigator for more detailed information about the study.

Those parents who gave consent were contacted by the investigator and were given an oral explanation of the study. Parents were reassured that all data would be kept confidential; that they could obtain final results, upon request, at the end of each evaluation session, as well as at the completion of the study; and, that they could withdraw their infant from the study at any time without prejudice. Informed consent was obtained after all parent questions were answered.

Infants of parents who agreed to participate were matched according to severity of hemorrhage (Stages One and Two IVH were labeled "mild", and Stages Three and Four IVH were labeled "severe") and birth weight. The treatment groups to which the infants were then randomly assigned consisted of: (1) a minimal intervention program consisting of routine medical care and referral to the NICU follow-up clinic, or (2) a more intensive intervention program which included routine medical care, referral to the NICU follow-up clinic, and an individual sensorimotor stimulation program (the

Curriculum and Monitoring System [CAMS] Motor Program) (Casto, 1979) beginning at 3 months corrected age (prematurity corrected to 40 weeks plus 3 months).

The **CAMS Motor Program** is designed to teach gross and fine motor skills to children who have delayed motor skill development. The program stimulates normal motor development patterns, beginning with raising the head and proceeding through running, hopping, and drawing squares and diagonals. The objectives of the curriculum program are developmentally sequenced beginning at birth and extending to five years of age.

The CAMS placement test identifies areas of developmental delay in the motor domain. A child development specialist administers the placement test and determines developmental level and appropriate goals for intervention in each domain. Parental concerns are also considered in developing intervention goals.

The physical therapist first assessed the child's intervention needs using the CAMS Motor placement test. The physical therapy consisted of development of sensorimotor function in the specific area(s) of need. For example, if the child displayed a motor weakness on the left side of the body, the physical therapist focused on increasing strength in that area. Clearly, children had different levels of need, and the therapist individualized treatment.

A typical intervention session would include the therapist working with the child with the parent present. The physical therapist also instructed the parent on exercises that the child could do at home, and the parent practiced and demonstrated competence on the exercises before the parent began home intervention.

The parents were told to work with the child at home at least 20 minutes per day, 5 days per week, on techniques they learned in the intervention sessions. The physical therapist telephoned the parent on weeks they did not meet to answer questions and provide guidance on implementation of intervention techniques.

Parents kept a record of the time spent with the child initiating the CAMS intervention. According to preliminary data, 90% of the parents in the initial sample completed the assigned time requirements, and provided accurate records of their intervention sessions. For those few parents who were not following the intervention criteria, the physical therapist maintained careful records of telephone calls and appointments made to get program compliance. The level of parental intervention and program involvement was used in analyzing the outcome for the children to determine if level of parent involvement affected the developmental outcome of the child.

Attendance and progress were monitored on an ongoing basis by the physical therapist's progress notes, and the CAMS placement test checklist was updated as goals are met. If a child required other equipment or services, for example, a child needed a walker or the family needed financial assistance to buy rehabilitation equipment, the physical therapist referred the family to agencies in the Salt Lake City area or obtained equipment no longer being used by other children. The physical therapist also kept a supply of equipment which she provided to parents on a no-cost basis.

The initial study sample had 24 subjects. The control group contained 14 subjects and the experimental group contained 10 subjects. Eleven (46%) subjects had a diagnosis of severe IVH (7 Grade III and 4 Grade IV) and 13 (54%) subjects had mild IVH (5 Grade I and 8 Grade II). All subjects were White, and there were a total of 12 (50%) males and 12 (50%) females in the study sample. The sample mean birth weight and gestational age were 1470 g (SD = 669) and 31 weeks (SD = 4.0), respectively.

Parents of subjects from both sites represented a cross-section of socioeconomic groups. One (4%) was unemployed, 6 (25%) were in the unskilled labor category, 9 (38%) were in the semi-skilled and blue collar category, 2 (8%) were in the semi-professional and technical category, and 6 (25%) were in the high level executive and professional category according to Duncan Socioeconomic Index (Duncan SEI) ratings (Miller, 1983).

Data Collection

Demographic data for each subject was obtained by the investigator and all ... :nostic evaluations were completed by examiners who were "blind" to the experimental and control group assignments. Information collected by the investigator from the medical records and questionnaires completed by the parent included: birth weight; gestational age; 1- and 5-minute APGAR scores; sex; race; mother's age; mother's parity (number of previous live births); number of abortions; type of birth (single, twin, or triplet); inborn/outborn status; type of delivery (vaginal versus cesarean); appropriate for gestational age (AGA) versus small for gestational age (SGA); presence or absence of hyaline membrane disease (HMD), bronchopulmonary dysplasia (BPD), ventriculo-peritoneal shunt (V-P shunt), patent ductus arteriosis (PDA), retinopathy, seizures, perinatal hearing impairment, and/or sepsis; total number of days on assisted ventilation; total days intensive care unit (ICU) status; total days intermediate care status; severity of IVH (I, II, III, or IV); total bilirubin above or below 15.0; parent socioeconomic status; marital status of mother; and total number of adults and children living in the home.

At 3- and 12-months corrected age, initial and follow-up developmental assessments, using the Battelle Developmental Inventory (BDI) (Newborg, Stock, & Wnek, 1984) were completed by "blind" examiners, and test protocols were checked for accuracy by the investigators. Mothers also completed the Parental Anxieties and Attitudes Scale (PAAS) (Field, 1978) when their infant was 3-months corrected age; the Carey Infant Temperament Scale (the Carey) (Carey & McDevitt, 1978) when their infant was between 6- and 9-months corrected age; and the Parenting Stress Index (PSI) (Abidin, 1983) when their infant was 12 months corrected age.

Instrumentation

The Battelle Developmental Inventory (BDI) (Newborg et al., 1984) is a relatively new, individually administered norm-referenced test. The test consists of 341 items grouped into the following domains: Personal-Social, Adaptive, Motor, Communication, and Cognitive. The BDI is behaviorally based and is primarily designed for identifying developmental strengths and weaknesses of handicapped and nonhandicapped children in infant, preschool, and primary programs; assessment of infants who are considered to be at-risk in any developmental area; and monitoring progress on a short- and long-term basis, as the test can be used to assess children from birth to age 8.

In addition to the BDI, the CAMS Motor Placement criterion-referenced test (Casto, 1979) was administered to those infants in the intensive sensorimotor intervention group, by a licensed physical therapist. It was given when the infant was 3-months corrected age, and was used to determine at what step the infant should begin in the individualized sensorimotor intervention program. The test items are the actual criteria from the final step of each of the 98 CAMS Motor Program objectives. At the completion of the intensive intervention program, when the infant was 12 months corrected age, the criterion test was readministered to this group of subjects. Control subjects were not given the CAMS Motor Placement test.

The Perinatal Anxieties and Attitudes Scale (PAAS) (Field, 1978) is a research instrument which was originally developed for the assessment of teenage mothers. It provides an assessment of attitudes and anxieties of the mother about herself and her infant during the pregnancy, labor, delivery, and postpartum periods. Examples of the 59 dichotomous "yes-no" questions were: "Were you angry when you found out you were pregnant?" and, "Did you want to be awake during the birth?" Mean scores reported by the author for teen (n = 90) and adult (n = 60) mothers were 20.3 and 16.5, respectively (Field, Widmayer, Stringer, & Ignatoff, 1980). Low scores represent low anxiety.

The Carey Infant Temperament Questionnaire (Carey & McDevitt, 1978) was used to assess temperament at 6- to 9-months corrected age. The authors of the scale define "infant temperament" as the emotional reactivity, or behavioral style (regardless of origin) that is displayed by an infant in the early months of life.

Results and Discussion

Scores on infant and maternal demographic and perinatal variables, the infant 3-month pretest BDI, the 6- to 9-month Carey, the infant 12-month posttest BDI, and the 12-month PSI were analyzed by the investigators in the following manner:

1. Means and standard deviations or incidence of each infant and maternal demographic and perinatal variable, and Carey and PSI ratings, were calculated for each group. Appropriate statistics (_t_ tests and tests of proportions) were used to evaluate the significance of the differences between groups.

2. Zero-order correlations were run between all demographic, perinatal, pretest BDI, posttest BDI, and PSI scores.

3. Appropriate predictor variables were entered into stepwise multiple regression equations with pretest BDI, posttest BDI, and PSI scores as dependent variables.

4. ANCOVAs were run on pretest BDI Total and subdomain scores by group. The first four predictor variables which entered on the respective multiple regression equation were used as covariates.

5. ANCOVAs were run on posttest BDI Total and subdomain scores, and Parenting Stress Index scores by group. The first five predictor variables which entered on the respective multiple regression equation were utilized as covariates.

The major results of these data (group by posttest BDI Total score ANCOVA) revealed the existence of a statistically significant positive relationship between intensive sensorimotor intervention and developmental outcome. Therefore, the research hypothesis was accepted. Table 50 displays these results. Although no other Battelle Developmental Inventory (BDI) scores were significantly different between groups, it is noteworthy that scores on the Personal-Social subdomain and the Fine Motor component approached significance. Experimental subjects earned slightly higher scores than control subjects.

Table 50

Adjusted and Observed Posttest BDI Total and Subdomain Mean Raw Scores, Observed Score Standard Deviations, and Observed Score Mean DQs by Group

	Control Group (n=14)				Experimental Group (n=10)			
Score	Adjusted	Observed	SD	DQ	Adjusted	Observed	SD	DQ
Total BDI*	153	152	19.4	74	159	164	22.0	81
Personal-Social	40	39	8.0	85	43	46	7.3	98
Adaptive	32	31	5.1	82	32	34	3.6	86
Total Motor	41	[illegible]	6.5	69	39	40	7.9	65
Gross Motor	25	26	4.6	69	24	24	6.0	65
Fine Motor	15	16	3.0	85	16	16	3.1	85
Communication	21	21	3.8	79	23	23	4.7	82
Cognitive	21	21	1.9	91	20	20	1.8	86

* Significant difference (p < .05)

As shown in Table 50, the major analyses of this study revealed a statistically significant difference between groups in performance on the posttest Total BDI which favored the experimental subjects. It is instructive that although subjects received differential treatment with respect to sensorimotor development, no significant differences were noted between groups with respect to Total Motor or Gross Motor scores, and Fine Motor scores only approached significance.

Experimental subjects earned slightly higher pretest BDI Motor scores as compared to control subjects (see Table 51), whereas at posttest, the control group had slightly superior Motor scores. These results may be explained in part by the limited ability of the Battelle Developmental Inventory (BDI) to make fine discriminations between subjects at 3-months of age. In addition, the overall posttest data indicates that 36.5% of subjects with severe IVH earned borderline DQ scores whereas only 23% of subjects with mild IVH earn similarly low scores. There were more than twice as many subjects with severe versus mild IVH in the experimental group, and the control group had more than twice as many mild IVH subjects.

An overall decline, from pre- to posttest, on observed BDI Total and subdomain DQ scores was noted. This drop was less than 1/2 standard deviation (SD) on the Personal-Social subdomain, and observed mean DQ scores remained in the average range at posttest. Observed mean DQ scores fell between 1 and 1-1/2 SDs on the BDI Adaptive subdomain and Fine Motor component, and scores dropped from the average to the low average range. Observed mean DQ scores on the BDI Total, Communication, and Cognitive subdomains fell between 1-1/2 and 2 SDs from pre- to posttest. Cognitive subdomain scores remained in the average range whereas BDI Total and Communication subdomain scores dropped from average to low average at posttest. The drop from pre- to posttest in observed mean DQ scores on the Total Motor subdomain and Gross Motor component was 2 SDs or more. Both of these scores dropped from an average to borderline DQ range.

Table 51

Pretest BDI Total and Subdomain Adjusted and Observed Mean Raw Scores, Observed Score Standard Deviations, and Observed Score Mean DQs by Group

	Control Group (n=14)				Experimental Group (n=10)			
Score	Adjusted	Observed	SD	DQ	Adjusted	Observed	SD	DQ
Total BDI	60	58	14.0	98	63	66	11.1	108
Personal-Social	17	17	4.2	97	17	18	4.8	99
Adaptive	13	12	3.4	102	12	13	3.2	104
Total Motor*	12	12	2.9	95	14	14	3.2	98
Gross Motor	9	9	1.8	96	9	9	1.7	96
Fine Motor*	4	4	1.7	99	5	5	2.1	102
Communication	10	9	2.8	103	11	11	1.6	115
Cognitive*	7	7	2.8	104	9	9	1.6	113

* Significant difference ($p < .05$)

These findings are consistent with the results of general low birth weight (LBW) and IVH outcome studies in the current literature (e.g., Escalona, 1982; Gaiter, 1982; Hack, Merkatz, McGrath, Jones, & Fanaroff, 1984; & Siegel, 1982) wherein infants demonstrated a general decline in developmental scores at 12 to 18 months when compared to an earlier assessment on which they had earned normal DQ scores.

In the IVH outcome literature, there were mixed results with respect to overall developmental outcome at 12 months or longer follow-up. Schub, Ahmann, Dykes, Lazzara, and Blumenstein (1981), at 34-month evaluation, found no significant difference between IVH and control LBW subjects, and only 12% of all subjects earned DQ scores below 70. Papile, Munsick-Bruno, and Schaefer (1983) report similar results; that is, approximately 10% of their LBW IVH sample demonstrated borderline or lower Bayley scores. Furthermore, they found that subjects with Grades III or IV IVH had significantly lower Bayley scores than either subjects with Grades I or II IVH or controls.

Other authors (e.g., Catto-Smith et al., 1985; Landry et al., 1984; Tekolste, Bennett, & Mack, 1985) investigating IVH outcome noted a significant difference between IVH and control subjects, in favor of the later. This difference was attributed to lower motor domain scores, which is consistent with the findings of general LBW research as well as those of the current study. Tekolste et al. (1985), at 36-month evaluation, noted that IVH subjects had significantly lower (though low normal) DQ scores, with differences attributed to lower motor domain scores in the IVH group. No significant differences were found between mild versus severe IVH subjects, although subjects with grade Four IVH demonstrated lower motor scores than those with Grades I through III IVH.

As mentioned previously, the Battelle Developmental Inventory has a limited number of items for assessing development in the birth to 11 month age range. As a consequence, there is a limited ability to make discriminations between subjects, in terms of development in each of the five domains. In the present study, 100% of control and experimental subjects earned average DQ scores (between 86 and 118) on the pretest BDI, and no subjects earned low average (between 71 and 85) or borderline (Between 65 and 70) DQ scores. However, on the posttest BDI, subjects in both groups earned average, low average, and borderline DQ scores. Differences between subjects may have been present at the time of pretesting, but were not identified because of the limitations of the BDI for this age group.

In the present study, subjects in both groups earned borderline DQ Motor scores, and low average to average DQ scores in other subdomains at posttest. Table 52 shows the developmental quotient scores at 12-months corrected age based upon severity of IVH. There were no subjects with Grade I IVH with DQ scores below 71, while no subjects with Grade IV IVH earned DQ scores above 85. Although the number of subjects evaluated in each of these categories was small, the trend is consistent with results of past research with IVH subjects (e.g., Gaiter, 1982; Tekolste et al., 1985). In addition, this study did not find significantly delayed scores in subjects with ventriculo-peritoneal shunts nor post-hemorrhagic hypertrophy (PHH) treated with lumbar punctures. This finding is in contrast to that of Landry et al. (1984) who found that those with PHH had the most significant delays in both mental and motor domains.

Table 52

Incidence of Posttest BDI Observed Score Total DQs by Severity of IVH

	I		II		III		IV	
DQ Score	N	%	N	%	N	%	N	%
65-70	--	--	3	38	1	14	3	75
71-85	4	80	4	50	3	43	1	25
86-102	1	20	1	12	3	43	--	--

Table 53 shows the results of a multiple regression analysis when maternal variables, perinatal variables, and demographic information were analyzed in a stepwise equation to determine significant predictors of developmental outcome at 12 months.

Table 53

Posttest Stepwise Multiple Regression Predictor Variables and Cumulative Variance by Outcome Measures

	Demographic, Perinatal, and Pretest BDI Predictor Variables					
Outcome Measures	Var.	Var.	Var.	Var.	Var.	Cum. r^2
Total BDI	FMRS	SEX	PDA	BILI	BIR	.91
Personal-Social	FMRS	LIH	ARS	BPD	GA	.92
Adaptive	BTRS	RET	GA	GMRS	CORS	.93
Total Motor	FMRS	SEP	PDA	ICU	RET	.85
Gross Motor	PDA	FMRS	SEP	SEX	BIR	.85
Fine Motor	FMRS	ICU	IO	SEP	---	.82
Communication	SEX	SEP	BILI	---	---	.81
Cognitive	INT	SEX	FMRS	CORS	---	.72
Total PSI	HMD	GMRS	ARS	ABPO	BIR	.80
Child Domain	HMD	ABPO	GMRS	ARS	PSRS	.77
Parent Domain	---					

BDI = Battelle Developmental Inventory
SEX = male or female
BIR = single, twin, or triplet
LIH = # living in home
BPD = bronchopulmonary dysplasia
SEP = sepsis
ICU = days intensive care unit
INT = days intermediate care
HMD = hyaline membrane disease
PSI = Parenting Stress Index
RET = retinopathy

FMRS = pretest Fine Motor score
PDA = patent ductus arteriosis
BILI = hyperbilirubinemia
ARS = pretest Adaptive score
GA = gestational age
BTRS = pretest Total BDI score
GMRS = pretest Gross Motor score
IO = inborn vs. outborn status
ABPO = age at posttest
PSRS = pretest Personal-Social score
CORS = pretest Communication score

These results show that severity of intraventricular hemorrhage, sex, patent ductus arteriosis, hyperbilirubinemia, type of birth (single, twin, or triplet), number of persons living in the home, bronchopulmonary dysplasia, gestational age, retinopathy, sepsis, days intensive care status, days intermediate care status, and inborn/outborn status were the only significant predictors of developmental outcome. These results are in contrast with those of prior studies on the LBW and IVH populations (e.g., Hack et al., 1984) wherein birth weight, days on assisted ventilation, and presence of hyaline membrane disease or ventriculo-peritoneal shunt significantly predicted outcome. This study also contrasted with the findings of Smith, Somner, and von Tetzchner (1982) in that a significant relationship between SES and/or birth weight and developmental outcome was not found.

Finally, this study found that subjects who participated in the sensorimotor intervention program, in addition to receiving routine medical follow-up, had significantly superior posttest Total Battelle Developmental Inventory (BDI) scores when compared to control subjects. This finding is congruent with the results of other early intervention research (e.g., Barrera, Rosenbaum, & Cunningham, 1986; McDaniels, 1977). Both of these studies presented well documented evidence of immediate benefits resulting from early intervention with at-risk populations. However, as noted by Bush and White (1983), it is essential to follow subjects longitudinally in order to evaluate the impact of participation in early intervention. Subjects in the current study will continue to receive sensorimotor stimulation and will be evaluated annually until the age of 7 in order to ascertain the benefits of participation in this particular program.

Products

Number in Product List	Title
Dissertation	The Effect of an Early Sensorimotor Intervention Program on the Development of Infants with Perinatal Intraventricular Hemorrhage
Manuscript Submitted	The Effect of an Early Sensorimotor Intervention Program on the Development of Infants with Perinatal Intraventricular Hemorrhage

INVESTIGATION 10:

PARENT AND CLINIC EARLY INTERVENTION FOR CHILDREN WITH LANGUAGE HANDICAPS: A COST-EFFECTIVENESS ANALYSIS

Most studies of early intervention report positive outcomes, although program effectiveness varies for reasons that are not yet entirely clear, but which may relate to program characteristics such as intensity and parent involvement (Bronfenbrenner, 1974; Casto & Mastropieri, 1986; Halpern, 1984; Ramey, Bryant, & Suarez, 1985). Efficacy is not the only consideration when choosing an early intervention program, however. The adoption of a new program that is only slightly more effective than an existing one might not be warranted if the additional costs are relatively high. Yet, very few economic evaluations have been conducted to examine the consequences of variations in program characteristics on economic efficiency ("cost-effectiveness") (Barnett, 1986; Barnett & Escobar, 1986).

Communication disorders are among the most common handicapping conditions of children who receive special education services (Kirk & Gallagher, 1983; Meyen, 1982). Early intervention appears to be particularly successful in remediating communication disorders. Such interventions have, on average, produced relatively large gains in language abilities (Arnold, Myette, & Casto, 1985). In addition, at least one study demonstrates that early language intervention can be economically efficient. The costs of adding early intervention for language handicapped and bilingual children to preschool and kindergarten programs were more than repaid by reductions in later special education costs (Weiss, 1981).

The primary hypotheses of the present study are that providing speech and language therapy through parents is an effective strategy for intervention and that intervention through parents is relatively inexpensive. Therefore, intervention through parents is potentially more economically efficient than traditional center-based intervention. A secondary hypothesis is that the addition of parent intervention to a center-based program results in greater efficacy. Given the relatively low cost of intervention through parents, the increased efficacy might be produced very efficiently. The rationale for these hypotheses is that parents are major contributors to the child's early language development. Previous research demonstrates that parental speech patterns influence the child's acquisition of language (Cross, 1978). Parents can have both positive and negative effects on the child's language, and language disorders may affect the parent-child relationship more generally (Bennett, 1982; Broen, 1972; Lahey, 1978; Moerk, 1972; Newport, 1976; Sachs, Brown, & Salerno, 1972).

Method

This study applied cost-effectiveness analysis to the results of an experimental design for the comparison of three alternative intervention programs and a no-treatment control group. All subjects were preschool children with mild to moderate language handicaps. The intervention programs focused specifically on the improvement of communication abilities.

Forty children and their families were randomly assigned to two groups. Half were assigned to begin the center-based clinic program immediately. The other half were placed on a waiting list for the following semester. Next, half of each of those two groups were randomly assigned to attend parent training and begin home-based therapy. Thus, four groups were formed in a two-by-two design: a center-based only group, a home-based only group, a group that participated in both programs and a group that participated in neither program.

Subjects

The sample consisted of preschool children who qualified for enrollment in the Brigham Young University (BYU) Communication Disorders Clinic in the Fall of 1985. Their families were middle income, and most parents had attended college. Children ranged in age between 35 and 59 months at entry to the study. The subjects were referred by pediatricians, friends, or family who had heard of the clinic. Children were screened for the program on language, articulation, and audition using two tests which are discussed below. Evidence of a delay of 20% or more in at least one of the three areas was required for enrollment. The study was explained to the parents of children who qualified. All agreed to participate in the study and to accept random assignment to treatment condition, with the understanding that the most effective services would be made available at the end of the experimental treatment period.

Program

Twenty children attended the BYU Communicative Disorders Clinic 2-1/2 hours a day, four days per week, for 13 weeks. The clinic day included 45 minutes of individual therapy, 35 minutes of small group therapy, and 60 minutes in a large group. Therapy was provided by a speech clinician with a Certificate of Clinical Competence (CCC) from the American Speech/Language and Hearing Association (ASHA) with the assistance of student aides. Supervision of student aides met or exceeded ASHA guidelines. All student aides were enrolled in a clinic class and had previous course work in language development and phonetics. Each student was assigned 1 or 2 children to conduct "individual" therapy with for the entire 13 weeks. Teams of 4 or 5 students were assigned to each small group (6 - 7 children) and took turns conducting therapy sessions. Large group time was spent in a regular preschool classroom with a teacher and aide.

The clinic program was guided by a pragmatics approach (Bates, 1976; Bloom & Lahey, 1978; Moerk, 1977). There was an emphasis on replicating the natural environment and on the social interaction of language. Children were heterogeneously grouped in the classroom and each child's strength was used as a model for another child's target behavior. The specific therapy received by each child was highly individualized, however.

Twenty children received therapy at home from their parents for 13 weeks. In the first 3 weeks of the intervention period, the parents (15 mothers, 4 couples, 1 father) attended four 2-1/2 hour training sessions. They attended five additional sessions over the course of the treatment period. The sessions were conducted by a second speech clinician with a CCC. The training sessions were designed to enable parents to provide a more stimulating home environment and to incorporate simple therapy techniques into ordinary activities at home. Again, the guiding perspective was that of pragmatics. Sessions taught communication theory, principles of instruction, techniques for observing and assessing children's language skills, techniques to facilitate language and speech development, and ways to create a more language-stimulating home environment. Parents were given assignments to implement with their children twice each day for 15 minutes.

The 10 children in the combined-treatment group comprised half of the participants in both the center- and home-based programs. In each type of program, the combined-treatment group was indistinguishable from the others and participated in exactly the same way as those receiving only one type of program. There was no significant difference in either childrens' attendance or parent participation in the training session between the combined-treatment group and the others. Make-up sessions were provided to ensure that parents completed all training sessions.

The remaining 10 children received no intervention during the 13 weeks. Their parents did attend the initial orientation meeting at which the study was explained and informed consent obtained. They were provided with services in the following semester.

Measures

All children were pretested and posttested using the Preschool Language Scale-Revised (PLS-R, Zimmerman, Steiner, & Pond, 1979) and the Arizona Articulation Proficiency Scale (AAPS, Fudala, 1974). The PLS-R is a measure of auditory comprehension and verbal ability. The PLS-R appears to have adequate reliability and validity based on data presented by Zimmerman et al. (1979). The PLS-R score is reported as a developmental age in months. The AAPS is a measure of articulation. The AAPS was validated for children with articulation ranging from normal to severely defective and has a high degree of reliability (Fudala, 1981). The AAPS is scored as a percentage of items completed. In this study, the measure was converted to the difference from the normal score for age so that the AAPS scores reported indicate the degree to which a child differs from normal.

Cost Estimation

The cost of each program alternative was estimated using an ingredients approach (Levin, 1983). The ingredients approach begins by identifying all of the resources used in each alternative. All program alternatives used professional staff, materials and supplies, children's initial evaluations, transportation, capital equipment, and facilities. Of course, the alternatives differed in the amounts of these used. In addition, the center-based program used student time and the home-based program used parent time to deliver treatment.

The costs estimated in a cost-effectiveness analysis are social costs, not accounting costs. In essence, that means that the analysis is concerned with the value of resources to society and not what was actually paid for them. For many resources, the social costs and the accounting costs are the same. This was true for professional staff, children's initial evaluations, materials, and supplies paid for by the clinic. Their costs were obtained from the clinic's budget and apportioned to each program based on actual use. Equipment cost for 13 weeks was estimated by annualizing (at 10%) the value of items on an inventory list for each program alternative (Levin, 1983). Facility costs were more problematic as the clinic paid a fixed overhead rate to the university in which it is housed. This fixed overhead rate may not accurately reflect the actual cost of the facility; thus, an average cost per child for facilities was estimated based on data from a national study of day care centers (Ruopp et al., 1979). Because equipment and facility costs were a very small fraction of total costs (0.3%), modest errors in these estimates would have little effect on estimated total cost.

The costs of the time of students and parents presented a different sort of problem. Although the clinic paid nothing for their time, there was an "opportunity cost" to society. Students and parents could have used their time in some other activity, and the value of that foregone activity was a cost to them. Without detailed information on each individual's alternatives, it was difficult to estimate opportunity cost precisely. Thus, two general estimates of time cost were used for both: the average wage rate, $8.74/hr., and the minimum wage rate, $3.35/hr. (U.S. Department of Labor, Bureau of Labor Statistics, 1986). Although imprecise, these two "ball park" estimates bracket a reasonable range.

The primary time costs were those of parents and student aides in delivering the interventions. Students' time costs were estimated based on the clinic schedule and staffing plan. Parents' costs were estimated based on the amount of time required for training and on a questionnaire which asked how much time parents spent in program-related activities. In addition, there was a time cost for transportation of children to and from the clinic, which parents had to provide. Parents' costs per hour of driving were estimated at the same rates as other time costs.

Results

The results are organized into two sections. One presents the findings for the effectiveness of the alternative intervention program designs (including no-treatment). The other presents the findings for the costs of the alternative designs. Together the results of the two sections provide a basis for judging the relative economic efficiency of the alternative designs.

Effects

Means and standard deviations are given in Table 54 for mother's age, child's age, and language test scores for each of the four groups at program entry. Duncan's (1955) Multiple Range Test was used to test for significant differences between pairs

Table 54

Sample Characteristics at Entry: Means by Treatment Group (Standard Deviations in Parentheses)

Variable	Home-based	Center-based	Combined	Control
Mother's Age (years)	32.20 (5.83) n=10	30.67 (4.95) n= 9	31.50 (5.78) n= 8	25.78* (4.87) n = 9
Child's Age at Pretest (months)	43.80 (8.60) n=10	45.30 (5.08) n=10	43.30 (5.50) n=10	43.00 (8.25) n= 9
PLS-R (pretest) (14.37)	42.38 (8.71) n=10	44.48 (9.70) n=10	43.50 (19.62) n=10	45.43 n= 7
AAPS (pretest)	-7.00 (12.10) n= 8	-17.90 (17.11) n=10	-17.54 (17.76) n=10	-6.32 (6.10) n= 6

*Significantly different from each of the other three groups ($\alpha = .10$)

of groups. Mother's age was the only family background available (others were father's age, parental education, and family income) for which there was a statistically significant difference across groups ($\alpha = .10$). Mothers in the no-treatment control group were younger then those in the other groups.

The effectiveness of alternative treatment strategies was investigated through a two-way analysis of covariance (ANCOVA). The dependent variables were posttest scores on the two language tests. The covariates were pretest scores and mother's age. Mother's age was considered a potential confounding influence as there is growing evidence that very young mothers are less knowledgeable and less skilled than older mothers (Field, 1981; Jones, Green, & Kraus, 1980; Ragozin, Bosham, Crnic, Greenberg, & Robinson, 1982). Within the range observed in this study, mother's age was not significantly related to any of the dependent measures nor did it interact with treatment. Results of the ANCOVA with pretest as the only covariate are presented in Table 55. The ANCOVA indicated that the home-based (parent) factor produced significant improvements in language development as measured by the PLS-R and AAPS. The center-based factor had no significant effects on either language measure, and there were no significant interaction effects. Children who received the home-based program had a 4.8 month higher mean gain on the PLS-R than those who did not. The children who received the home-based program also passed 6.6% more items correctly on the AAPS. For further information, the change scores on both tests are presented in Table 56 for each of the four groups of subjects.

Table 55

ANCOVA of Interventions' Effects on Language Development with Pretest Scores as Covariates

	APPS			PLS-R		
Source of Variation	df	MS	F	df	MS	F
Center-based (A)	1	14.24	.18	1	.14	.01
Home-based (B)	1	346.51	4.43*	1	169.54	8.44**
A x B	1	.24	.00	1	25.16	1.25
Pretest	1	4049.20	51.81***	1	3446.54	171.48***
Error	29	78.16		31	20.10	

* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 56

Mean (Standard Deviation) Change in Language Measure Score from Pretest to Posttest

Variable	Home-based	Center-based	Combined	Control
PLS-R	10.58 (6.89) n=10	5.85 (4.82) n=10	8.58 (3.92) n= 9	3.82 (4.69) n= 7
AAPS	7.34 (8.53) n= 8	4.64 (12.27) n=10	11.26 (6.16) n=10	0.83 (8.80) n= 6

There was a slight amount of attrition. One child assigned to the no-treatment control group was lost to the study before pretesting. For the 39 children who began the study, PLS-R scores were not obtained for two children at either pre- or posttest and one child at posttest only. AAPS scores were not obtained for three children at either pre- or posttest and for two children at pretest. Given the small numbers and the occurrence of most attrition at pretest, a quantitative analysis of attrition was not appropriate. However, as most of the attrition was in the control group, it could be expected to have little effect on the comparison of home- and center-based factors.

Costs

Cost estimates for each intervention program are presented in Tables 57 and 58. Only costs paid by the program are included in Table 57. It does not represent the full cost of the programs to society, because student and parent time have an opportunity cost. Two alternative estimates of social costs are shown in Table 58, one using the average wage and the other using the minimum wage to estimate cost per

Table 57

Costs for Program Alternatives: Zero Volunteer and Parent Costs

Resources	Home-based	Center-based	Combined
Personnel	$4,296	$ 9,595	$13,891
Facilities	14	156	170
Equipment	120	986	1,106
Materials/Supplies	95	300	395
Evaluation	350	350	350
Total Cost	4,875	11,387	15,912
Cost Per Child	488	1,139	1,591

Table 58

Costs for Program Alternatives: Volunteer and Parent Time Valued at Average Wage (and at Minimum Wage)[a]

Resources	Home-based	Center-based	Combined
Personnel	$4,296	$ 9,595	$13,891
Student Aide Time	0	13,634 (5,226)	13,634 (5,226)
Parent Time	1,595 (611)	0	1,595 (611)
Facilities	14	156	170
Equipment	120	986	1,106
Materials/Supplies	95	300	395
Evaluation	350	350	350
Transportation[b]	104 (64)	182 (112)	286 (176)
Total Cost	6,574 (5,550)	25,203 (16,725)	31,427 (21,925)
Cost Per Child	$ 657 ($ 555)	$ 2,520 ($ 1,673)	$ 3,143 ($ 2,193)

[a]Numbers in parentheses reflect the use of minimum wage to value time.

[b]Transportation includes cost of driver's time.

hour of time. Transportation costs were based upon information obtained from the parent questionnaire. Home-based intervention was substantially less expensive than center-based intervention no matter how parent and volunteer time was dealt with. Adding in the costs of parents and student volunteers actually increased the difference in costs between the home- and center-based programs. The average time per week parents in the home-based group reported spending in intervention was slightly more than the 1.75 hours "required" by the program, although response varied considerably within the group (x = 2.0 hours, SD = 1.6 hours).

Discussion

The results are relatively easy to interpret from an economic perspective. The parent-delivered, home-based intervention was more economically efficient for the outcomes measured. The center-based intervention was more expensive and less effective. Combining the two types of intervention added only to costs. For this particular sample, the training of parents to conduct language therapy at home was clearly superior. There are a few caveats to bear in mind, however. Only immediate effects were measured, and outcome variables were limited to measures of communication abilities. Language development is a complex process with many determinants, and there may be a variety of intervention strategies that obtain the same results (Horowitz & Sullivan, 1981). Nevertheless, this study contributes to the literature on program design and parent involvement. Substantial improvements in communication abilities were achieved and significant differences were found between the treatments, despite the small sample size, limited duration of the experiment, and small amount of time spent in formal intervention activities by the home-based group. The findings should give impetus to further research on parent-delivered programs as a low cost and effective alternative to center-based programs.

One possible explanation for the superior performance of the parent-delivered intervention can be found in Tizard's (1981) findings that child-adult conversations differ between preschool center and home. Tizard found that conversations were more frequent and more sustained at home. Also, children more frequently and persistently asked questions at home. Those findings suggest greater opportunity for effective language therapy in the home. Researchers should consider observing the frequency and type of language interactions in the home and center in future studies of alternative interventions.

The measurement of time use and estimation of time costs are crucial issues in cost-effectiveness analysis of early intervention programs that involve parents and volunteers. The valuation of their time is particularly problematic for two reasons. First, there is no market transaction establishing the price of the time spent in intervention activities. Second, there is often joint production--students receive an education at the same time that they provide therapy, and parents may engage in more than one activity at a time.

The student aides were not paid for their time because it was part of their coursework. The value of their time was inferred from what other students were paid. Even then, it was impossible to determine how much of their time cost should be allocated to the intervention because their time was simultaneously used in their education. Ultimately, any division of cost between education and intervention was to some extent arbitrary.

The most common time costs to parents were for transportation and participation in intervention activities. In this study, data on time use were obtained from questionnaires. The questionnaires revealed that many parents car-pooled. In the absence of those data, transportation costs might have been estimated based on the assumption that every family drove every day. Estimating the time cost of parent activities, even with a questionnaire, is difficult. Parents may substitute intervention activities for other types of interaction with their child, may engage in other activities simultaneously, or may incorporate intervention strategies in many of the activities that they ordinarily engage in with their child. There are indications from other studies that surveys which ask how much time was spent in one specific activity (as in this study) provide less accurate information than other methods (Robinson, 1985). Thus, researchers measuring the time use of intervenors may wish to consider such methods as short-recall time diaries (Juster & Stafford, 1985), or direct observation (Hoge, 1985). Better data on time use might provide insights into the reasons for the differential efficacy of alternative interventions as well as improve the precision of cost estimates.

Products

Number in Product List	Title
5	Economic Costs and Benefits
50	Language Intervention Efficacy
102	Parent and Clinic Early Intervention

INVESTIGATION 11:

AN ANALYSIS OF SPECIAL EDUCATION EARLY CHILDHOOD PROJECTS APPROVED BY THE JOINT DISSEMINATION REVIEW PANEL

In recent years, the number of early intervention programs for handicapped, at-risk, and disadvantaged children has increased dramatically. There are now legal mandates in 18 states to provide intervention services for handicapped preschoolers, and similar legislation is pending in several other states. Money allocated to state educational agencies for the development of preschool special education programs (through the Preschool Incentive Grant Program) has doubled from $12.5 dollars in FY 1978 to $25 million in FY 1984.

This increased level of activity again focuses attention on questions concerning the efficacy of early intervention for handicapped and disadvantaged preschoolers. Although review articles examining this question abound, their conclusions are equivocal and several previous reviewers (Dunst & Rheingrover, 1981; Simeonsson, Cooper, & Scheiner, 1982) have suggested that methodological weaknesses in much of the early intervention efficacy research has made it difficult to reach definite conclusions.

One source of early intervention efficacy data which many have suggested might not exhibit the same degree of methodological weakness consists of early intervention projects approved for national dissemination by the Joint Dissemination Review Panel (JDRP). Indeed, Odom and Fewell (1983) concluded that JDRP projects "were among the best the field has to offer" (p. 445). According to Tallmadge (1977), the criteria by which the panel judges effectiveness require that the project demonstrate positive impact in terms of statistical significance and educational importance, that measures of impact be reliable and valid, and that evidence be presented that the project can be replicated in other sites. Given these criteria, the systematic evaluation process to which all JDRP-approved projects are subjected, and the widespread perception that JDRP-approved projects are of an exemplary nature (Datta, 1977), it was felt that a careful examination of JDRP projects could provide valuable information to the early intervention field.

Odom and Fewell (1983) recently completed a similar analysis in which they described the service delivery models, target populations, goals, evaluation designs, outcome measures, and costs of projects funded by the Handicapped Children's Early Education Program (HCEEP) which has been approved by JDRP. They concluded that although many of the projects had weaknesses from an experimental design perspective, most of these weaknesses were either unavoidable or understandable. Odom and Fewell went on to suggest ways in which early intervention efficacy research designs could be improved. The present article is both a replication and an extension of Odom and Fewell's work. Whereas their article was primarily descriptive, the present report is primarily a critical analysis which seeks answers to the following questions.

1. What are the characteristics of projects approved by JDRP as being exemplary?
2. What is the quality of the research conducted by JDRP-approved early intervention projects, and are there implications of the research conducted by these projects for future early intervention efficacy research?
3. What can be concluded about the efficacy of early intervention based on the reports of JDRP-approved projects?

Description of the Joint Dissemination Review Panel

The Joint Dissemination Review Panel is an interagency panel established by the federal government in 1975 for the purpose of determining if specific educational programs have sufficient evidence of effectiveness to justify including them in an official government publication entitled Programs That Work (National Diffusion Network, 1983). The 22-member Joint Dissemination Review Panel is comprised of equal numbers of people from the National Institute of Education and the Department of Education. Panel members are chosen for their expertise in education, as well as their ability to evaluate the effectiveness of educational programs. Each application is reviewed by at least seven members of this panel and approved for national dissemination if the panel concludes that the project has demonstrated educationally significant effects based on reliable and valid data which were obtained using well-documented and replicable procedures (Fang, 1981; Tallmadge, 1977). Each project is judged on a written application (limited to 10 pages) which follows a standardized format describing the context in which the project was implemented, the goals and objectives of the project, specific claims of effectiveness made by the project, the unique components and theoretical/conceptual base of the project, a detailed description of the evidence of effectiveness, an explanation of dissemination/replication activities, and the costs of the project (Tallmadge, 1977).

Once the project is approved by the JDRP, the federal government disseminates information about the program through a publication entitled, Programs that Work (National Diffusion Network, 1983). In some cases, federal funding is provided to assist new agencies with the cost of replicating the approved program and/or to support "outreach" efforts by the approved program (Fang, 1981). The perceived benefits of the JDRP process are that local educators are able to obtain information about the consequences of specific interventions and to select those that have a demonstrated capability to address their specific needs and problems (Datta, 1977); and the costs of developing similar projects can be substantially reduced because of the ability to replicate previously validated projects.

Procedures

Collection of Data

Since the establishment of the panel in 1975, over 500 applications for JDRP approval encompassing all types of educational programs and products have been reviewed (Fang, 1981). During this time, 21 early intervention programs for handicapped children have been approved by JDRP--all of which were funded by HCEEP.[1] Copies of the original application materials for each of these projects were obtained from the Technical Assistance Development System (TADS) Project at the University of North Carolina. In addition, letters were written to each of the project directors listed on these applications to determine if there was additional information about the efficacy of the program which could not be included in the JDRP application because of the 10-page space limitation. Ten of 21 project directors responded. In all cases, the additional information provided was descriptive information about the program, and not additional information on the effectiveness of the program.

The application materials for each project were coded along the following dimensions.

1. The characteristics of the subjects included in the research (e.g., child's IQ prior to the intervention, SES, type and severity of handicap).

2. The type of intervention used (e.g., home- or center-based, educational or medical degree to which parents were involved).

3. The type and quality of research design employed (e.g., whether design was pre-post, quasi-experimental, or true experimental; presence of various threats to the internal validity of the design; whether data collectors were "blind").

4. The type of outcome measures and the procedures used (e.g., who collected the outcome information, when was it administered, what type of test was used).

5. The conclusions reached by the study (e.g., the magnitude of the effect, the conclusions of the author).

The magnitude of the effect attributed to each intervention project was determined using a standardized mean difference effect size, defined as $(X_E - X_C) \div SD_C$ (Glass, McGaw, & Smith, 1981). This "effect size" measure is essentially a Z score and has been widely used in recent years to describe the impact of educational programs (Cohen, 1977; Glass, 1976; Horst, Tallmadge, & Wood, 1975; Tallmadge, 1977).

[1]One additional HCEEP-funded project has been approved by JDRP but is not included in this analysis. FEED Project (Bloomington, Indiana) was designed to teach seventh and eighth grade students about the normal developmental process and the consequences of early childhood handicapping conditions. Since the project was not designed to provide early intervention services, it was not included in this analysis.

Results and Discussion

Characteristics and Accomplishments of JDRP-Approved Projects

Table 59 contains information describing the 21 projects considered in this analysis. A brief description of each project is given, along with the setting in which the services were delivered, the age ranges and types of handicap of children enrolled in the project, the major accomplishments of the project, and an estimate of the number of times each project has been officially replicated. Projects are listed in alphabetical order of the states in which they were located, and an indication is given of the data at which each project was approved by the JDRP.

As can be seen, projects cover a wide geographical area and serve children with a mixture of handicapping conditions. The earliest projects were approved in September, 1975, shortly after the creation of JDRP. The most recent project was approved in February, 1983. The projects have resulted in the development of numerous curriculum materials, parent training and teacher training materials, handbooks, and assessment instruments. In many cases, these written materials have been disseminated even more broadly than the official replication sites. The fact that these 21 projects have been replicated in over 2,000 sites across the country suggests that the JDRP approval contributes substantially to the type of special education early childhood intervention programs that are implemented across the United States.

Most of the projects for which a determination could be made utilized professionally certified people as the primary intervenors, and all of them used some type of educational intervention as opposed to a medical, dietary, or sensory stimulation intervention. Findings with respect to other frequently cited variables are summarized below.

Table 59

Descriptions of Early Intervention Projects Approved by the JDRP

Project Name and Location (Date of Approval)	Description	Types of Handicap, Ages	# of Replication Sites	Major Accomplishments
Rutland Center Project Athens, GA (9/75)	A center-based project for preschoolers with severe emotional problems	Emotionally disturbed, ages 2-8	78	Curriculum for emotionally disturbed Developmental therapy textbook
PEECH Project Champaign, IL (11/75)	A center-based program for children and families	Mixed handicaps, ages 3-5	42	Manuals on classroom planning Family involvement manuals
Macomb 0-3 Project Macomb, IL (6/80)	A home-based program for children and their families	Mixed handicaps, ages 0-3	15	Baby buggy series of books/papers Use of mobile van to deliver services Rural network
Peoria 0-3 Project Peoria, IL (2/79)	A home-based program for children and their families	Mixed handicaps ages 0-3	127	Assessment instruments Slide-tapes on normal and abnormal development
ERIN Project Dedham, MA (7/78)	Early recognition and home- & center-based intervention program	Mixed handicaps, ages 207	40	Preschool screening systems Developmental Inventory Developmental checklist
High/Scope Project Ypsilanti, MI (3/79)	A cognitively oriented center-based intervention program	Mixed handicaps, ages 4-6	61	Cognitively oriented curriculum Teacher training manual
UNISTAPS Project St. Paul, MN (9/75)	A family oriented home-based program for deaf/hard of hearing	Deaf and hearing impaired, ages 0-5	--	Home activities guide Preschool planning materials
Central Institute Project St. Louis, MO (11/75)	Parent as teacher home-based project for hearing impaired	Hearing impaired ages 0-4	--	Teacher training materials Training institutes
Regional Demonstration Program, Yorktown Heights New York (6/81)	A center-based interdisciplinary team intervention program	Mixed handicaps, ages 3-5	10	Curriculum guides Parent volunteer manual Manual and activity catalog
Preschoolers/Families Project, Fargo, ND (7/79)	A four-part home- & center-based prevention/intervention program	Develop. delayed Emotionally disturbed, ages 0-6	--	Magic Kingdom Screening Program Parent training materials
Chapel Hill Project Chapel Hill, NC (2/83)	A center-based assessment/intervention program for mildly handicapped	Mixed handicaps, ages 4-6	900	Learning accomplishment profile Mainstreaming materials Training materials
Teaching Research Project Monmouth, OR (3/78)	Home- & center-based individualized skills instruction program for moderately/severely handicapped	Mixed handicaps, ages 1-8	150	Book on data-based classroom Teaching research curriculum Teaching materials
Good Samaritan Hospital Project Portland, Oregon (6/81)	A center-based diagnostic/prescriptive classroom program	Multi-handicapped Physically handicapped ages 0-6	10	Task analyzed curriculum materials Parent training materials
DEBT Project Lubbock TX (10/80)	A home-based/parent taught intervention program	Mixed handicaps, ages 0-2	40	DEBT Developmental Scale DEBT Teaching Activities Training materials
PEECH Project Wichita Falls, TX (7/79)	A home-based/parent taught intervention program	Mixed handicaps, ages 1/2-6	48	Teachers handbook Parents handbook
MAPPS Project Logan, UT (6/80)	A home- and center-based program for children in remote areas	Mixed handicaps, ages 0-5	25	CAMS Curriculum Materials Criteria-Referenced placement test Program evaluation materials

Table 59 (continued)

Descriptions of Early Intervention Projects Approved by the JDRP

Project Name and Location (Date of Approval)	Description	Types of Handicap, Ages	# of Replication Sites	Major Accomplishments
SKI*HI Project Ogden, UT (7/78)	A home-based diagnostic and intervention program	Hearing impaired, ages 0-6	50	Curriculum materials Identification materials Total communication curriculum
Down Syndrome Project Seattle, WA (9/75)	A center-based program for Down syndrome children	Down Syndrome ages 0-6	52	Parent involvement procedures Teacher training materials
Communication Project Seattle, WA (9/75)	A home- & center-based program for communicatively disordered	Communicatively disordered, ages 0-6	40	Teacher training materials Classroom observation systems
C.P. Project Milwaukee, WI (9/75)	A home-based multidisciplinary program for physically disabled	Physically disabled ages 0-3	85	Pre-speech assessment scale Training materials
Portage Project Portage, WI (11/75)	A home teaching program for multi-categorical handicapped	Mixed handicaps ages 0-6	70	Portage Guide to Early Education Portage Parent Program Portage checklist

Parental Involvement

Many people believe that parental involvement is an important contributor to the success of early intervention programs (Bronfenbrenner, 1974; Comptroller General, 1979; Goodson & Hess, 1975; Hewett, 1977; Weikart et al., 1978). That position is consistent with the characteristics of JDRP-approved projects. Eighty-one percent of the projects involve parents to a substantial degree, and 33% of the projects are solely home-based program.

Age at Which Intervention Starts

The projects approved by JDRP are consistent with the common belief that early intervention programs are more effective the earlier they begin (Bronfenbrenner, 1974; Comptroller General, 1979; Garland et al., 1981; Marquis, 1976). Fourteen of the 21 projects (67%) provide services beginning before 1 year of age. Ninety percent provide services beginning at age 3 or before. Two of the JDRP-approved projects--Logan, Utah, and Seattle, Washington--states in their application materials that children who began the program earlier did better.

Types of Children Served

Most projects provides only vague information about the types of handicapped children served. Thirteen of the 21 projects noted only that the services were provided to mixed handicaps. Very little information was given about the demographic characteristics of the populations (e.g., SES, IQ prior to program initiation). The vagueness with which subject populations are described in JDRP application materials is unfortunately characteristic of many early intervention studies (Casto et al., 1983). Because of the extreme heterogeneity of children served in early intervention programs, it is very important to describe specifically the characteristics of participating children. To say that an intervention is effective with "children with mixed handicaps" or "children with Down syndrome" makes it very difficult to judge whether that particular intervention will be effective with other "mixed handicap" or "Down syndrome" children.

Degree of Structure

eral people have noted that early intervention curricula which are more st . red are more effective (Bronfenbrenner, 1974; Karnes et al., 1970; Lazar & Darlington, 1982), while others have claimed that the type of curriculum is unrelated to program effectiveness (Jason, 1975; Weikart et al., 1978). From the 21 projects considered in this analysis, the curriculum was sufficiently described in 15 cases to make a determination of the degree of structure. Of those, 20% were very structured, and 73% were somewhat structured. Thus, programs which are somewhat or very structured are more frequent than unstructured programs.

Type of Outcomes Measured

The most frequently measured outcomes were IQ (45%), language (28%), motor (13%), and preacademic skills (13%). It is unfortunate that outcomes measured for this body of literature are as narrow as outcomes reported in the general early intervention efficacy literature (Casto et al., 1983; Simeonsson et al., 1982). Even though most people agree that it is important to measure the effect of early intervention on variables such as family functioning, children's social/emotional status, or daily living skills, such measures are infrequently reported. Also, JDRP reports for these projects contained no data about long-term effectiveness. Yet, questions about long-term effectiveness are some of the most frequently asked questions about the efficacy of early intervention.

Strengths and Weaknesses of JDRP-Approved Projects

Each of the 21 JDRP-approved projects provided some information about the effectiveness of their project, and all the authors claimed that there was clear evidence that the interventions resulted in important benefits. This is not surprising, since the primary purpose of the Joint Dissemination Review Panel is to determine whether claims for effectiveness made by the project can be supported by data which are reliably and validly collected so that others can be confident that the observed results are really attributable to the treatment. Contrary to our initial expectations, however, the analysis of experimental design characteristics of each project, as shown in Table 60, revealed that all of the projects suffered from major internal validity threats. Indeed, on a 5-point rating scale with 1 being high and 5 being low, all of the projects were related as either 4s or 5s, indicating methodological problems of such a serious nature that one could not be confident that the observed effects were due to the treatment and not to extraneous factors. The problems, as shown in Table 60, was not with the types of outcome used or the size of samples but the types of design employed (15 of the 21 projects utilized only pre-post designs) and the serious threats to the internal validity of all studies.

As mentioned earlier, each of the seven threats identified by Campbell and Stanley (1966) to the internal validity of experiments (maturation, history, testing, instrumentation, regression, selection, and experimental mortality) was coded as to whether it was not a threat, was a minor threat, or was a major threat. The most serious threats to the internal validity of the JDRP-approved studies were

Table 60

Efficacy Data for Early Intervention Projects Approved by the JDRP

Project	Duration[a]	Sample Size	Design	Major Threats to Validity[b]	Study Quality[c]	Outcome Measures	Average Effect Size
Rutland Center Project Athens, GA	5 months	49	Pre-post	M H T I R	5	Developmental Therapy Objectives,[d] Referral Form Checklist[d]	NC
PEECH Project Champaign, IL	7 months	27	Experimental-Control	h t I em	4	Metropolitan Readiness Test Frostig Visual Perception Test[d] Stanford-Binet, Calif. Achievement Test	.34
MACOMB 0-3 Project Macomb, IL	18 months	34	Pre-post	m h I r em	5	Alpern-Boll, Receptive-Expressive Emergent Language Scales	.28
Peoria 0-3 Project Peoria, IL	12 months	77	Pre-post	m h t I r	4	Functional Profile	1.33
ERIN Project Dedham, MA	6 months	25	Pre-post	h t I R em	5	McCarthy Scales Metropolitan Reading Tests,[d] Preschool Screening System[d]	.70
High/Scope Project Ypsilanti, MI	9 months*	16	Pre-post	h I r em	4	McCarthy Scales, High/Scope Preschool Productive Language Assessment Tasks[d] High/Scope Child Observation Record[d]	.66
UNISTAPS Project St. Paul, MN	9 months	25	Pre-post	m h t I r EM	5	Preschool Attainment Record Merrill-Palmer Scale of Mental Tests[d]	.66
Central Institute Project St. Louis, MO	24 months 30 months 36 months 42 months	25 21 29 32	Pre-post	m h T I r	5	Scales of Early Communication Skills[d]	NC
Regional Demonstration Program Yorktown Heights, NY	48 months	102	Pre-post	m t I r em	4	McCarthy Scales	.55
Preschoolers/Families Project Fargo, ND	7 months	39 32	Pre-post	m h t I r em	5	Alpern-Boll, Therapeutic Evaluation & Treatment Center Skills Assessment, Parents and Children Together[d]	.44
Chapel Hill Project	8 months	93, 21	Quasi-Experimental	t I s em	4	McCarthy Scales, Special Education Placement	.52
Chapel Hill, NC	8 months	93	Pre-post	m h t I em	4	McCarthy Scales	
Teaching Research Project Monmouth, OR	—	20	Pre-post	M H T I	5	Teaching Research Curriculum for the Moderately & Severely Handicapped[d]	NC
Good Samaritan Hospital Project	9 months	28	Quasi-Experimental	T I S em	5	Student Progress Record	.57
Portland, OR	9 months	28	Pre-post	m h T I em	5	Student Progress Record	.82
DEBT Project Lubbock, TX	15 months	103	Pre-post	h t I r	5	Koontz Child Development Program Follow-up Placement[d]	.55
PEECH Project Wichita Falls, TX	44 months	98	Pre-post	m h t I r EM	5	Alpern-Boll, Stanford-Binet	1.24
MAPPS Project	18 months	120	Pre-post	h I r	e	Bayley Scales, Peabody Picture Vocabulary Test, Visual-Motor Integration, Assessment of Children's Language Comprehension Test	1.65

(continued)

Table 60 (continued)

Efficacy Data for Early Intervention Projects Approved by the JDRP

Project	Duration[a]	Sample Size	Design	Major Threats to Validity[b]	Study Quality[c]	Outcome Measures	Average Effect Size
SKI*HI Project Ogden, UT	10 months	30, 35	Quasi-Experimental	m h t I r	4	Receptive Expressive Emergent Language Scale	1.02
	10 months	61	Pre-post	m h t I R	5	Receptive Expressive Emergent Language Scale	1.00
Down Syndrome Project Seattle, WA	9 months	66	Pre-post	M h I EM	5	Denver Developmental Screening Test[d] Gesell Preliminary Inventory,[d] Down Syndrome Assessment Form[d]	NC
Communication Project	10 months	39	Pre-post	m h I r EM	5	Sequenced Inventory of Communication Development Peabody Picture Vocabulary Test,[d] Preschool Profile, Language Sample	2.27
C. P. Project Milwaukee, WI	9 months	36	Pre-post	m h T I r em	5	Receptive Expressive Emergent Language Scale,[d] Preschool Attainment Record,[d] Mecham Verbal Language Development Scale,[d] Peabody Picture Vocabulary Test[d]	NC
Portage Project Portage, WI	9 months	57	Pre-post	m h I r	5	Stanford-Binet, Cattell, Peabody Picture Vocabulary Test, Utah Test of Language Development[d]	

[a] Duration of the project for which an effect size was calculated. In some cases, projects lasted longer.

[b] Each study was coded as to whether the following threats to internal validity were not a problem, a minor problem, or a major problem. Capital letters indicate a major problem; lower case letters a minor problem; no letter, no problem: M = maturation; H = history; T = testing; I = instrumentation; R = regression, S = selection; EM = experimental mortality.

[c] Based on the number and severity of threats to the internal validity of the study, an overall rating of study quality was made. Each study was coded as: 1 = excellent, 2 = good, 3 = fair, 4 = weak, 5 = very weak.

[d] Indicates those outcome measures which were used by the study but for which an effect size could not be computed because insufficient data were reported, or the measure duplicated another outcome also reported (e.g., when two IQ tests were administered, only the results of the highest-quality test were used for an effect size).

instrumentation, testing, and experimental mortality. Each threat to internal validity is discussed briefly below, with examples of the kinds of problems present in studies included in this sample.

Maturation

Maturation is a threat to the internal validity of an experiment when biological, physiological, and/or psychological changes that occur simply with the passage of time could have accounted for the gains otherwise attributable to the intervention. Maturation is frequently a threat to interpretation in early intervention programs, since the variables being measured change naturally so rapidly during the first few years of life. Of the 21 JDRP-approved projects, 13 exhibited minor problems and three exhibited major problems with maturation.

For example, several JDRP projects reported positive gains on criterion-referenced assessments of motor development. Since the treatments took place over a period of 9 to 18 months, it is plausible that the children would have progressed on such criterion-referenced measures even if the treatment program were completely ineffective. When such pre- to posttest gains are measured on norm-referenced tests, maturation is not such a serious threat because the norming process on such measures supposedly controls for normal maturation. However, recent work (Powers, Slaughter, & Helmick, 1983; White, Taylor, Carcelli, & Eldred, 1981) suggests the even on the best of tests, norms do not control for all of the maturation effect.

The best way to eliminate maturation as an alternative explanation for an observed effect is to utilize an untreated control group of comparable children. Then, maturation will occur equally in both groups, and any differences between the groups can more confidently be attributed to the treatment. In addition, since unnormed measures are more susceptible to maturation threats, researchers would do well to use assessment measures which provide norms appropriate for the population which is being studied.

History

History is a plausible rival hypothesis whenever children in the treatment group participate in activities which are not a part of the treatment but could have led to the increased scores. Eighteen of the 21 projects exhibited minor history threats, and two exhibited major history threats.

For example, one project reported an 8.1-month mean gain in language on the Bzoch-League Receptive Expressive Emergent Language Scale for children enrolled in their program for one year. However, it is altogether possible that events in the children's lives outside the intervention program contributed tot heir increased performance on the language measure (for example, increased family interaction with the child independent of the intervention program, greater exposure to educational television programs, or increased association with neighborhood children and consequent exposure to appropriate modeling). Since very little information was given about the sample or the treatment setting, it was difficult to exclude history as an alternative explanation. History threats are usually more serious in pre-post designs than in experimental designs. When comparable groups of children from the same general area are assigned to experimental and control groups, it is more realistic to believe that experiences outside of the treatment conditions will be similar for each group.

Testing

Testing is a threat to the internal validity of a study when a child's increased scores are due to practice in taking the test rather than the effect of the treatment. Particularly in situations where the intervention program is developed using behaviors drawn directly from the test items used as the outcome measures, this can be a very serious threat. Repeated practice of specific behaviors included in the test items will lead to increases in scores without necessarily indicating that the child has improved in the domain of behaviors of which the test item is one very limited example. Testing threats can also occur when children become more comfortable with testing procedures and more highly motivated to perform for the examiner. For example, Zigler, Abelson, and Seitz (1973) found gains of as much as 10 points upon repeated testing of Head Start children with no intervention. They attributed these increases to children's increased motivation and comfort with the task. Ten projects had minor problems, and five had major problems with testing as a threat to the internal validity of the study.

An example of testing threats in this sample was seen in a project in which children were administered the scales of Early Communication Skills for Hearing Impaired Children (Moog & Geers, 1975), a criterion-referenced instrument, at least every six months for a period of 2-1/2 to 3-1/2 years. Evidence of program effects were inferred from the differences between the initial pretest and the last posttest. In this situation, changes in scores could easily have been at least partially attributable to the child's increased comfort, motivation, and skill at test-taking rather than to the intervention.

Instrumentation

Changes in tests, testing procedures, or test administrators may result in instrumentation threats. In addition, if researchers involved in the intervention conduct the outcome assessments, they may unconsciously or consciously bias the results in favor of children in the experimental groups. All of the projects had major problems with the instrumentation as a threat to the internal validity of the study.

For example, one of the projects measured child progress via change in teacher ratings of the children's high priority problems. Although the project found that the percent of problems perceived by teachers decreased from entry to termination, it is entirely plausible that the teachers' judgments were biased after having worked with the children for a year. A relatively simple way to protect against this type of instrumentation threat is the utilization of "blind" data collectors or test administrators (people who do not know the purpose of the experiment, which children are in which groups, and who have no invested interest in the outcome). Unfortunately, none of the 21 projects approved by JDRP utilized "blind" examiners. Even so-called objective tests or standardized tests are subject to instrumentation threats due to examiner bias. Small nonverbal cues, often unconscious, can have dramatic effects on test scores.

Regression

Whenever pre- and posttest scores are not perfectly correlated and subjects are selected on the basis of extreme scores, their scores upon remeasurement will be less extreme than the scores on which they were selected. This natural phenomenon has been demonstrated repeatedly and can now be accepted as a statistical truism (Hopkins, 1969). Regression towards the mean, as it is sometimes called, was a minor threat in 13 of the studies and a major threat in three.

In one of the JDRP-approved projects, 12 children with the most severe problems on a criterion instrument were selected from a pool of 23 moderately to severely handicapped preschoolers. Although this approach is admirable in terms of providing direct services to the children in greatest need, it seriously hinders the interpretation of the experimental results. For example, in this particular project, children who were treated showed an average gain of 5.1 points. However, based on the fact that children were selected because they were 26 points below the mean, and assuming a correlation of .70 between the pre- and posttest, one would have predicted a gain of almost 7 points due to regression alone (Hopkins, 1969). Thus, all of the observed "growth" attributed to the intervention could more plausibly be explained in terms of regression towards the mean.

Selection Bias

Many research projects judge the effect of the treatment as the differences in scores between the children who received the treatment and children who did not. Selection bias is a threat to such interpretations whenever the groups are not comparable before the treatment begins. Very few of the JDRP projects suffered from selection bias because there were so few projects which utilized a comparison or control group. Because selection bias is a problem with noncomparability of comparison groups, it is by definition impossible to have selection bias in pre/posttest designs. However, the very best way to control for many of the threats experienced with the pre/post designs in this body of literature (maturation, history, regression, and testing) is to utilize appropriate comparison groups. The believability of the effects claimed by these JDRP-approved projects would have been strengthened substantially had comparison groups been used in place of the pre-posttest designs.

For example, the project mentioned above had 23 children available for the intervention project. Due to limited project resources, only 12 children could be served. Instead of picking the 12 children they considered to be most needy (since all of the children met the eligibility criteria for the project), they could have randomly assigned children to experimental and control groups, resulting in a much stronger experimental design. By not choosing the experimental/control group comparison, the results c the experiment were basically uninterpretable.

Simply having a control group is not the entire answer, however. Control groups or comparison groups must be selected in such a manner that children in both groups are truly comparable before the treatment begins. In another of the JDRP-approved projects, children in an intervention program conducted at a private hospital in a large metropolitan area were compared to a fairly small random sample of all handicapped children served throughout the remainder of the state. Since there was no effort to match children in the two samples, and since families who utilized the private hospital facilities were probably quite different from the average family in the remainder of the state (particularly those in the more rural areas), selection becomes a very serious threat to the interpretation of that experiment.

Experimental Mortality

Whenever subjects are lost from the experiment before outcome data are collected, experimental mortality is a potential threat to the interpretation of the results. If, for example, in a pretest/posttest design all of the children who had very low pretest scores dropped out of the project before it was completed, the average posttest score would be higher than the pretest score even when there is absolutely no treatment effect. None of the JDRP projects reported sufficient data to be sure that experimental mortality was creating spurious results which were being attributed to the treatment; however, several projects had very large attrition rates. For example, one project started the intervention with 203 subjects, but finished with only 92. Although one cannot be sure of the basis of the data reported, it is not implausible that the children who stayed to complete the program were those who were being the most successful. If so, the reported gains are artificially inflated over what they would have been had all children remained to complete the program.

An Analysis of Efficacy Data from JDRP-Approved Projects

The serious methodological problems identified above for all of the JDRP-approved early intervention projects cast doubt on the efficacy data produced by those projects. The analysis of 21 projects resulted in 40 estimates of program impact, since several projects utilized two experimental designs (e.g., both a pre-post design and a quasi-experimental design) or multiple assessment instruments. Unfortunately, there were several projects for which no standardized mean difference effect size could be computed based on the reported information (see Table 64). Considering all of the information about intervention effects, the average effect size was .96, with a standard error of measurement of .10. This means, using IQ as an example, that the average project reported an effect of the treatment of 14.4 points.

Although these effect size results appear very promising in terms of the immediate effects for the participating handicapped children, it is critical to emphasize that the serious methodological problems with all the research designs of these projects create serious doubt as to the confidence with which these effects can be attributed to the treatments.

Conclusions

The field of early intervention for handicapped, disadvantaged, and at-risk children is relatively new. Those projects approved by the Joint Dissemination Review Panel for national dissemination have made substantial contributions to expanding the field, developing service models, and encouraging the provision of services to preschool children. Curricula have been developed, assessment instruments which are useful in diagnosing areas of need and prescribing services have been created, and the feasibility of providing preschool services to handicapped children has been repeatedly demonstrated. The fact that these projects have been replicated in over 2,000 sites demonstrates that such programs are feasible, manageable, and well-liked by teachers, program administrators, and parents. Those are important and legitimate accomplishments.

Unfortunately, the methodological weaknesses of the research conducted by all JDRP-approved early intervention projects are so serious that the resulting efficacy data are of questionable value. The pervasiveness of methodological problems also raises questions about the procedures and standards used by the Joint Dissemination Review Panel. Based on these data, it appears that the panel may be approving projects which do not have convincing evidence of educationally significant effects.

The problems with research methodology, however, do not imply that the projects themselves are of questionable worth. Indeed, the widespread acceptance of the procedures and materials developed by these projects is an indicator of their exemplary nature. In critiquing the quality of research conducted by these projects, it is important to remember that all of the projects were funded as demonstration projects and not research projects. In fact, the HCEEP funding guidelines strongly discouraged the use of HCEEP funds for conducting research. It would be inappropriate to fault these project directors for not conducting high-quality research, given the mandate to demonstrate services and the position of the funding agency. It would also be inappropriate and unfortunate not to learn from these projects on now to improve future efficacy research.

We agree with Odom and Fewell (1983) that the methodological weaknesses of JDRP-approved early intervention studies are understandable but not that they are generally unavoidable. The weaknesses are understandable given the funding agency's emphasis of demonstration and de-emphasis of research during the time period that these projects were conducted. However, many of the methodological problems with the research conducted by these projects could be feasibly avoided with proper planning.

The most serious problem with the research conducted by JDRP-approved early intervention projects is the almost total reliance upon pretest/posttest designs. The difficulties of interpreting data from these projects would have been substantially reduced had control groups or alternative treatment comparison groups used. Interestingly, the simplest approach is also the most defensible. By randomly assigning children to groups at the beginning of the experiment, the need for complex statistical manipulations at the end of the experiment is largely eliminated. Even though many people have suggested that the use of experimental/control groups is impossible in early intervention research, such designs have been successfully used by many researchers (see Andrews et al., 1982; Gordon, 1969; Gray & Klaus, 1970; Ramey & Haskins, 1981; Williams & Scarr, 1971). Particularly in those cases where the number of children in need of services far exceeds the capacity of the service agency to provide services (an almost ubiquitous occurrence, if one is to believe funding rests to state legislators and federal offices), or in those cases where alternative treatment programs are being considered (e.g., half- vs. full-day programs), random assignment to treatment/no treatment groups or to alternative treatment groups is both feasible and advantageous.

Another procedure which would result in substantial improvements in the quality of data collected by early intervention research projects is the use of "blind" examiners. As noted earlier, instrumentation threats were the single most frequent threat to the internal validity of the projects in this sample. The use of "blind" examiners would have solved virtually all of these problems.

Unfortunately, data from these projects provide very little information which can be used in answering questions about the efficacy of early intervention. Programs included in this data set are consistent with current thinking about what factors contribute to effective intervention (involvement of parents, beginning intervention as early as possible, and using structured intervention programs), but do not provide data to corroborate those positions. Fortunately, there are other data suggesting that early intervention programs for handicapped and disadvantaged children do have immediate beneficial effects (Casto et al., 1983; Ramey, Bryant, & Suarez, in press). However, the question of long-term effects, particularly with handicapped children, remains largely unresolved. Furthermore, issues of what types of intervention programs are most effective for which children still need much additional research.

The projects thus far approved by the Joint Dissemination Review Panel have made important contributions to the field of early intervention for handicapped children. They have established models that have been widely disseminated and have created a great deal of awareness and interest in providing effective intervention services. If the field is to continue to progress, that early commitment to program development and demonstration of services must now give way to the type of rigorous research which enables us to determine the long-term impact of early intervention and the most effective ways of providing services.

Products

Number in Product List	Title
14	Efficacy Review
94	Analysis of JDRP Products

INVESTIGATION 12:

THE ECONOMICS OF EARLY INTERVENTION FOR HANDICAPPED CHILDREN: WHAT DO WE REALLY KNOW?

Claims that early intervention is "cost-effective" and that intervention is more cost-effective the earlier it begins are frequently found in the literature available to public policy makers and practitioners (California State Department of Education, 1982; Edmiaston & Mowder, 1985; Haring & Hayden, 1981; McNulty, Smith, & Soper, 1983; Select Committee on Children Youth and Families, 1985; Strain & Smith, 1984; U.S. Department of Education, 1985; Wood, 1981). These claims are good news and have considerable common sense appeal. Yet, to researchers, such claims must surely seem too broad given the kinds of incremental knowledge research generally yields. To separate hyperbole from conclusions supported by research, the investigators reviewed and analyzed the empirical literature.

Procedures

A computer-assisted search of the literature found 16 articles germane to this review. Fifteen were reports of primary research, and one reported secondary research which used data from other sources to estimate the costs and benefits of alternative early intervention strategies. Each study's design, analyses, economic methods, and results were analyzed in order to assess the validity and importance of its conclusions. So that readers may make their own judgments regarding the importance of any given threat to a study's validity, the grounds for each judgment have been explicitly stated.

This review was limited to studies in which the subjects were identifiably handicapped as preschoolers. This led to the inclusion of two evaluations of the Perry Preschool Project (Berrueta-Clement et al., 1984). Although the subjects in the Perry Preschool Project were economically disadvantaged, most were classified as mentally retarded at the time. One other study (Weiss, 1981) was included because the subjects were described as handicapped in the report, although test scores suggested that most of the subjects would not have been identified as handicapped before school entry. Economic evaluations of early intervention for disadvantaged children have been reviewed elsewhere (Barnett & Escobar, 1987).

Economic Methods

A complete introduction to economic evaluation is beyond the scope of this paper. However, a brief explanation of terminology and key concepts may be of use to those unfamiliar with economic methods. A more extensive discussion of issues in economic evaluation of early intervention has been provided by Barnett (1986b). For further information, readers are referred to the fine texts by Gramlich (1981), Levin (1983), and Thompson (1980).

The purpose of economic evaluation is to assess <u>economic efficiency</u>. An early intervention program is economically efficient if it produces a greater gain than loss for society as a whole. The most efficient intervention is the one that produces the greatest gain. Efficiency is always assessed comparatively. Economic evaluation may also be used to examine the <u>equity</u>, or fairness, of the distribution of gains and losses. Economics provides no criterion for equity as it does for efficiency, however.

Economic evaluations of early intervention have been of two general types: cost-benefit analysis and cost-effectiveness analysis. These two differ in their treatment of outcomes and are typically used to make different types of comparisons. Cost-benefit analysis requires monetary valuation of a program's effects and is usually employed to compare a program to no program. Cost-effectiveness analysis requires the researcher to estimate a program's costs and effects, but monetary valuation of effects is not attempted. Cost-effectiveness analysis is typically used to compare two or more programs seeking to produce the same effect. Cost-effectiveness analysis is relatively less useful for program-no program comparisons or when programs have multiple outcomes which are difficult to compare.

An essential concept in cost estimation is opportunity cost, which is defined as the value of a resource in its best alternative use. Opportunity cost frequently differs from the cost of a resource in a program's budget. For example, buildings may be donated or parent time volunteered at no cost to an early intervention program. These resources may have a value in other uses, however. The building might be used commercially or for a charitable activity. Parents might devote their time to a job, productive activities at home, or leisure. On the other hand, programs may employ persons who would otherwise be unemployed; their opportunity cost is likely to be lower than the wages they are paid.

Another essential concept in economic evaluation is discounted present value. In a multiyear study, dollars from one year are not equivalent to dollars from another year, and two adjustments must be made to produce equivalent dollars. First, inflation causes changes in the value of dollars over time. Thus, a price index, such as the Gross National Product Deflator is used to convert nominal dollars to a standard purchasing power, or "real" dollars. Second, dollars have an opportunity cost over time. A dollar received today is worth more than a dollar received next year, because today's dollar can be invested to yield more than a dollar next year. Real dollars from different years are converted to a common present value using a discount rate. Roughly speaking, the discount rate represents the rate of interest that people must be paid to trade a dollar today for a dollar one year later. Both deflating and discounting reduce the estimates of later costs or benefits more than earlier costs or benefits. These adjustments are especially important in economic evaluation of early intervention as benefits may follow costs by many years and the timing of intervention is an issue.

The last concept to be introduced is net present value, which equals discounted benefits minus discounted costs. Usually, net present value is the appropriate criterion for judging economic efficiency. The program with the greatest net present value is the most efficient, and programs with negative net present values are not sound economic investments. This criterion holds for cost-effectiveness analysis as well as cost-benefit analysis. Contrary to popular belief, the program with the highest cost-effectiveness ratio is not necessarily the most efficient. Consider two programs. Program A increases language test scores by 10 points at a cost of $2,000. Program B increases scores by 15 points for $5,000. Program A has a cost-effectiveness ratio of 10/2,000, Program B 15/5,000. Nevertheless, Program B may yield a greater net present value. The correct way to choose between programs A and B is to decide if five extra points are worth $3,000. In sum, consumers of cost-effectiveness analysis must value effects and estimate net present value in order to identify the most efficient alternative, just as the economist would do in cost-benefit analysis.

Results

The design characteristics and methods of economic analysis used in each primary research study are summarized in Tables 61 and 62, respectively. Table 61 describes sample size, ages of children in the study, research design, and the types of statistical analyses applied to the data. Table 62 specifies whether the study was a cost-benefit analysis or a cost-effectiveness analysis, the alternatives compared, and the completeness of the cost estimates.

As can be seen, many of the studies used relatively weak designs and produced incomplete estimates of intervention costs. Three studies (Barnett, 1985a; Weber, Foster, & Weikart, 1978; Weiss, 1981), were based on experimental designs. Three studies (Rule et al., 1987; Skeels, 196?; Taylor, White, & Pezzino, 1984) employed nonequivalent group designs in which the comparison group was formed by obtaining matches for the experimental subjects. Seven studies (Liberman, Barnes, Ho, Cuellar, & Little, 1979; Stile & Thompson, 1982; Stock et al., 1976; Weiss &

Table 61

Design Characteristics

Study	Sample Size n (exp,con)	Ages	Research Design	Statistical Analyses
Barnett (1985a, b)	121 (63,58)	3-4	Experimental	Multivariate with prest & demographic cov.
Weber et al. (1978)	123 (65,58)	3-5	Experimental	ANCOVA and others
Weiss (1981)	114 (59,55)	3-5	Experimental	ANCOVA w/ pretest as cov.
Skeels (1966)	25 (13,12)	0-3	Nonequiv. group	t-test
Rule et al. (1985) (1985)	24 (12,12) 18 (9,9)	2-5	Nonequiv. group	ANCOVA on matched pairs with pretest as cov.
Taylor et al. (1984)	50 (25,25) 44 (22,22)	2-4	Nonequiv. group	ANOVA on matched pairs
Liberman et al. (1979)	97	0-21	1-group pre-post	Wilcoxin T & t-test
Stile & Thompson (1982)	8 10	0-3 3-5	1-group pre-post	None
Stock et al. (1976)	130 (28 progs.)	0-6	1-group pre-post	t-test, actual posttest v. posttest est. from pretest
Weiss & Jurs (1979)	72	2-5	1-group pre-post	t-test, actual posttest v. posttest est. from pretest
Casto & Tolfa (1981)	60 60	0-3 3-5	1-group pre-post	Correlated t-test
Walker (1981)	15	0-5	1-group pre-post	None
Hutinger (1981)	unknown	0-3	1-group pre-post	unknown (p values reported)
Macy & Carter (1980)	819 (15 progs.)	>3	1-group posttest only	none
Snider et al. (1981)	10	unk.	1-group posttest only	none

Table 62

Economic Methodology

Study	Type*	Comparison	Completeness of Cost Analyses
Barnett (1985a,b)	CBA	Program v. no Program	All program costs, Parents' time
Weber et al (1978)	CBA	Program v. no Program	All program costs, Parents' time
Weiss et al. (1981)	CBA	Program v. no Program	Unknown
Skeels (1966) (1966)	CBA	Intervention v. deprivation	Program costs except for capital
Rule et al. (1985) (1985)	CBA	Mainstream daycare v. self cont./Head Start	Program costs except for capital
Taylor et al. (1984)	CEA	Half-day v. Full-day	All program costs, Parents' costs Volunteers & donations
Liberman et al (1979)	CBA	Program v. no Program	Budgeted program costs
Stile & Thompson (1982)	CBA	Program v. no Program	Unknown
Stock et al. (1976)	CEA	Various program characteristics	Costs of direct service staff
Weiss & Jurs (1979)	CEA	Program v. no program	All program costs
Casto & Tolfa (1981)	CEA	Program v. no program	Program costs except capital and administ.
Walker (1981)	CEA	Program v. no program	Program costs except capital and administ.
Hutinger (1981)	CEA	Program v. no program	Program costs except capital and administ.
Macy & Carter (1980)	CEA	Program v. no program	Staff salaries
Snider et al. (1974)	CBA	Program v. no program	Unknown

*CBA = Cost-benefit analysis
CEA = Cost-effectiveness analysis

Jurs, 1984; Casto & Tolfa, 1981; Walker, 1981; Hutinger, 1981) used single group, pre-post designs, and two studies (Macy & Carter, 1980; Snider, Sullivan, & Manning, 1974) used single group, posttest-only designs.

Experimental Designs

The studies from the Perry Preschool Project (Barnett, 1985a; Weber et al., 1978) provided a relatively strong basis for valid inferences. The subjects were 123 Black children selected on the basis of low IQ and low parental socioeconomic status. Subjects were randomly assigned to an experimental or control group in five waves. The first wave of the experimental group received one year of intervention at age 4. The other waves received two years beginning at age 3. All subjects entered kindergarten at age 5. Minor exceptions to random and independent assignment to groups occurred and were examined by Weikart, Bond, and McNeil (1978, pp. 16-18). The exceptions to random assignment had no statistically significant effects on the findings.

The intervention was center-based, 2-1/2 hours per day, five days per week from October through May. The teachers also visited each home weekly for 90 minutes. The teacher-child ratio was approximately 1:5, with two teachers in a classroom. The program was implemented by a public school special education division with teachers who were trained in both special education and early childhood. The curriculum was cognitive-developmental; a detailed description is available in Weikart, Kamii, and Radin (1967).

The economic analysis attempted to identify all costs and benefits, direct and indirect. Monetary estimates were made for all of the costs identified, but for only some of the benefits. Nevertheless, the estimation of economic benefits was the most complete in the early intervention literature. Cost and benefit estimates were adjusted for inflation and discounted to calculate net present value as a measure of the intervention's economic efficiency. In addition, the distribution of costs and benefits between taxpayers and program participants was estimated and sensitivity analysis was conducted to examine the effects of variations in key assumptions on the study's conclusions.

The major finding of Barnett's study was that the intervention was economically efficient under a wide range of economic assumptions. Both taxpayers and participants gained more than they lost. Reductions in education costs because of a reduced need for special education were an important source of economic benefit to taxpayers. The largest monetary benefit to the participants was from increased earnings. Earnings were projected over a lifetime based on census data and the participants' educational attainment. These projections were also supported by observed higher earnings at age 19. At a 5% discount rate, the net present value of two years of preschool education was estimated to exceed $10,000 (in 1986 dollars). To illustrate the magnitude of the effects underlying that estimate: Placement in classrooms for mentally retarded students was reduced from 35% to 15%; high school graduation was increased from 49% to 67%; and welfare assistance was reduced from 32% to 18%.

The Perry Project sample consisted of disadvantaged Black children, most of whom would not meet current standards for mental retardation based on their original IQ scores. However, the Perry Project used the 1960 norms for the Stanford-Binet. Evidence from a similar study suggests that IQ scores would be about 10 points lower using the 1972 norms, which would place most of the sample within the range for mild mental retardation with a mean IQ of 69 (Ramey, Bryant, & Suarez, 1985). In view of the issues surrounding the IQ scores of disadvantaged children, generalization of the findings to all mentally retarded children may be problematic (Barnett, 1986a; Mercer, 1973; Zigler, Balla, & Hodapp, 1984), but the conclusions may apply to much of that population.

The other economic evaluation of the Perry Preschool program (Weber et al., 1978) was based on earlier more limited outcome data, but came to much the same conclusions regarding economic efficiency. One salient difference between the studies exists. Weber et al. concluded that one year of intervention was more economically efficient than two years; Barnett concluded that, although differences were not statistically significant, the small number of subjects who had received one year (n=13) provided little statistical power for comparing one and two years. For example, the estimated effect on IQ at kindergarten entry was 50% (four points) higher for those who had two years of intervention. This difference was not statistically significant, but it would be risky to assume that one year was as effective as two based on these data.

Weiss (1981) studied the effects of preschool special education on 3- to 5-year-old "language handicapped and bilingual" children attending public preschool, day care, Head Start, and kindergarten programs (Weiss, 1981, p. 40). The research design matched seven pairs of classrooms on age, sex, bilingual status, and socioeconomic background of children, and on class meeting time. One classroom from each pair was randomly assigned to the INREAL (IN-class REactive Language) treatment group. INREAL specialists provided language therapy in each classroom for a school year. They used an inatist approach (Lindfors, 1980) in a naturalistic way that sought to avoid stigmatizing the children with language difficulties. Specialists served one preschool program in the morning and one or two kindergartens in the afternoon. Additional information regarding the INREAL model is available in Weiss (1981).

One-third of the children (n = 172) from each classroom were randomly selected for pre- and posttesting. Language abilities were assessed immediately before and after the school year using the FIPS (Fluarty, 1974; Weiss, 1981), an individually administered language screening test. The experimental group significantly outperformed the control group on the posttest, with statistical adjustment for pretest differences. In addition, a three-year follow-up was conducted to assess subsequent special education placement and retention in grade. Sample attrition over the three years was a substantial 34%, but did not differ between the experimental and control groups. The experimental group experienced less grade retention and fewer and less restrictive special education placements. Over the three years, the control group experienced 44% more speech and language placements, 33% more resource room placements, and 13% more grade retentions.

To examine the economic efficiency of the program, Weiss estimated INREAL's cost and the cost savings from improved educational placements in the three-year follow-up period. The estimated cost savings exceeded INREAL's estimated cost by over $1,100 per child. Unfortunately, it appears that there are problems with the economic evaluation. The $175 cost estimate seems likely to cover only the salary of direct service staff; actual cost may have been 30% to 50% higher. In addition, cost savings occurred over several years, but were not adjusted for inflation or discounted. Those adjustments reduce benefit estimates by about 20%. In the case of INREAL, more appropriate estimates of real costs and benefits sustain the conclusion that "ordinary" preschool programs can be made more efficient through the addition of a language intervention program. However, the same methodological problems might produce misleading results in another study.

Interestingly, the posttest scores of the control group appear to contradict the statement that the subjects were handicapped as preschoolers. The control group mean score was above the age-adjusted pass/fail cut-off for each subscale and the full test (Weiss, 1981, p. 43). The subjects might more appropriately be called "at risk." The degree of risk is indicated by the control group's special services in the third follow-up year: 25% were in speech-language therapy and 24% in resource rooms. Weiss and Heublein (1983) analyzed data on subjects who were identifiably handicapped (n = 36) separately. By Year 3, 94% of the handicapped control group received special services, compared to 62% of the handicapped experimental group. Comparison of INREAL's effects on handicapped and other subjects suggests no clearly greater overall impact on either group (Weiss & Heublein, 1983, p. 22).

Nonequivalent Group Designs

Perhaps the first study to systematically examine the economic implications of early intervention was Skeels' (1966) 30-year follow-up study. Skeels' study compared the adult outcomes of an experimental early intervention group to those of a comparison group. The subjects were White orphanage children under age 3 when the experiment began. The experimental group had low IQ's initially (mean = 64). As an experiment they were transferred from the orphanage to a residential facility for retarded persons where they were "house guests" and given a center-based intervention program. The comparison group was composed of children who had remained in the orphanage during the experimental period. These children had higher IQs initially (mean = 87).

Initial comparisons showed large IQ gains for the experimental group and large declines for the comparison group. At adult follow-up, the experimental group was clearly more successful than the comparison group in terms of independent living, educational attainment, social status, income, and family formation. For example, median educational attainment was 12 years for the experimental group, 2.75 years for the comparison group. Skeels' economic evaluation of the differences was rudimentary, but it provided meaningful indications of the magnitude of the economic differences between the two groups.

Skeels' study has been extensively criticized (Clarke & Clarke, 1976; Longstreth, 1981) on methodological grounds. Regression to the mean and preexisting differences between the groups posed threats to validity. Yet, the magnitude of the differences in outcomes between the two groups appears too great to be entirely explained by known problems in the study. The findings suggest that early intervention that results in long-term changes in the environment can alter the course of a life that begins in severe deprivation and that this has significant economic benefits. The environmental deprivation of the subjects and the nature of the intervention limit generalization of the findings, however.

Rule et al. (1987) compared 3- and 4-year-old mentally retarded children mainstreamed into day care centers to matched children in self-contained programs and in Head Start programs. The day care intervention used special education professionals to provide direct services to children and to train day care teachers in special education techniques. Special educators, teachers, and aides provided instruction in basic skills in groups, individual microsessions, and "coincidental teaching." The program also addressed the development of social skills and included a home support component. A complete description of the program is given by Rule et al. No statistically significant differences in cognitive or social gains over the school year were found between the day care group and the other two groups. Mainstreamed day care was substantially less expensive, and provided more hours of service, than the self-contained program. Head Start's cost and hours of service were not compared. Rule et al. concluded that the day care model was a viable alternative and that children's gains were comparable to those from other models.

Although the viability of the day care model and its potential to reduce costs were demonstrated, the study's design had two noticeable weaknesses. First, the small sample size limited the study's power to detect differences between the groups. There were only 12 matched pairs in the comparison to self-contained programs and 9 matched pairs in the Head Start comparison. Second, preexisting differences between the groups may have masked differences in intervention efficacy. Although children were roughly matched on age, mental age, and AAMD classification, it appears that the effects of other potentially important variables such as gender, secondary handicapping conditions, previous intervention experience, and family background were not assessed.

Taylor et al. (1984) compared the cost-effectiveness of half-day and full-day programs using matched pairs of handicapped preschool-age children. The 15 classrooms studied were operated by nine public school districts in Iowa. Taylor et al. found no statistically significant differences in posttest measures of cognitive development (22 pairs) or expressive language (25 pairs) between the two groups. The estimated public cost of half-day programs was very high compared to the estimated cost of full-day programs. In fact, it was greater than the cost of full-day programs for children with communication disorders. Taylor et al. concluded that full-day programs were more cost-effective for children with communication disorders while half-day programs were more cost-effective for children with mental retardation.

The study's conclusions can be questioned on at least two grounds. First, Taylor et al. drew conclusions based on differences that were statistically insignificant. Some of these differences were too small to be meaningful even if they had been statistically significant. For example, it was concluded that full-day programs were "most cost-effective" for children with communication disorders because half-day programs cost 2.5% more. Second, the difference in length of day was confounded with other program differences. As Taylor et al. noted, the half-day programs provided more therapy than the full-day programs, more than twice as much in the case of children with communication disorders. Thus, the study's implications regarding the economic efficiency of length of day generally are quite limited, although it may be a useful analysis of these programs in their local context.

Taylor et al. conducted a detailed and exceptionally complete study of costs. Equipment, transportation, and facilities costs were included. They also estimated the costs to parents of time and other resources donated to the programs. The analysis of opportunity cost is a sometimes confusing aspect of economic evaluation, and it may be instructive to review two minor problems that arose in this study. The cost of parent time was estimated by the cost of replacing parents with "paraprofessionals." However, the relevant opportunity cost was the value of the activities parents gave up. A more appropriate estimate of hourly cost would have been parents' hourly earnings, not paraprofessionals' earnings. Taylor et al. also estimated costs for space used in home intervention and for lunches prepared by parents. No such costs would have been estimated if the concept of opportunity cost had been correctly applied. In all likelihood, parents did not purchase larger homes to accommodate home visits and would have provided lunch to their children in any case.

Single-Group, Pre-Post Designs

The seven single-group pre-post studies were problematic in that they measured outcome variables that would have increased over time in the absence of intervention, child development, and, in one study, family income. Moreover, because initial low performance (or income) is generally a criterion for intervention, regression to the mean may also have been a problem. In none of these studies were the increases in outcome variables sufficiently high that maturation and regression to the mean were implausible explanations.

Two of the studies (Stock et al., 1976; Weiss & Jurs, 1984) used pretest scores to estimate posttest scores in the absence of treatment. Although that method addressed the maturation problem, it did not avoid problems due to regression to the mean. Moreover, it is susceptible to problems from initial differences between children who differ in age. For example, children who e er intervention later seem likely to differ in type and severity of handicapping condition from those who enter earlier. On the other hand, in any given year, the pretest scores of older children with previous intervention experience will reflect the effects of intervention. The studies presented no evidence that these problems had been addressed.

In terms of economic analysis, the single-group studies gave few details regarding cost estimation and tended to underestimate costs by including only the personnel costs of direct services and omitting the costs of supervision and administration, transportation, facilities, and equipment. The study by Weiss and Jurs (1984) was an exception in that its cost estimation procedures were highly inclusive and clearly explained. None of the stud s considered potential costs to parents, however. Four of the six studies making program-no program comparisons conducted cost-effectiveness analyses. Their findings are difficult to interpret because readers must estimate the value of the outcome variable for themselves in order to judge program efficiency.

Perhaps the most interesting single-group study was the Battelle (Stock et al., 1976) investigation of the relationships between program characteristics and cost-effectiveness. A random sample of 130 subjects between the ages of 2 and 7 was selected from 32 intervention programs sponsored by the Handicapped Children's Early Education Program (HCEEP). Cost was estimated for each intervention, and effectiveness was estimated in five developmental domains. Cost included salary costs alone and, thus, was underestimated. Effectiveness was estimated by comparing actual posttest with posttest predicted from pretest data. It was found that gains were 2.3 times the predicted gain for personal-social skills; 1.6 times the predicted gains for adaptive, cognitive, and communications skills; and 1.3 times the predicted gain for motor skills.

The Battelle investigators concluded that medium-cost interventions were more effective than high- or low-cost interventions and that interventions with medium child-staff ratios were more effective than those with high or low ratios. Those conclusions are much the same since the primary determinant of variation in cost was child-staff ratio. The validity of the conclusions is undermined by the well-known relationship between cost (child-staff ratio) and type/severity of handicapping condition (Kakalik et al., 1981). The statistical analyses did not attempt to adjust for variation in handicapping condition, and the information reported was not sufficient to determine its importance.

Stock et al. also concluded that home-based interventions, interventions with substantial parent involvement, and interventions using more structured curricula were more effective. With respect to service delivery mode and parent involvement, the finding is potentially explainable by other uncontrolled intervention or child characteristics. A high correlation between service mode and age of child would be expected, for example. No analyses supporting the conclusion regarding structure were reported in the Battelle study, although programs that had developed their own curriculum were associated with greater child gains.

The two single-group, posttest-only studies (Macy & Carter, 1980; Snider et al., 1974) were the least interpretable. Macy and Carter studied 15 early intervention programs serving children ages birth to 2 and reported cost estimates and later special education placements. Snider et al. studied the Regional Intervention Program (RIP) and concluded that RIP had a highly favorable benefit-cost ratio of 7.79 based on reduced costs of later special education placements. However, Macy and Carter followed children no further than kindergarten and Snider et al. relied entirely on projections of future residential placements based on clinical judgments for 10 selected subjects. Neither study produced statistical estimates of program outcomes.

Economic analyses were incomplete in the posttest-only studies. Macy and Carter included only the salaries of direct service personnel. It appears from the brief report that Snider et al. did not discount benefits and (incorrectly) inflated benefit estimates. The seemingly precise benefit-cost ratio produced by Snider et al. gave no hint of the small sample size, difficulties in measuring costs to parents, or the range of conclusions that might be produced by reasonable variations in assumptions such as the discount rate.

Secondary Data Analysis

One secondary data analysis was found, Wood's (1981) benefit-cost analysis of age at start. The study's goals were to determine if early intervention was economically efficient and how efficiency varied with age at start. Estimates from several primary sources were combined to estimate educational costs through grade 12 for handicapped children entering special education at three different ages: before age 2, age 2, and age 6. Based on estimated cost-savings, Wood concluded that early intervention is economically efficient and that intervention before age 2 is ›re efficient. Wood's conclusions have been cited widely (California State Department of Education, 1982; Edmiaston & Mowder, 1985; Haring & Hayden, 1981; McNulty et al., 1983; Select Committee on Children, Youth and Families, 1985; Strain & Smith, 1984). Unfortunately, there were serious problems with the underlying estimates and methods of economic analysis.

One problem with Wood's analysis was the use of incomplete cost estimates. The costs of early intervention were understated relative to the costs of later education because the cost estimates for infant and preschool intervention were less comprehensive than the estimates for school-age costs. For example, the cost of 12 months of preschool was estimated to be $2,310 while the cost of school-age special education was estimated to be $4,445 in 1978-79. Yet a comprehensive national study of special education costs (Kakalik et al., 1981) for the previous year found that the costs of preschool special education exceeded the cost of elementary school special education and almost equaled the average cost of all special education. In addition, the costs of infant intervention appeared to have been more seriously underestimated (as low as $32 for 12 months) than the costs of preschool intervention. Finally, comparison to national estimates (Kakalik et al., 1981) indicated that Wood overestimated the cost difference between regular and special education.

The assumptions made to estimate the effects of intervention on special education placement were also problematic. The estimated effects were from single-group studies which assumed that all subjects would have entered special education without intervention. This tended to overestimate the effects of early intervention. Furthermore, in the absence of empirical estimates for special education beyond kindergarten, it was assumed that of children in special education at age 6, 1/3 would receive 6 years of special education, and 2/3 would receive 12 years of special education. That assumption does not correspond to observed patterns of special education placement by age (U.S. Department of Education, 1986).

Finally, the present value of educational costs was not calculated. Costs over the years were simply added up and the totals compared. This exaggerated the benefits and biased the comparison in favor of an earlier starting age because the present value of costs would be lower for later interventions. When this is added to the other problems, it appears that the findings were substantially biased in favor of earlier intervention. Despite the biases, the estimated difference between intervention before and after age 2 was only $327 over a child's entire school career. Clearly, Wood's study is an unsound basis for the conclusion that earlier is better.

Discussion

The fragmentary and sometimes problematic evidence available regarding the economics of early intervention for handicapped children, contrasts sharply with the confident conclusions offered to policy makers and practitioners (e.g., Edmiaston & Mowder, 1985; McNulty et al., 1983; Strain & Smith, 1984; U.S. Department of Education, 1985). The lack of evidence regarding age at start is particularly striking. This contrast can be attributed to broad generalization of a few strong studies and reliance on studies with significant methodological problems. A lack of familiarity with economic methods may have contributed to the acceptance of highly questionable conclusions. Important elements of cost were frequently omitted from estimates, encouraging policy makers and others to believe that effective early intervention programs could be provided at unrealistically low costs.

Undoubtedly, many of the studies reviewed were attempts to make the most of available data, and their design problems were unavoidable. Nevertheless, the existing research is a weak basis for deciding on economic grounds who should receive publicly funded early intervention, and what kinds of intervention should be provided at what ages for how long. It does, however, provide a starting point for investigating the economic efficiency of early intervention for children with various handicapping conditions. The Perry Preschool Project offers strong evidence that early intervention can be a sound investment for some children with mental retardation. Similarly strong studies need to be conducted for children with other types and severities of handicapping conditions. Researchers should also consider new studies of age at start because the economic implications of this question are great. Even information about the short-term benefits of different starting ages would substantially improve the field's knowledge. As Weiss (1981) has demonstrated, valuable economic evaluations can be conducted without the massive efforts required by studies like the Perry Project. Given current funding levels, those large-scale efforts are likely to be rare.

If the empirical basis for public policy regarding early intervention for handicapped children is to be strengthened, researchers must place greater emphasis on sound methods such as the more interpretable quasi-experimental (Cook & Campbell, 1979) or experimental designs. Although it can be argued that internal validity has been overemphasized generally (Cronbach, 1981), much of the research on the economics of early intervention has had such weak internal validity that it is quite unconvincing. Other desirable characteristics for future research would be larger sample size, collection of longitudinal data, and the use of recently-developed statistical procedures that may increase the interpretability of quasi-experimental designs (Kenny, Lee, Maddalla, & Trost, 1979; Magidson & Sorbom, 1982). Large samples are difficult to obtain because of the low incidence of handicapping conditions among preschoolers, but lack of statistical power and limited ability to covary program and child characteristics have posed serious problems for previous studies.

Finally, researchers should adopt accepted procedures for economic analysis. Failure to do this has produced misleading and inconsistent results. Cost estimates should be at least approximately complete and include all personnel and capital such as facilities. Costs and benefits should be discounted when costs are incurred at different times or when longitudinal data on benefits or effects are involved. These should be minimum requirements for economic evaluations.

Research on the economics of early intervention is a new endeavor in a field that is still quite young and rapidly developing (Bailey & Bricker, 1984). It shares with many other types of research on early intervention substantial theoretical and practical problems (Meisels, 1985; Sheehan & Keogh, 1982). The economic dimension adds its own difficulties. It is hoped that the promise of economic evaluation is sufficiently great that early intervention researchers will apply this approach with greater frequency and enthusiastically tackle the problems that have limited past efforts.

Products

Number in Product List	Title
5	Economic Costs and Benefits
34	Economics of Early Intervention
35	Economic Analysis
51	Methodological Issues

INVESTIGATION 13:

INTRAVENTRICULAR HEMORRHAGE FOLLOW-UP STUDY

Introduction

During the past 20 years, the problem of predicting developmental outcome from characteristics apparent in infancy has taken on new interest among the health professions. The possibility of early prediction is particularly relevant to the population of low-birth-weight (LBW) infants. These children have been shown to be at risk for a range of developmental disabilities and handicapping conditions (Caputo & Mandell, 1970; Murphy, Nichter, & Liden, 1982; Ross, Schechner, Frayer, & Auld, 1982). Major advances in medical care and the development of neonatal intensive care facilities have greatly enhanced the treatment and follow-up of LBW infants. As a result, the mortality rate has declined, but subsequently, it is also becoming progressively more obvious that the same conditions that once caused death in LBW infants are also responsible for a host of adverse neurological, medical, and behavioral sequelae in the survivors (Stewart, Reynolds, & Lipscomb, 1981).

In a recent world survey and review of follow-up studies involving LBW infants, Stewart et al. concluded that intraventricular hemorrhage (IVH) is by far one of the most common causes of developmental disabilities in LBW infants. In modern neonatal intensive care facilities, the incidence of IVH in LBW infants (less than 1500 grams or born after less than 35 weeks gestation) is approximately 40% to 45% (Volpe, 1981).

An understanding of neuropathology of IVH has been made possible largely through the use of ultrasound, computerized tomography (CT), and more recently Positron Emission Tomography (PET) (Grants, Borts, & Schellinger, 1981; Papile, Munsick-Bruno, & Schaefer, 1983; Volpe, Herscovitch, Perlman, & Raichle, 1983). Briefly, periventricular hemorrhage (surrounding the ventricles, but not filling the ventricles) and IVH emanate from the small vessels, or capillaries in the subependymal germinal matrix, located near the head of the caudate nucleus (Hambleton

& Wigglesworth, 1976). Approximately 80% of the cases of periventricular hemorrhage involve a rupturing through the ependyma and fill the ventricular system. In more severe forr , the hemorrhage extends into the cerebral parenchyma which is often followed by the development of a porencephalic cyst, and other complications (Pasternak, Mantovani, & Volpe, 1980). The pathogenesis of IVH is related to several factors concerned with the distribution and regulation of cerebral blood flow, intravascular pressure, vascular integrity, and extravascular environment. These factors combine in the premature LBW infant, partic larly in the infant subjected to asphyxial insult, to result in periventricular or intraventricular hemorrhage (Volpe, 1987).

The prognosis of IVH is best considered in terms of the short- and long-term outlooks. Volpe (1987) reports that all the important factors in determining outcome are not clearly understood, but there is a distinct relation between the severity of the hemorrhage and the prognosis. It is known that more severe hemorrhaging generally produces more short-term medical complications and more long-term developmental complications, but the types of long-term disabilities have not yet been clearly established.

The recent contributions by Papile et al. (1983), through the use of CT, have made it possible to classify IVH into four grades, which allow for follow-up based on severity. The classifications are: Grade I, germinal matrix hemorrhage; Grade II, IVH with normal ventricular size; Grade III, IVH with ventricular dilation; Grade IV, IVH with parenchymal hemorrhage. Grades I and II are considered to be less severe, while Grades III and IV are more severe because of the ventricular dilation and parenchymal involvement.

It has only been in the last four to six years that longitudinal and follow-up studies have begun to look specifically at contributions of IVH to developmental deficits and handicapping conditions. This is partly due to the fact that ultrasound and CT as methods of assessment, detection, and diagnosis of IVH are rather new methods. As a result, recent studies that specifically investigate the after-effects of IVH (Gaiter, 1982; Ment, Scott, & Rothman, 1978; Papile et al., 1983; Papile, Munsick, Weaver, & Pecha, 1979; Williamson et al., 1982) have typically been with children 36 months or younger.

Almost unanimously, there has been a call for more long-term follow-up at late preschool and early school age to determine the residual effects of IVH. The majority of investigations cited earlier have generally agreed that Grades I and II may or may not result in a significant handicap. While it is known that up to 80% of those children experiencing Grades III and IV IVH may exhibit moderate to severe handicapping conditions by the time they are 3 years old, some children in this group appear to be minimally affected (Papile et al., 1983). It is not clearly known which cognitive, physiological, and behavioral functions are the most effected by differences in severity of IVH and which functions have been possibly reacquired due to the plasticity and equipotentiality of the developing brain (Rourke, Bakker, Fisk, & Strang, 1983). There has also been some speculation in the literature that IVH children exhibit learning and behavior problems later in life.

There emerge three basic questions that need further investigation: (1) what are the long-term outcomes for IVH infants; (2) are there longitudinal differences in outcomes that occur between mild and severe IVH; and (3) if there are differences, where do these differences occur neuropsychologically? By determining the differentiating effects of IVH, early intervention and rehabilitation programs might be in a better position not only to identify those children at risk, but more importantly, they will be better equipped to plan specific strategies for helping these children.

The purpose of the present study was to conduct a neuropsychological assessment of late preschool and early school-age children who had a history of neonatal IVH, in order to determine whether there were residual effects, and to differentiate these effects according to severity of the hemorrhage. The major objective of the study was to determine whether children ages 4 and 5 who had been diagnosed by CT and/or ultrasound as having experienced Grades I, II, III, or IV IVH had generally poorer outcome than a normal population, and whether grades I and II differed neuropsychologically from Grades III and IV.

Methods

This study was conducted as a cooperative effort between Utah State University's Early Intervention Research Institute (EIRI), Primary Children's Medical Center (PCMC), and the University of Utah Medical Center. Children born from 1980 to 1981 were eligible for the study. Both medical centers had a complete medical record on each child describing the incident of IVH, how it was diagnosed (ultrasound and/or CT), the severity of IVH, the type of treatment received in the NICU, as well as family history and medical data.

Sample

Medical records available for each child eligible for the study were made available from both medical centers. The medical centers serve as the major referral source for intermountain area infants with life-threatening conditions. Children eligible for the study met the following qualifications:

1. Birth-weight less than 2700 g.
2. Gestational age less than 36 weeks.
3. Diagnosed IVH by ultrasound or CT.
4. Currently 4 or 5 years old.

Potential children, based initially on information obtained from the medical records, were grouped into a mild IVH category (Grades I and II) and a severe IVH category (Grades III and IV).

Verification that children were still living was obtained through the State Department of Vital Statistics in the respective states. This verification was established before any contact was made with the parents of potential subjects.

The parents of children eligible for the sample were then contacted with an initial mailing from the respective medical center. This mailing contained information on the nature and purpose of the study. A request for parents to include their child in the study was made by the director of the NICU where the children had been treated after birth. Parents who indicated an interest in the study by returning an enclosed postcard were then contacted by the researchers who presented the purpose of the study in more detail. Parents were further briefed on the nature of the study and given more information relating to its importance. Parents who decided to include their child in the study were contacted by the Early Intervention Research Institute to schedule an appointment for testing. Twenty-five subjects were obtained from this initial effort.

A second mailing from the respective medical center was sent to parents who had not responded to the initial mailing. The same process of recruitment was carried out with parents who responded to this mailing. Four more subjects were contacted from this effort. The final sample consisted of 13 for the mild IVH group and 16 for the severe IVH group for a total sample of 29.

All children were treated in NICUs to control for early post-natal differences. It is known from past studies (Stewart et al., 1981) that children treated in NICUs usually have a better chance at recovery from post-natal complications than those who have not been in NICUs.

Instrumentation

Because there has been only limited follow-up done on survivors of IVH at ages 4 and 5, it was deemed necessary to use a broad-spectrum battery that covered a variety of neurobehavioral functions and combination of functions. A group of tests was combined to form a battery that was clinically appropriate for assessment of residual effects of IVH. This battery was selected on the basis of a review of current practice in child neuropsychology (Filskov & Boll, 1981; Goldman, Englestein, & Guerry, 1983; Rourke et al., 1983; Rutter, 1982).

Two types of data were collected for each child in the sample. The first type of information was taken from the child's medical record. The second type of data came from the results of the neuropsychology battery. The instruments included in the battery were the McCarthy Scales of Childrens Abilities, the Preschool Language Scale, The Ravens Progressive matrices, and the Peabody Picture Vocabulary test. A brief description of these measures follows.

The McCarthy Scales of Children's Abilities (MSCA). The MSCA was designed to satisfy the need for a single instrument that could be used to determine general intellectual level as well as a child's strengths and weaknesses in important abilities (McCarthy, 1972). Scores derived from systematic observation of a variety of cognitive and motor behaviors are provided for six scales: verbal, perceptual-motor, quantitative, general cognitive, memory, and motor. The scales are appropriate for children from 2-1/2 through 8-1/2 years of age. The content of the tasks is suitable for both sexes, as well as for children from various ethnic, regional, and socioeconomic groups.

Preschool Language Scale (PLS). This instrument is a widely-used instrument for systematically appraising the early stages of language development. The scale is especially useful for evaluating maturational lags, strengths, and deficiencies as they pertain to developmental progress.

The two parts of the PLS are based on the natural dichotomy between auditory comprehension and verbal ability (Zimmerman et al., 1979). The scale consists of a series of auditory and verbal language tasks, each of which is assigned to a certain age level. According to Zimmerman et al., all items in the PLS have been selected on the premise that in any language, a child's auditory comprehension and verbal ability develop according to capacity, maturation, and life experiences in a spiraling, sequential advancement.

Raven's Progressive Matrices (RPM). The RPM provide a means to assess a person's present ability to think clearly, irrespective of past experiences or present ability for verbal communication (Ravens, 1977). The colored progressive matrices, which is an additional set of problems used with younger children, was constructed to assess in greater detail the ability to complete continuous patterns which, towards the end of the set, change first in one and then in two directions at the same time. A second set of problems requires the ability to see discrete figures as spatially related wholes and to choose a figure which completes the missing part. A third set contained problems requiring abstract thinking.

Peabody Picture Vocabulary Test (PPVT). The PPVT is designed primarily to measure receptive (hearing) vocabulary (Dunn & Dunn, 1981). It has been found useful for a number of school, clinical, and research purposes. Since the PPVT is a reasonably good measure of scholastic aptitude, it should also be useful as an initial screening device in scanning for bright, low-ability, and language impaired children who may need special attention.

Not requiring subjects to read or write makes the scale especially fair for non-readers and other persons with written language problems (Dunn & Dunn, 1981). Also, because neither pointing or oral responses are essential, even severely handicapped individuals are able to be tested.

Data Collection

Descriptive medical data was collected from the medical records. Based on a review of current literature and on consultation with the neonatologists at PCMC and University of Utah Medical Center, the following information was collected: sex, age at time of testing, gestational age, and birthweight. Information was also collected on the following sequelae that commonly occur with LBW and/or IVH: seizure disorder birth asphyxia, post-hemorrhagic hydrocephalus, apnea, hyaline membrane disease, and hyperbilirubinemia. Other information was collected on the type of treatment that occurred in the NICU following an IVH. Three medical procedures were relevant: whether or not the infants received an exchange transfusion, lumbar puncture, or ventriculo-peritoneal shunt, The test protocols were administered and scored by trained examiners. The examiners were "blind" about all previous information for each child, including severity of IVH. Scoring and completeness of administration was checked by the research team project coordinator.

Results

Two types of data were collected and analyzed. First, demographic data were collected from the medical records. Second, and because IVH typically occurs within the context of LBW, medical data indicating the presence or absence of other LBW sequelae were also collected from the medical records. These data were analyzed in terms of their relationship to the performance of the mild and severe IVH groups on the neuropsychological testing. The second type of data collected were based on the outcome measures of each of the scales used in the neuropsychological assessment battery. These data were analyzed to answer the primary research question. The demographic data for the 29 subjects is shown in Table 63.

Table 63

Demographic Characteristics of the Sample

Variables	Mild IVH (N = 13)			Severe IVH (N = 16)		
	Mean	SD	%	Mean	SD	%
Sex: male			62			65
female			38			35
Age at testing (mo.)	54.0	4.5		57.5	7.1	
Gestational Age (wks.)	30.5	2.8		31.1	3.5	
Birthweight (grams)	1495.4	473.3		1526.7	521.9	

Medical data, collected from the subjects' medical records, is shown in Table 64. This data represents common sequelae that typically occur along with LBW and IVH. Post-hemorrhagic hydrocephalus and porencephalic cyst do not usually occur with just LBW; they are sequelae that follow the occurrence of severe IVH. The percent of occurrence of each sequelae for the mild and severe IVH group is indicated.

Table 64

Medical Characteristics of the Sample

Sequelae	Mild IVH (n = 13)		Severe IVH (n = 16)	
	No.	%	No.	%
Apnea	9	69	10	63
Birth Asphyxia	1	8	4	25
Seizure Disorder	1	8	4	25
Hyperbilirubinemia	11	85	10	63
Hylane Membrane Disease	13	100	12	75
Post-Hemorrhagic Hydrocephalus	2	15	9	56*

* = Significant difference ($p < .05$)

The initial assessment results indicated that both groups of IVH preschoolers are significantly behind their normal age mates at 4 and 5 years of age. Table 69 presents the means and standard deviations for both groups on all four measures.

From Table 64 it may be seen that the IVH preschoolers in this sample continue to lag behind their age mates by over one standard deviation. Their delays are most apparent in both motor and perceptual performance, the areas which appear to be most affected by intraventricular hemorrhage.

Table 65

Means and Standard Deviations for Mild and Severe Groups on Four Tests

	Mild (N = 13)		Severe (N = 16)	
	$\bar{X}$	SD	$\bar{X}$	SD
McCarthy Scales				
General Cognitive	79.3	33.3	81.1	39.1
Verbal	42.6	18.2	44.6	21.3
Perceptual Performance	35.6	17.1	36.8	19.8
Quantitative	40.3	16.2	40.8	20.4
Memory	42.3	16.2	42.0	20.5
Motor	30.5	15.7	32.6	15.8
Peabody Picture Vocabulary	84.0	33.7	82.8	38.6
Ravens Progressive Matrices	8.3	6.0	9.1	6.0
Preschool Language Scale				
Language Quotient	94.8	33.4	86.8	40.0
Auditory Comprehension	96.3	33.[illegible]	88.1	40.2
Verbal Ability	93.2	34.6	84.9	39.7

In previous follow-up studies, there appeared to be differences in outcomes between Grades I, and II IVH and Grades III and IV. A series of analyses were done next to ascertain if the two groups differed on outcome measures.

A discriminant analysis was first performed to determine if the mild and severe IVH groups could be distinguished on the basis of the test scores. This technique provides a measure of success with which the discriminating variables actually discriminate between two or more groups (Kerlinger, 1979; Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975).

The results of the discriminant analysis of mild and severe IVH groups and the neuropsychological testing yielded no statistical significance (chi-square = 21.16, d.f = 19, $p < .3$). It was not possible to distinguish performance on the neuropsychological testing based on severity of the hemorrhage. Prediction of group membership (mild or severe IVH) could not be made based on these test results.

Individual t-test analysis of each scale used in the neuropsycho-logical testing further indicated no significant difference between the mild and severe IVH groups. The severe IVH group actually performed slightly better on most of the scales than the mild IVH group. Although it is known that severe hemorrhaging typically causes more severe motor handicaps, one would expect the severe group to exhibit more severe motor delays, these results indicate that even on the McCarthy motor subscale there was no significant difference between the mild and severe groups, although scores for both groups were below normal.

Pearson Correlational Coefficients were computed between the neuropsychological test results and the medical characteristics of the sample listed earlier. As shown in Table 66, significant correlations were found between those subjects that had documented episodes of seizure disorder in the early post-natal period and performance on all measures. These correlations ranged from .35 to .51. Other significant correlations, although not as strong, were found between those subjects that experienced birth asphyxia and performance on the PPVT and the PLS. The lowest correlation on these language related measures was .33 for the PLS verbal ability and the highest was .34 for the PPVT. Significant correlations were also found between those subjects diagnosed with an apneic episode and their performance on the Ravens (.27) and on the McCarthy Performance subtest (.27). There were no significant correlations between test performance and hyperbilirubinemia, hyaline membrane disease, or post-hemorrhagic hydrocephalus.

These results indicate a relationship between a history of both seizure disorder and birth asphyxia and performance on the neuropsychological testing. The presence

Table 66

Pearson Correlational Coefficients Between Other Sequelae of LBW/IVH and Neuropsychological Test Performance

	A	BA	SD	HBR	HMD	PHH
McGenCog	.09	.21	.44C	-.09	-.08	-.03
McVerbal	-.03	.21	.35B	-.10	-.07	-.01
McPerformance	.27A	.18	.47C	-.13	-.08	-.09
McQuantitative	.09	.22	.46C	-.10	-.10	-.01
McMemory	-.06	.21	.39C	-.12	-.07	.06
McMotor	.20	.24	.43C	-.14	-.09	.00
PPVT	.00	.33B	.46C	.06	-.12	-.09
PLSLQ	.00	.34B	.45C	.06	.12	-.05
PLSAC	.04	.33B	.45C	-.09	-.16	.05
PLSVA	-.05	.34A	.45C	-.01	-.06	.05
RPM	.27A	.01	.51C	.01	.01	-.06

Note: A = Apnea, BA = Birth Asphyxia, SD = Seizure Disorder, HBR = Hyper-bilirubinemia, HMD = Hylane Membrane Disease, PHH = Post-Hemorrhagic Hydrocephalus, Mc... = McCarthy Scales

A = $p < .10$, B = $p < .05$, C = $p < .01$

or absence of neonatal seizure disorder had the strongest correlations with outcome measures on the neuropsychological assessment.

Summary of the Findings

The results indicate that IVH groups continue to perform below that of their age mates up to age 6. Based on these results, one general conclusion that could be drawn is that the occurrance of an IVH at birth does predict children's cognitive

A more relevant issue is that the severe group had a greater incidence of seizure disorder than the mild group (25% and 8%, respectively). Although this did not show up as a statistically significant difference, it would be expected that this may have further impaired the severe IVH group in their performance. The results, however, did not show this to be the case.

There was also a significant correlation found between the incidence of birth asphyxia and performance on the PPVT and PLS (.18 - .34). Although not as strong a relationship as shown with seizure disorder, children who had documented episodes of birth asphyxia performed below group averages. Again, there was a greater incidence of birth asphyxia in the severe IVH group (25%) than in the mild IVH group (8%). This difference was not found to be statistically significant, and it did not seem to have a deleterious effect on the performance of the severe group.

In general, there was no evidence shown from the results that the lack of significant findings was due to differences in the medical characteristics of the group as described by the data collected from the medical records. There may have been other differences between the groups that were not identified within the parameters of the study. At the present time, it is not known what these other differences are, and how they could potentially effect the findings. Based on the information and findings of this study, however, the present results conflict with at least some of the findings of previous follow-up that indicate some recovery of cognitive functioning with continued development.

The lack of significant differences found between severity of IVH and test performance is urprising based on previous follow-up studies. Unlike several previous studies, these results indicate that both IVH groups are still behind their normal age mates. It is not known whether this difference will continue to be evident as these children continue to develop. Further follow-up is planned as these children move into the first and second grade to determine how IVH at birth affects later academic performance.

A more relevant issue is that the severe group had a greater incidence of seizure disorder than the mild group (25% and 8%, respectively). Although this did not show up as a statistically significant difference, it would be expected that this may have further impaired the severe IVH group in their performance. The results, however, did not show this to be the case.

There was also a significant correlation found between the incidence of birth asphyxia and performance on the PPVT and PLS (.18 - .34). Although not as strong a relationship as shown with seizure disorder, children who had documented episodes of birth asphyxia performed below group averages. Again, there was a greater incidence of birth asphyxia in the severe IVH group (25%) than in the mild IVH group (8%). This difference was not found to be statistically significant, and it did not seem to have a deleterious effect on the performance of the severe group.

In general, there was no evidence shown from the results that the lack of significant findings was due to differences in the medical characteristics of the group as described by the data collected from the medical records. There may have been other differences between the groups that were not identified within the parameters of the study. At the present time, it is not known what these other differences are, and how they could potentially effect the findings. Based on the information and findings of this study, however, the present results conflict with at least some of the findings of previous follow-up that indicate some recovery of cognitive functioning with continued development.

The lack of significant differences found between severity of IVH and test performance is urprising based on previous follow-up studies. Unlike several previous studies, these results indicate that both IVH groups are still behind their normal age mates. It is not known whether this difference will continue to be evident as these children continue to develop. Further follow-up is planned as these children move into the first and second grade to determine how IVH at birth affects later academic performance.

Products

Number in Product List	Title
Dissertation	A Neuropsychological Follow-up of Infants Experiencing Intraventricular Hemorrhage
Manuscript Submitted	A Neuropsychological Follow-up of Infants Experiencing Intraventricular Hemorrhage

INVESTIGATION 14:

DOWN SYNDROME PARENT INVOLVEMENT STUDY

The initial results of the meta-analysis of the early intervention research literature were somewhat startling in the area of parental involvement. Briefly, the findings suggested that parents could be effective intervenors, that they were not essential to intervention success. In fact, although there have been a relatively small number of parent involvement studies with handicapped preschoolers, the meta-analysis findings indicated that there was very little difference between programs which utilized parents extensively and those which did not (Casto & Lewis, 1984).

Data from nine studies which made direct comparisons between different levels of parental involvement (Abbott & Sabatino, 1975; Bidder, Bryant, & Gray, 1975; Gordon, 1969; Karnes et al., 1970; McCarthy, 1968, Miller & Dyer, 1975; Nedler & Sebra, 1971; Radin, 1971; Ramey & Bryant, 1983) were examined. When all 134 effect sizes from these studies were considered, there was a slight advantage for programs which involved parents more extensively (.08 standard deviations). However, these findings were heavily weighted by the Gordon study which found an average advantage of .18 favoring the involvement of parents. The other eight studies found an average effect size of .05 favoring programs which did not involve parents.

Taken together, the data from these different sources of information suggested that programs which involve parents extensively can be effective, but they are no more effective than programs which do not involve parents. There is little support for the position that involvement of parents leads to more effective intervention programs. Admittedly, this is counter to what most people assume and what is intuitively logical. It is important to note that most of these data come from the disadvantaged population literature. Very little information was available on whether involving parents leads to more effective early intervention programs for handicapped children. The results of the meta-analysis also suggested a need for more tightly controlled studies utilizing true experimental designs and led to the next phase of the research program.

Experimental Studies

Further support for the lack of parental involvement effect as indicated by the meta-analysis came from an experimental study conducted by DCHP staff. Peterson, Casto, and Lindauer (1985) conducted a study in which 50 developmentally delayed preschoolers were randomly assigned to treatment and no-treatment conditions. Children in the treatment condition group received direct instruction from an itinerant teacher using the CAMS materials, while the no-treatment group received traditional Head Start intervention. The treatment group performed significantly better on developmental tests than did the no-treatment group after the 16-week intervention in the absence of any parental involvement.

Faced with additional evidence that a no-parent involvement (Casto & Mitchell, 1977) intervention produced gains similar to the original parent-as-intervenor findings, Pezzino and Lauritzen (1986) designed a more sophisticated parent training package (PIE) to provide parents with a systematic conceptual and hands-on experience in areas such as child development, observation and recording, targeting intervention behaviors, teaching processes, decision making, and communicating with professionals.

The impacts of implementing the PIE training package with parents was assessed in two separate recently conducted students. In the first study, Pezzino and Bradley (1986) investigated the effectiveness of three levels of parental involvement with handicapped preschoolers on measures of child developmental progress, children's adaptive behavior, family rearing attitudes, marital adjustment, and family relationships. Forty-five mildly handicapped preschoolers and their parents were selected to participate as subjects in this study. Subjects were matched prior to random assignment to treatment groups according to chronological age and child's type of handicapping condition. Subjects were then assigned to one of three treatment groups which represented different levels of parental involvement in their child's preschool intervention programs. In the "high parent involvement group" parents and children participated in a center-based intervention program that included a home component and structured parent training sessions using the PIE materials. In the "low parent involvement group," children participated in the same type of center-based program, but parents were not involved. The third treatment group was a comparison group in which neither parents nor children participated in a center-based classroom or other interventions. Results indicated no significant group differences on any of the child progress measures or the parent and family outcome measures.

In the second, study Pezzino, Mott, and Waidler (1986) conducted an investigation to determine the child and family impacts of adding a structured parent involvement program to an existing center-based early intervention program which included a minimal level of parental participation. Fifty-one moderately and severely retarded preschool-age children and their parents were randomly assigned to either the parent training or non-parent training group. Prior to random assignment, subjects were stratified on children's chronological age and performance on developmental pretest measures. The parent training sessions were based on the PIE training curriculum, and the treatment lasted approximately six months.

Posttest child outcome measures included standardized tests of child development and progress in achieving individual education plan objectives. Posttest family outcome measures included standardized tests of stress and coping, general family functioning, degree of social support and resources, and parental knowledge and attitudes.

Pre- and posttest gain scores on the standardized child developmental measures indicated that the basic intervention program was effective for both groups of children. Both groups of children demonstrated developmental progress equivalent to one month for every month of intervention. However, with the exception of a parent knowledge survey which was based directly on the PIE training materials, the child progress and family outcome measures exhibited virtually no statistically significant differences between the two groups. There was some indication, based on one Parenting Stress Index (Abindin, 1983) subtest, that the non-parent training group mothers were experiencing significantly lower levels of stress than the parent training group mothers.

Based upon the integrative review findings, the findings of EIRI primary research studies, and input received from numerous experts in the field, we have now concluded that family involvement can be categorized into programs which attempt to train family members or programs which provide support for family members. Each of these broad categories can be easily subdivided, as described below.

Training Programs

Programs which attempt to train family members as a part of early intervention for a handicapped child, generally focus training in one of the following areas, or in some combination of these areas.

1. Training family members to provide specific therapeutic activities focused on those areas in which the child is delayed or experiencing difficulty. Such training may include speech and motor therapy· activities designed to increase cognitive functioning; child management skills such as dressing, feeding, and bathing; or activities designed to enhance social/emotional functioning. Training in this category may ask parents to provide structured therapy in a specifically set-aside time each day, or may ask parents to incorporate the therapeutic techniques into normal day-to-day activities.

2. Training designed to enhance the parent/child interactions in order to solidify parent/child bonding and attachment, and to create an atmosphere in which the child is viewed as an important contributing member of the family who has different skills and abilities, but is, nonetheless, an important person.

3. Training in behavior management skills. Programs in this category are similar to the widely disseminated packages such as Systematic Training for Effective Parenting (STEP), or Parent Effectiveness Training (PET). These training programs focus on helping parents to effectively manage their children's behavior by providing consistent feedback, teaching appropriate rules, and teaching the children responsibility.

4. Teaching child development skills. Some programs provide parents information about child development so that parents can be aware of the sequence in which children acquire skills and the age appropriate milestones

at which such skills are generally acquired. The philosophy behind these programs is that parents can be more effective at helping their handicapped child if they understand the typical developmental sequence for non-handicapped children.

Family Support Programs

As opposed to programs which expect family members to acquire specific skills, the second category of programs is designed to provide support and encouragement to the family in a variety of areas. The activities in such programs can be categorized as follows:

1. **Group Support for Parents.** Parent groups are formed in which parents of handicapped children are assisted in understanding that there are many other parents who have dealt successfully with having a handicapped child. Activities in such programs focus on issues such as working through the grieving process, dealing with the frustrations and adjustments of having a handicapped child, learning how you are not alone in having a handicapped child, understanding your legal rights and programmatic opportunities, and understanding the rationale behind various service programs.

2. **Individual Parent Support.** Some programs provide one-on-one support services to parents of handicapped children, sometimes referred to as pilot parent programs or parent-to-parent programs. The activities in this category focus on many of the same topics as the above, but do it on a one-to-one setting, where parents who have been through the process meet individually with parents of newly identified handicapped children to assist them in understanding the consequences of having a handicapped child, and learning about the support systems that are available. A fairly recent addition to the area of group parent support, is the involvement of fathers as well as mothers of the handicapped children in this process.

3. **Sibling support.** Usually conducted in group settings, siblings of handicapped children are convened to provide them with a network of friends who have similar frustrations and experiences with a handicapped brother or sister. The rationale behind such programs appears to be to communicate to siblings of handicapped children that they are not alone, and that other children have successfully dealt with having a handicapped sibling, in addition to helping them understand the valuable attributes of their handicapped sibling.

4. **Respite Care.** Some programs provide respite care services as a part of the family support program for both short-term and long-term situations. Care for the handicapped child is arranged so that parents and/or family members can be freed from the constant responsibility of caring for the handicapped child.

5. **Access to Services.** Another form of support is to assist parents and family members in accessing available services which will assist in dealing with their handicapped child. Such services range from food stamps, to nutrition counseling, to learning about public transportation opportunities, to

learning about available therapeutic programs. Some programs have even organized to assist parents in enhancing their educational or employment status. These programs are usually conducted by an individual case manger, who meets individually with the family to assess their needs and to then acquire the needed resources to meet those needs.

Based upon the above conceptualization, we now feel that our research in parental involvement should focus on: (1) comparing different intensities of parental support in much of the same manner that we have compared different intensities of parent training, (2) ascertaining which component of parental involvement (training versus support) results in the most favorable outcomes for child and family, and (3) conducting further research which teases out which specific components of training and support are most effective. The research reported here focused on area one, comparing different intensities of parental support.

Experimental Design

A pretest/posttest control group experimental design was used. Children who qualified for the program were matched on age and sex and randomly assigned to one of two experimental conditions. Condition 1 was an intensive program of support services, while Condition 2 was a minimal program of support services consisting of the present group support-only program.

Sample

The sample for the study included 29 preschoolers ages 0-3. All subjects will were diagnosed as having Down syndrome.

The research sample included the first 29 children enrolled in the Baton Rouge, Louisiana, intervention program. The enrollment period ended on November 1, 1986, and treatment began immediately after.

Experimental Conditions

The high intensity program of support services included the following components.

1. **Individual Parent Support.** This component provided support services to parents by having the parents of older Down syndrome preschoolers provide one-on-one support to the parents of the children enrolled in the experimental condition. The experimental group parents received information from the parents of older children which allowed them to know the services available, their legal rights, and provided information designed to help them understand their child's handicapping condition better.

2. **Group Support for Parents.** This level of support was currently being offered in the program. Parent groups were formed in which parents of handicapped children were assisted in understanding that there are many other parents who have dealt successfully with having a handicapped child. Activities in this program focused on issues such as working through the grieving process, dealing with the frustrations and adjustments of having a handicapped child, having parents learn they are not along in having a handicapped child, understanding their legal rights and programmatic opportunities, and understanding the rationale behind various service programs.

3. **Sibling Support.** The level of support focused on the siblings of the handicapped child. The siblings of the experimental group preschoolers were convened in order to provide them with a network of friends who had similar frustrations and experiences with a handicapped brother or sister. The rationale behind such programs is to communicate to siblings of handicapped children that they are not alone, and that other children have successfully dealt with having a handicapped sibling, in addition to helping them understand the valuable attributes of their handicapped sibling.

4. **Respite Care.** This program provided short-term respite care for the families of experimental group children. Care for the handicapped preschooler was arranged on a regular basis so that parents and family members could be freed from the constant responsibility of caring for the handicapped child.

Instrumentation

The core measures selected for the pretest represented a range of variables which are thought to be important outcomes of increased family support. For example, it may be that supports add very little for a mild to moderately mentally retarded child who lives in a close-knit, well-educated, financially secure family; whereas it may be very beneficial for a similar child who lives in a single parent low socioeconomic family. Such aptitude by treatment interactions or 'value added' hypotheses, were able to be investigated because adequate demographic and family functioning data of the type proposed for the pretest core battery was collected.

The core measures selected for posttesting represent the ones with the greatest potential for reflecting overall differences between experimental and control groups. The Battelle Developmental Inventory was selected as a core child measure based on its positive characteristics, as described in the initial proposal, as well as the support obtained from expert recommendations and the validity studies conducted by the institute over a five-year period. The Vineland Adaptive Behavior Scales--Revised was selected as a core posttest child measure. Staff agreed that the child's social and adaptive behavior should be a primary outcome variable for parent support programs.

Assessment of family integrity prior to participation in the research project was accomplished through pretest administration of FACES III and the Family Inventory of Life Events and Changes. Aspects of family functioning which are expected to be impacted by the experimental interventions were assessed pre- and posttest. As their titles imply, family stress, resources, and support were assessed through administration of the Parenting Stress Index, Family Support Scale, and Family Resource Scale. A measure of parent satisfaction with services (which includes a description of all additional services received by the family that might be expected to assist with the conditions caused by the child's handicapping condition), as well as a report of the child's health during the past year, was also collected at posttest time.

The pretest measures were given the weeks of November 3, and treatment began immediately after. Posttesting was done during September, 1987.

Data Analysis

Pre- and posttest data collected were analyzed using multivariate procedures to answer the following questions.

1. Does a program of intensive parental support result in more favorable child outcome?
2. Does a program of intensive parent support result in lowering stress for families?
3. To what extent did experimental group parents participate in each support component?

Results and Discussion

Due to the fact that posttesting was done in September, 1987, all data have not been entered for all the children and their families. Preliminary analysis of data received indicates that child outcome, as measured by the Battelle Developmental Inventory, was significantly different for both groups at posttest, as shown in Table 67.

Table 67

Down Syndrome Parent Involvement Study Difference Between Pre- and Posttest Score Means

Test	2-tailed Probability (Significance of Difference Between pre- and posttest
Battelle Total	.002
Parenting Stress Index Total Score	.509
Parenting Stress Index Child Report	.029
Parenting Stress Index Parent Report	.003

It is important to note that these trends appear strong, and should be further confirmed when all data has been received and coded. This information will be sent as an addendum to this report and will be included in articles written for publication.

When experimental and control groups were compared on the Battelle total score, an effect size of .80 was found favoring the control group.

Although the N is small, it appears that the parent support activities had either a negative or minimal influence on child progress.

The parent support program does result in lowering family stress as measured by the PSI. When experimental and control groups were compared on the parenting stress index, the experimental group had significantly lower stress levels (Es = .77) than the control group.

Experimental group participation is shown in Table 68. This data indicates that although families participated and made positive evaluation of the activities, attendance varied considerably across components.

Table 68

<u>Attendance for Down Syndrome Parent Involvement Components</u>

Activity	Sessions Attended
Lead Parent Contacts	111
Father Groups	17
Teenage Sibling Groups	24
Elementary-Age Get-Together	2
Sibling Preschool Daycare	19

Table 69 shows the contacts made to experimental group parents by parents designated as lead parents. Due to family pressures, these contacts were not made each week as planned. Total individual contacts made by the staff to families in the experimental group was 190.

Table 69

Down Syndrome Parent Involvement Study Number of Contacts By Lead Parents by Month

Month	# of Lead Parent Contacts
January	0
February	9
March	26
April	24
May	22
June	20
July	10

Parent support at best is difficult to define, and more difficult to implement. Activities planned for this project were the result of brainstorming by staff who had worked with families for some time, and activities planned of necessity were those that could be offered with present resources. Preliminary results of the study indicate that these activities may be helpful to families, but families themselves need to have more input into planning parent support activities. In order to provide more extensive parent support, funding for parent support activities may need to be at a higher level, including some form of payment to parents who serve as lead parents.

The preliminary results of this study are difficult to explain. Clearly, the parental support programs had either no impact or a negative impact on child change. The chief observed result of the support program was to decrease parental stress. A replication of the study with a larger number of subjects is needed before conclusions are drawn.

Products

Number in Product List	Title
13	Counseling Families
14	Down Syndrome Review
24	Psycho-Social Development
47	Becoming a Parent Plus
66	Early Intervention Helps Parents

PRODUCT LIST

CURRICULA MATERIALS, GUIDELINES, INSTRUMENTS

1. Cost-Effectiveness Training Manual
2. Meta-Analysis Training Manual
3. Battelle Training Manual
4. Research Site Review Guidelines

BOOKS, CHAPTERS, MONOGRAPHS

Books

5. Pueschel, S. M., Tingey, C., Rynders, J., Crocker, A., & Crutcher, D. (Eds.) (1986). New perspectives on Down syndrome. Baltimore, MD: Brooks Publishing.

6. Tingey, C. (Ed.) (1987). Down syndrome: A resource handbook. San Diego, CA: College Hill Press.

7. Tingey, C. (Ed.) (in press). Early intervention: Hands-on strategies. Baltimore, MD: Brooks Publishing.

8. Worthen, B. R., & White, K. R. (Eds.) (1987). Evaluating educational and social programs: Guidelines for proposal review, on-site evaluation, evaluation contracts, and technical assistance. Boston, MA: Kluwer-Nijhoff.

Chapters

9. Barnett, W. S., & Escobar, C. M. (in press). Economic costs and benefits of early intervention. In S. J. Meisels & J. P. Shonkoff (Eds.), Handbook of early intervention: Theory, practice and analysis. Cambridge, MA: Cambridge University Press.

10. Bell, C., Casto, Y., & Casto, G. (1986). A treatment program for pregnant adolescents and their infants. In. R. Feldman, & A. Stiffman (Eds.), Advances in adolescent mental health (pp. 37-46). Greenwich, CT: JAI Press.

11. Bowell, T., & Tingey, C. (1987). My brother, Jim. In C. Tingey (Ed.), Down syndrome: A resouce handbook. San Diego, CA: College Hill.

12. Carey, J., & Tingey, C. (1987). Counseling families. In C. Tingey (Ed.), Down syndrome: A resource handbook. San Diego, CA: College Hill.

13. Casto, G. (1987). Plasticity and the handicapped child: A review of efficacy research. In J. Gallagher & C. Ramey (Eds.), Plasticity in development, (pp. 103-113). Baltimore, MD: Brookes Publishing.

14. Casto, G. (1897). Early intervention for infants and preschoolers with Down syndrome. In C. Tingey, C. (Ed.), Down syndrome: A resource handbook. San Diego, CA: College Hill.

15. Casto, G. (in press). Research and program evaluation in early childhood special education. In S. Odom & M. Karnes (Eds.), Research and program evaluation in early childhood special education.

16. Casto,G., Gaynard, L., Mobasher, H., Chan. G., Dolcourt, J., Levkoff, A., & Saylor, C. (1987). The efficacy of early intervention programs for low birth weight infants. In G. Casto, F. Ascione, & M. Salehi (Eds.), Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

17. Denenberg, V. (1987). Early stimulation and brain organization in the rat. In G. Casto, F. Ascione, & M. Salehi (Eds.), Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

18. Goncu, A. (1987). The role of adults and peers in the socialization of play during preschool years. In G. Casto, F. Ascione, & M. Salehi (Eds.), Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

19. Greenspan, S. I., & White, K. R. (1987). Conductinng research with preventive intervention programs. In I. F. Berlin & J. Noshpitz (Eds.), Basic handbook of child psychiatry New York: Basic Books.

20. Lewis, M., & Hadzimichalis, D. (1987). Children's attention as a measure of central nervous system integrity. In G. Casto, F. Ascione, & M. Salehi (Eds.), Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

21. Meisels, S. (1987). Using criterion referenced assessment data to measure the progress of handicapped children in early intervention programs. In G. Casto, F. Ascione, & M. Salehi (Eds.), Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

22. Tingey, C. (1987). Cutting the umbilical cord: Parental perspectives. In S. Pueschel (Ed.), The person with Down syndrome: Transition from adolescence to adulthood. Baltimore, MD: Brooks Publishing.

23. Tingey, C. (1987). The elementary school years. In C. Tingey (Ed.), Down syndrome: A resource handbook. San Diego, CA: College Hill.

24. Tingey, C. (1987). Psycho-social development in persons with Down syndrome. In S. Pueschel, C. Tingey, J. Rynders, A. Crocker, & D. Crutcher (Eds.), New perspectives on Down syndrome, (pp. 311-334). Baltimore, MD: Brooks Publishing.

25. Tingey, C. (in press). Social and emotional enhancement. In C. Tingey (Ed.), Early intervention: Hands-on strategies. Baltimore, MD: Brooks Publishing.

26. White, K. R. (in press). Cost analyses in family support programs. In H. Weiss & F. Jacobs (Eds.), Evaluating family programs. Hawthorne, NY: Aldine.

27. White, K. R. (1983). The relation between socioeconomic status and academic achievement. In R. J. Light (Ed.), Evaluation studies review annual (Vol. 8). Beverly Hills, CA: Sage Publications.

28. White, K. R. (1985). Review of Comprehensive ability battery. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook (pp. 377-379). Highland Park, NJ: Gryphon Press.

29. White, K. R. (1985). Review of Assessment of skills in computation. In J. V. Mitchell (Ed.), The ninth mental measurements yearbook (pp. 103-105). Highland Park, NJ: Gryphon Press.

30. White, K. R., & Greenspan, S. I. (1987). An overview of effectiveness of preventive intervention programs. In F. Berlin & J. Noshpitz (Eds.), Basic handbook of psychiatry. New York: Basic Books.

Monographs

31. Casto, G. (1985). Common outreach indicators. Monograph of the Technical Assistance Development system. Chapel Hill, NC.

32. Casto, G., Ascione, F., & Salehi, M. (Eds.) (1987). Current perspectives in infancy and early childhood research. Logan, UT: Early Intervention Research Press.

33. Casto, G., & Lewis, A. (1985). A planning process for recruiting and retaining early interventio personnel in rural areas. Monograph of the Rural Network. Western Illinois Univesity Press.

ARTICLES, REVIEWS, PROCEEDINGS

34. Barnett, W. S. (in press). The economics of early intervention under PL 99-457. Topics in Early Childhood Special Education.

35. Barnett, W. S. & Escobar, C. M. (in press). The economics of early intervention for handicapped children: What do we really know? Journal of the Division for Early Childhood.

36. Barnett, W. S. (in press). A cross cultural investigation of economics and mental retardation. In S. C. Hey (Ed.), Proceedings of the Society for Disability Studies. Salem, OR: Willamette University.

37. Barnett, W. S., & Pezzino, J. (1987). Cost-effectiveness analysis for state and local decision making: An application to half-day and full-day preschool special education programs. Journal of the Division for Early Childhood, 11, 171-179.

38. Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., & Weikart, D. P. (1987). The effects of early educational intervention on crime and delinquency in adolescence and early adulthood. In J. D. Burchard and S. N. Burchard (Eds.). Prevention of delinquent behavior (pp.220-240). Beverly Hills, CA: Sage.

39. deAyora, P., & White, K. R. (in press). Using the Minnesota Child Development Inventory as a measure of developmental progress with handicapped children. Journal of Psychoeducational Assessment.

40. Fifield, M., & Casto, G. (1987). Utah State University Developmental Center for Handicapped Persons. Journal of the Association for Persons with Severe Handicaps, 11, 271-275.

41. Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). Early intervention for behavior disordered children: The synthesis of single subjct research. Behavioral Disorders Monographs, 9.

42. Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). Reply to Owen White. Remedial and Special Education, 8, 49-52.

43. Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). Response to Salzberg, Strain, and Baer. Remedial and Special Education, 8, 49-52.

44. Tingey, C. (1987). What's in an A? Early Years, 17(3), 85-87.

45. Tingey, C. (1987). Becoming a parent plus. Early Childhood Development & Care, 28(1).

46. Tingey, C. (1987). Advice on good grooming. The Exceptional Parent, 17(3), 10-26.

47. Tingey, C., Boyd, R., & Casto, G. (1987). Parent involvement in early intervention: Becoming a parent plus. Early Child Development and Care, 28, 91-105.

48. White, K. R. (1987). Causes of academic failure among children with normal IQ. A review of S. Broman, E. Bien, and P. Shaughnessy Low achieving children: The first seven years. Contemporary Psychology, 32(3).

49. White, K. R., & Mott, S. E. (in press). Conducting longitudinal research about the efficacy of early intervention with handicapped children. Journal of the Division for Early Childhood.

50. Arnold, K. R., Myette, B., & Casto, G. (1986). Relationships of language intervention efficacy to certain pretreatment subject characteristics in mentally retarded preschool children: A meta-analysis. Education and Treatment of Children.

51. Barnett, W. S. (1986). Methodological issues in economic evaluation of early intervention programs. Early Childhood Research Quarterly, 1(3), 249-268.

52. Barnett, W. S. (1986). The transition from state residential schools for retarded people to custodial facilities: An economic explanation. Disability, Handicap and Society, 1(1), 53-71.

53. Barnett, W. S., (1986). Definition and classification of mental retardation: A reply to Zigler, Balla, and Hodapp. American Journal of Mental Deficiency, 91, 111-116.

54. Bush, D. W., & White, K. R. (1986). Reply to Weiss and Rosenberg. Psychological Reports, 58, 534.

55. Casto, G. & Mastropieri, M. (in 1986). Strain and Smith doth protest too much: A reply. Exceptional Children, 53, 266-268.

56. Casto, G. & Mastropieri, M. (1986). Much ado about nothing: A reply to Dunst and Snyder. Exceptional Children.

57. Casto, G. (1986). Family assessment. DEC Communicator, 12, 1-2.

58. Casto, G., & Mastropieri, M. (1986). The efficacy of early intervention programs for handicapped children: A meta-analysis. Exceptional Children, 52, 417-424.

59. Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research: Methodology and validation. Remedial and Special Education, 8.

60. Mastropieri, M. A., & Scruggs, T. E. (1986). Early intervention for socially withdrawn children. Journal of Special Education, 19, 429-442.

61. Mastropieri, M., White, K. R., & Fecteau, F. (1986). Introduction to special education textbooks: What they say about the efficacy of early intervention. Journal of the Division for Early Childhood, 11, 59-66.

62. Roopnarine, J. L., Mounts, N. S., & Casto, G. (1986). Mother's perceptions of their children's supplemental care experience: Correlation with spousal relationship. American Journal of Orthopsychology, 56, 581-588.

63. Scruggs, T. E., Bennion, K., & White, K. R. (1986). Training test-taking skills to elementary school children: A meta-analysis. Elementary School Journal, 87, 69-82.

64. Tingey, C. (1986). Early intervention--learning what works. The Exceptional Parent, 16(7), 32-37.

65. Tingey-Michaelis, C. (1986). The importance of structure in early education programs for disadvantaged and handicapped children. Early Childhood Development and Care, 23, 283-297.

66. Tingey-Michaelis, C. (1986). Early intervention helps parents, too. The Exceptional Parent, February, 51-52.

67. White, K. R. (1986). The efficacy of early intervention. Journal of Special Education, 19(4), 401-416.

68. White, K. R., & Pezzino, J. (1986). Ethical, practical, and scientific considerations of randomized experiments in early childhood special education. Topics in Early Childhood Special Education, 6, 100-116.

69. Barnett, W. S. (1985). Benefit-cost analysis of the Perry Preschool program and its policy implications. Educational Evaluation and Policy Analysis, 7, (4).

70. Bush, D. W., & White, K. R. (1985). Questionnaire distribution: A method that significantly improved return rates. Psychological Reports, 56, 427-430.

71. Casto, G., & Lewis, A. (1985). Selecting outcome measures in early inter ···on. Journal of the Division of Early Childhood.

72. Casto, G., & White, K. R. (1985). The efficacy of early intervention programs with environmentally at-risk infants. Journal of Children in Contemporary Society, 17, 37-50.

73. Cook, S., Scruggs, T. E., Mastropieri, M. A. & Casto, G. (1895-86). Handicapped students as tutors: A meta analysis. Journal of Special Education, 19, 489-506.

74. Greenspan, S. E., & White, K. R. (1985). The efficacy of preventive intervention: A glass half full? Zero to Three, 5, 1-5.

75. Schweinhart, L. J., Berrueta-Clement, J. R., Barnett, W. S., epstein, A. S., & Weikart, D. P. (1985). The promise of early childhood education. Phi Delta Kappan, 66(8), 548-554.

76. Tingey-Michaelis, C. (1985). Stress: Getting out from under. Early Years K-8, 15(5), 35.

77. Tingey-Michaelis, C. (1985). Early intervention: Is certification necessary? Teacher Education and Special Education, 8(2), 91-96.

78. Tingey-Michaelis, C. (1985). Psychological mirror on the wall. Early Years, 15(8), 16+.

79. White, K. R., Bush, D. W., & Casto, G. (1985). Let the past be prologue: Learning from previous reviews of early intervention efficacy research. Developmental Review.

80. Casto, G., & Lewis, A. (1984). Parent involvement in infant and preschool programs. Journal of the Division for Early Childhood, 9(1), 49-56.

81. Mastropieri, M. (1984). Age at start and early intervention effectiveness: A correlational analysis. Pediatrics.

82. Mastropieri, M. A. & Scruggs, T. E. (1984). Generalization of academic and social behaviors: Five effective strategies. Academic Therapy, 19, 427-431.

83. Mastropieri, M., Scruggs, T. E., & Casto, G. (1984). Early intervention for behaviorally disordered children: An integrative review. Behavioral Disorders.

84. Mastropieri, M. A., Scruggs, T. E., & Levin, J. R. (1984). research in progress: Mnemonic strategies for handicapped and gifted learners. Exceptional Children, 50, 559.

85. Osguthorpe, R. T., Scruggs, T. E., & White, K. R. (1984). Handicapped children as tutors. Exceptional Children, 50, 466.

86. Scruggs, T. E., & Mastropieri, M. A. (1984). Issues in generalization: Implications for special education. Psychology in the Schools, 21, 397-403.

87. Taylor, C., White, K. R., & Pezzino, J. (1984). Cost-effectiveness analysis of full-day versus half-day intervention programs for handicapped preschoolers. Journal of the Division for Early Childhood, 9(1), 76-85.

88. Tingey-Michaelis, C. (1984). Developing the whole child. Early Years K-8, 15(4), 8+, 28-29.

89. Tingey-Michaelis, C. (1984). Gifted kids are not all gift wrapped. Early Years, 14(5).

90. Tingey-Michaelis, C. (1984). Day one! How to handle that all-important first day of school. Early Years K-8, 15(1), 37-38.

91. Tingey-Michaelis, C. (1984). Ordinary is alright. Early Years K-8, 15(3), 28-29.

92. White, K. R. (1984). Is early intervention effective? Outreach Project Newsletter, 6.

93. White, K. R., & Casto, G. (1984). An integrative review of early intervention efficacy studiesa with at-risk children: Implications for the handicapped. Analysis and Intervention in Developmental Disabilities.

94. White, K. R., Matropieri, M., & Casto, G. (1984). An analysis of special education early childhood projects approved by the Joint Dissemination Review Panel. Journal of the Division for Early Childhood, 9, (1), 11-26.

95. White, K. R., & Worthen, B. W. (1984). Improving state and local education programs: Two outcomes of a statewide evaluation. Educational Evaluation and Policy Analysis, 6, 253-266.

96. Casto, G., White, K. R., & Taylor, C. (April 1983). Studies of the efficacy and cost effectiveness of early intervention. Journal of the Division of Early Childhood.

97. Casto, G., White, K. R., & Taylor, C. (1983). An Early Intervention Research Institute: Efficacy and cost studies in early intervention. Journal of the Division for Early Childhood, 7, 5-17.

98. Tingey-Michaelis, C. (1983). Make room for movement. Early Years, 13(3), 26-29+.

99. Tingey-Michaelis, C. (1983). Repetition, relaxation, routine. The Exceptional Parent, 13(3), 52-54.

100. Tingey-Michaelis, C. (1983). Home work vs. school work. The Exceptional Parent, 13(4), 36-37.

OTHER MANUSCRIPTS

101. Barnett, W. S., & Escobar, C. M. (1987). Economic evaluation of early intervention: A Critcal Revciew. Manuscript submitted for publication.

102. Barnett, W. S., Escobar, C. M., & Ravsten, M. (1987). Parent and clinic early intervention for children with language handicaps: A cost-effectiveness analysis. Manuscript submitted for publication.

103. Boyd, R. D., Welge, P., Sexton, D., & Miller, J. H. (1987). Concurrent validity of the Battelle Developmental Inventory: Relationship with the Bayley Scales in young children with known or suspected disability. Unpublished manuscript, Louisiana State University, New Orleans.

104. Escobar, C. M., Barnett, W. S., & Keith, J. E. (1987). Contingent valuation of an early intervention program for handicapped preschoolers. Manuscript submitted for publication.

105. Casto,G., Mitchell, H., & Corey, W. (1986). A comparative analysis of head start efficacy research. Unpublished manuscript, Utah State University, Early Intervention Research Institute, Logan.

106. Casto, G., Tingey-Michaelis, C., & Crutcher, D. (1986). Early intervention for preschoolers with Down syndrome: A review. Unpublished manuscript, Utah State University, Early Intervention Research Institute, Logan.

107. Mott, S. E. (1986). Concurrent validity of the Battelle Develomental Inventory for speech and language disordered children. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

108. Mott, S. E., & Casto, G. (1986). Annotated bibliography of self-report measures of family functioning. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

109. Pezzino, J., Mott, S., & Waidler, J. (1986). Concurrent validity of the Battelle Developmental Inventory. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

110. Mehren, M., & White, K. R. (1986). Parent tutoring as a supplement to compensatory education for first grade children. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

111. Casto, G. (1985). The relationship between program intensity and duration and efficacy in early intervention. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

112. Austin, A. (1984). Establishing a philosophical framework for programs for young children. Unpublished Manuscript, Utah State University, Early Intervewntion Research Institute, Logan.

113. Reeder, D., & Casto, G. (1984). Parental involvement in early intervention: A review and critique. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

114. Casto, G., & Lewis, A. (1984). Early intervention: Is earlier better? Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

115. Goodrich, G. (1984). The research of Harold Skeels: Contributions to psychology and early education. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

116. Peterson, A., Casto, G., & Lindauer, S. L. (1984). Improving pre-academic skills in developmentally delayed preschoolers through use of a highly structured cognitive intervention program. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

117. Watkins, S., & Casto, G. (1984). A longitudinal study of early interveniton with hearing impaired children. Unpublished Manuscript, Utah State University, Early Intervention Research Institute, Logan.

MAJOR PRESENTATIONS AND WORKSHOPS

1987

118. Barnett, W. S. (1987, July). The economics of compensatory education: New evidence from experimental studies. Paper presented at the Annual Meeting of the Western Economic Association, Vancouver, B.C.

119. Barnett, W. S., & Escobar, C. M. (1987, May). Costs and benefits of early intervention: Implications of recent legislation. Paper presented at Focusing on the Future: Linking Research, Policy, and Practice in Early Intervention, Arlington, VA. A conference sponsored by the U.S. Office of Special Education Programs and the Early Childhood Research Institutes.

120. Barnett, W. S., & Frede, E. C. (1987, April). The efficacy of public preschool programs and the relationship of program quality to efficacy. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

121. Barnett, W. S., & Frede, E. C. (1987, April). The South Carolina Preschool Study: Small sample analysis. Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

122. Casto, G. (1987, September). The importance of long-term intervention: Models and methods. Invited paper presented at the Conference on Early Intervention, Anchorage, AK.

123. Casto, G., & Barnett, W. S. (1987, March). Economic analysis of early intervention. Invited presentation at the 19th Banff International Conference on Behavioral Science, Banff, Alberta, Canada.

124. Casto, G., Millard, J., Dudley, K., Hoagland, V., & Peterson, A. (1987, May). Prospective and follow-up studies of medically fragile infants. Paper presented at the Focusing on the future: Linking Research, Policy, and Practice in Early Intervention Conference, Washington, DC.

125. Casto, G., Tingey, C., & Haring, K. (1987, May). The long-term efficacy of early intervention. Paper presented at the American Association of Mental Deficiency, Twenty-Nine Palms, CA.

126. Escobar, C. M., Barnett, W. S., & Keith, J. E. (1987, July). Use of contingent valuation methods in benefit cost analysis of preschool education. Paper presented at the Annual meeting of the Western Economic Association, Vancouver, B.C.

127. Tingey, C. (1987, April). Parental Involvement in Early Intervention: What is Known. Paper presented at the Second Annual Regional Conference, Association for Retarded Citizens, Birmingham, AL.

128. Tingey, C. (1987, September). Social skills development and retention in Down syndrome. Plennary Session, International Down Syndrome Symposium, Mexico City.

129. Tingey, C. (1987, September). Teaching social skills at home. Paper presented at the International Down Syndrome Symposiuim, Mexico City.

130. Tingey, C. (1987, April). How Children With Down Syndrome Learn Through the Years. Paper presented at the Down Syndrome Project, Baton Rouge, LA

131. Tingey, C. (1987, April). Increasing Parental Involvement. Paper presented at the National Head Start Conference, New Orleans, LA

132. Tingey, C. (1987, April). Transition Into Adulthood. Paper presented at the Down Syndrome Foundation, Ogden, UT

133. Tingey, C. (1987, May). Developmental Study of Triplets, One of Which has Down Syndrome. Paper presented at the Update--Down Syndrome Speical Interest Group, AAMD, Los Angeles, CA

134. Tingey, C. (1987, May). Early Intervention Research: The Longitudinal Studies. Paper presented at AAMD, Los Angeles, CA

135. Tingey, C. (1987, May). Assessment in Early Intervention. Paper presented at AAMD, Los Angeles, CA

136. Tingey, C. (1987, June). Early Intervention With Infants and Children Who Have Down Syndrome. Paper presented at the Utah State University, 3rd Annual Conference Research and Practice in Down Syndrome, Snowbird, UT

1986

137. Barnett, W. S. (1986, December). The economics of early intervention: A brief review. Invited presentation at Schools as Partners in Urban Development, conference sponsored by Far West Laboratory for Educational Research and Development and the ERIC Clearinghouse on Urban Education, Los Angeles, CA.

138. Barnett, W. S. (1986, September). El proyecto de investigacion longitudinal "Perry Preschool Project," de la Fundacion de Investigaciones Educationales High/Scope. Invited paper presented at the Instituto Nacional del Nino y la Familia, Quito, Ecuador.

139. Barnett, W. S. (1986, September). Investment in programs for young children as a strategy for national development. Invited paper presented at the Instituto Nacional del Nino y la Familia, Quito, Ecuador.

140. Barnett, W. S. (1986, June). Economics of early childhood care and education. Seminar conducted at the Westfalische Wilhelms-Universitat Munster, Federal Republic of Germany.

141. Barnett, W. S. (1986, April). A cross-cultural investigation of economics and mental retardation. Paper presented at Society for the Study of Chronic Illness, Impairment and Disability Annual Meeting, Reno, NV.

142. Barnett, W. S. (1986, April). Economic evaluation of early intervention programs. Paper presented at the American Educational Research Association Conference, San Francisco, CA.

143. Casto, G. (1986, June). Long-term effectiveness of early intervention in Down Syndrome. Presented at the Second Annual Research and Practice in Down Syndrome, Logan, UT.

144. Casto, G. (1986, May). The efficacy of early intervention: Separating fact from fiction. Paper presented to the Gulf Coast Early Intervention Conference, Delta Shores, AL.

145. Casto, G. (1986, May). The efficacy of early intervention with Down syndrome preschoolers. Paper presented to the American Association of Mental Deficiency, Denver, CO.

146. Casto, G. (1986, May). What we know about early intervention efficacy. Keynote speech at the Annual Conference on Early Intervention, Billings, MT.

147. Casto, G., & Barnett, W. S. (1986, October). The efficacy of early intervention: What we know, what we don't know. Paper presented at the National Early Childhood Conference on Children with Special Needs, Louisville, KY.

148. Scruggs, T. E., & Mastropieri, M. A. (1986, May). Early intervention for social withdrawal: A quantitative synthesis of single subject research. Paper presented at the annual meeting of the Association for Behavior Analysis, Milwaukee, WI.

149. Casto, G. (1986, April). The relationship of program intensity and duration to the efficacy of early intervention. Paper presented to the Council for Exceptional Children annual conference, New Orleans, LA.

150. Tingey, C. (1986, May). Developmental Study of Triplets, One of Which has Down Syndrome. Paper presented at the Down Syndrome Special Interest Group, AAMD, Denver, CO.

151. Tingey, C. (1986, May). Early Intervention Research Institute Parental Involvemlent and Longitudinal Studies. Paper presented at AAMD, Denver, CO

152. Tingey, C. (1986, September). Cutting the Umbilical Cord--Parental Perspectives. The Person With Down Syndrome: Transition From Adolescence to Adulthood. Paper Presented at New Paltz, NY

153. Tingey, C. (1986, April). Adaptations in learning for individuals with down syndrome. Paper presented at the Utah Foundation for Down Syndrome, Ogden, UT.

154. Tingey, C. (1986, October). Social Skills to Avoid Isolation. Plenary Session, National Down Syndrome Congress Conference, Kansas

155. Tingey, C. (1986, April). Adaptations in learning for children with Down Syndrome from birth to adult. Presented at the Utah Down Syndrome Foundation Conference, Ogden, UT.

1985

156. Barnett, W. S., (1985, October). The economics of early childhood education: The Perry Preschool study and its implications. Paper presented at the Utah Association for the Education of Young Children Fall Conference, Salt Lake City, UT.

157. Barnett, W. S. (1985, October). The economics of early intervention: A critical review of the literature. Paper presented at the Division for Early Childhood (CEC) Conference, Denver, CO.

158. Barnett, W. S. (1985, August). Early childhood intervention for handicapped children: Efficacy, economics, and the need for evaluation of state programs. Presentation to the Nebraska Select Committee on Services to Children Under Age 5 and their Families.

159. Barnett, W. S. (1985, June). Public policy and the Perry Preschool Study. Paper presented at the Utah Early Childhood Interinstitutional Conference, Salt Lake City, UT.

160. Casto, G. (1985, October). Children with special needs. Paper presented to the National Early Childhood Conference on Children with Special Needs, Denver, CO.

161. Casto, G. (1985, October). Early intervention--is it better? Paper presented at the Rocky Mountain Educational Research Association conference, University of New Mexico, Las Cruces, NM.

162. Casto, G. (1985, October). Efficacy of early intervention. Presented to the National Institute of Child Health and Human Development Conference on Behavioral Intervention with High Risk Infants, Bethesda, MD.

163. Casto, G. (1985, October). Implications for public policy of the Early Intervention Research Institute's meta-analysis. Paper presented to the National Early Childhood Conference on Children with Special Needs, Denver, CO.

164. Casto, G. (1985, October). Policy implications of early intervention efficacy research. Paper presented to the Nebraska Council for Exceptional Children conference, Grand Island, NB.

165. Casto, G. (1985, October). Using research to affect policy decisions on early childhood special education. Presented to the Utah State Board of Education, Salt Lake City, UT.

166. Pezzino, J. (1985, October). Efficacy, cost and policy implications of early intervention research with special needs children. Paper presented to the New Mexico State Council for Exceptional Children conference, Santa Fe, NM.

167. Tingey, C. (1985, November). Sisters and brothers of individuals with Down Syndrome. Paper presented to the National Down Syndrome conference, Anaheim, CA.

168. Tingey, C. (1985, April). Socialization and maturation. Paper presented at the Down Syndrome State-of-the-Art Conference, Boston, MA.

169. White, K. (1985, November). How research findings should affect policy decisions in early childhood special education. Paper presented to the U.S. Congressional Staff, Washington, DC.

170. White, K. R., & Mott, S. (1985, July). Pilot programs for comprehensive services to handicapped children ages birth to three years. Paper presented to the Illinois State Board of Education and Governor's Planning Council on Developmental Disabilities, Writer's Workshop for Pilot Programs for Handicapped Children up to Three Years of Age, Springfield, IL.

171. Barnett, W. S. (1985, June). The Perry Preschool Study: Implications for policy and practice. Paper presented at the Utah Interinstitutional Tenth Annual Early Childhood Conference, Salt Lake City, UT.

172. Casto, G. (1985, June). Early intervention programs for Down syndrome. Paper presented at the Conference on Research and Practice in Down Syndrome, Logan, UT.

173. Casto, G. (1985, June). Plasticity and the handicapped child. Paper presented at the Malleability of Children Conference Agenda, Rougemont, NC.

174. Frede, E. (1985, June). Using systematic observation as a teacher-training tool. Paper presented at the Utah Interinstitutional Tenth Annual Early Childhood Conference, Salt Lake City, UT.

175. Kutz-Sivill, S., & Peterson, A. (1985, June). Sensory-motor component of early intervention. Paper presented at the Conference on Research and Practice in Down Syndrome, Logan, UT.

176. Barnett, W. S. (1985, April). The long-term effects of preschool programs: Implications for research and public policy of the Perry Preschool Program's long-term effects. Paper presented at the annual conference of the American Educational Research Association, Chicago, IL.

177. Casto, G. (1985, April). The efficacy of early intervention. Paper presented at the annual conference of the Council on Exceptional Children, Anaheim, CA.

178. Casto, G. (1985, April). The efficacy of early intervention for handicapped infants. Paper presented at the Iowa Early Intervention Conference, Cedar Rapids, IA.

179. Casto, G. (1985, April). Efficacy research with infant populations. Paper presented at the Oklahoma State Early Childhood Conference, Oklahoma City, OK.

180. Pezzino, J. (1985, April). An analysis of intervention programs of varying intensities. Paper presented at the Oklahoma State Early Childhood Conference, Oklahoma City, OK.

181. Pezzino, J. (1985, April). A cost-effectiveness comparison: Professionals versus paraprofessionals as intervenors for young handicapped children. Paper presented at the Council for Exceptional Children Annual Convention, Anaheim, CA.

182. Pezzino, J., & Barnett, W. S. (1985, April). Cost-effectiveness analysis of two programs of different intensities for handicapped preschoolers. Paper presented at the Iowa State Conference on Special Education, Cedar Rapids, IA.

183. Pezzino, J., & Lux, J. (1985, April). Cost-effectiveness analysis of two early intervention programs of different intensity. Paper presented at the Iowa State Conference on Innovative Practices in Special Education.

184. Barnett, W. S. (1985, March). Cost-effectiveness of early intervention programs for disadvantaged and handicapped children. Paper presented at the conference of the National Consortium of Early Childhood/Special Education Coordinators, Denver, CO.

185. Casto, G. (1985, March). The efficacy of early intervention with handicapped preschoolers. Paper presented at the conference of the National Consortium of Early Childhood/Special Education Coordinators, Denver, CO.

186. Mastropieri, M. A., White, K. R., & Casto, G. (1985, March). Efficacy of early intervention for the handicapped and disabled: A meta-analysis. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

187. Pezzino, J., & Lauritzen, V. (1985, March). A description of the P.I.E. parent-training curriculum. Paper presented at the Assessment and Intervention Strategies for Developmentally Disabled and Mentally Retarded Infants and Preschoolers Conference, Salt Lake City, UT.

188. Pezzino, J., & Lauritzen, V. (1985, March). A training curriculum for parents of handicapped preschoolers. Paper presented at the Utah State Conference of Strategies for Developmentally Disabled and Mentally Retarded Infants and Preschoolers, Salt Lake City, UT.

189. Casto, G. (1985, February). The efficacy of early intervention with medically at-risk infants. Paper presented at the Medical University of Charleston, Charleston, SC.

1984

190. Casto, G., Barnett, W. S., & Pezzino, J. (1984, December). Efficacy studies in early intervention. Paper presented at the Handicapped Children Early Education Program Conference, Washington, DC.

191. White, K. R. (1984, December). The efficacy of early intervention: Separating fact from folklore. Invited address presented in the Charles Stewart Mott Foundation Lecture Series, Harvard University, Cambridge, MA.

192. White, K. R. (1984, December). Guidelines for conducting early intervention efficacy research. Paper presented at the Handicapped Children Early Education Program Conference, Washington, DC.

193. Mastropieri, M. A., & Casto, G. (1984, November). Early intervention for behavior disorders: An integrative review. Paper presented at the Eighth Annual Conference on Severe Behavior Disorders, Tempe, AZ.

194. Casto, G. (1984, October). The relationship of age at start and the degree of parental involvement to intervention effectiveness. Paper presented at the Rocky Mountain Educational Research Conference, Oklahoma City, OK.

195. White, K. R. (1984, October). Evaluating early intervention programs: Conclusions from previous research. Invited address presented at Infants at Risk: A New England Institute, Portland, ME.

196. Casto, G. (1984, September). The efficacy of intervention programs for severely handicapped preschool children. Paper presented at the Midwestern Conference on Deaf-Blind, Chicago, IL.

197. Casto, G. (1984, April). A report on EIRI meta-analysis. Paper presented at the 62nd Annual Convention of the Council on Exceptional Children, Washington, DC.

198. Pezzino, J. (1984, April). Cost-effectiveness of early intervention programs. Paper presented at the 62nd Annual Convention of the Council on Exceptional Children, Washington, DC.

199. Pezzino, J. (1984, April). A critique of cost-effectiveness research. Paper presented at the 62nd Annual Convention of the Council on Exceptional Children, Washington, DC.

200. Pezzino, J., Goudie, K., & Casto, G. (1984, April). A comparison of two service delivery modes in delivering speech, occupational and physical therapy to handicapped children. Paper presented at the 62nd Annual Convention of the Council on Exceptional Children, Washington, DC.

201. White, K. R. (1984, April). Applications of meta-analysis to special education: Efficacy of early intervention with handicapped and at-risk children. Paper presented at the 62nd Annual Convention of the Council on Exceptional Children, Washington, DC.

202. White, K. R. (1984, April). Efficacy of early intervention. Presentation given to the annual conference of the Utah Chapter of the American Association on Mental Deficiency, Provo, UT.

203. White, K. R. (1984, February). Contributions of research to the development and implementation of early intervention programs. Paper presented at the meeting of the National Consortium of State Agency Preschool Coordinators, Denver, CO.

1983

204. White, K. R. (1983, December). Conducting efficacy research with early intervention programs. Paper presented at a conference of the National Center for Clinical Infant Programs, Washington, DC.

205. White, K. R., & Casto, G. (1983, December). A meta-analysis of the efficacy of early intervention with handicapped and at-risk children. Paper presented at the annual meeting of the Division of Early Childhood Handicapped Children Early Education Program, Washington, DC.

206. White, K. R., Casto, G., & Mastropieri, M. A. (1983, December). Meta-analysis: Early intervention research literature and HCEEP validated projects. Paper presented at the Directors' Conference of the Division of Early Childhood, Washington, DC.

207. White, K. R., & Watkins, S. (1983, December). Longitudinal effects of various types of early intervention with hearing impaired children. Paper presented at the annual meeting of the Division of Early Childhood Handicapped Children Early Education Program, Washington, DC.

208. Casto, G., & Casto, Y. (1983, April). Intervening with high risk infants. Paper presented at the Fourth Annual Montana Symposium on Early Education of the Exceptional Child, Billings, MT.

209. Casto, G., & Clarkson, D. (1983, April). Selecting outcome measures in early intervention. Paper presented at the Fourth Annual Montana Symposium on Early Education of the Exceptional Child, Billings, MT.

210. Casto, G., Shearer, D. E., Cavaleri, T. (1983, April). Critical issues in early intervention: A view from the field. Paper presented at the Rocky Mountain Psychological Association Conference, Snowbird, UT.

211. Pezzino, J., & Taylor, C. (1983, April). A critical review: Cost-effectiveness analysis in human service research. Paper presented at the Rocky Mountain Psychological Association Conference, Snowbird, UT.

212. Casto, G., & Shearer, D. (1983, March). Previous reviewers' conclusions about the effectiveness of early intervention. Paper presented at the Montana Conference for Severely Handicapped, Billings, MT.

213. Shearer, D. (1983, February). The Early Intervention Research Institute. Presentation at the Research in Action II Conference, Lubbock, TX.

DISSEMINATION ACTIVITIES

During its five-year history, the institute received 1,331 letters of inquiry. These inquiries came from all 50 states and the following countries:

- Argentina
- Australia
- Belgium
- Brazil
- Canada
- China
- Costa Rico
- Czechoslovakia
- Dominican Republic
- East Germany
- England
- Finland
- France
- Hungary
- India
- Ireland
- Israel
- Mexico
- Netherlands
- Nepal
- Nigeria
- Nova Scotia
- Spain
- Sweden
- West Germany
- West Indies
- Yugoslavia

As a result of the inquiries received, 3,054 products/documents were disseminated. The EIRI mailing lists included the mailing lists from:

- Developmental Center for Handicapped Persons
- EIRI Field Reviewers
- Handicapped Children's Early Education Program Inventory
- State Directors of Special Education
- MCH and NICHD Funded Infant and Preschool Programs
- EIRI Inquiry File
- University Affiliated Facility Directory

The major audiences which received EIRI products included:

- Early intervention personnel in 50 states and 27 countries
- University departments of education, psychology, and child development
- State Directors of Special Education
- State Directors of Early Childhood
- Early Intervention Researchers
- State Legislatures

- U.S. Congress
- National Parent Organizations
- Neonatal Intensive Care Unit Directors
- State Departments of Health
- State Family Service Agencies

In addition, a national dissemination conference in Washington, D.C., drew 400 registered participants.

TRAINING

GRADUATE/RESEARCH ASSISTANTS

An integral part of the EIRI activity was to provide advanced research training for research and graduate assistants from several disciplines. Some of this training was offered at no cost to EIRI because of students' participation in existing DCHP interdisciplinary training programs. The research training consisted of formal course work in developmental disabilities, inservice training in procedures for specific research (meta-analysis, longitudinal research, cost-effectiveness analysis), research seminars on the studies conducted, actual data collection, analysis and dissemination, individual tutorials with senior research faculty, the planning and implementation of training sessions, and meeting specific research competency requirements.

Recruitment of Graduate and Research Assistants

The majority of assistants came from five university areas: Special Education, Psychology, Communicative Disorders, Social Work, and Family and Human Development. Research assistants, as a term employed in this report, will refer to both staff research positions and graduate students.

Training Plans

EIRI research assistants had two training modes at their disposal. First, interdisciplinary training is a primary mission of the DCHP as a UAF, and, therefore, all students who work on DCHP projects are enrolled for 9 credits of interdisciplinary course work. All graduate students meet with the DCHP director of interdisciplinary training to construct interdisciplinary training plans (ITPs) that outlined specific course work and internship experiences (other than the project to which they were assigned). Second, prior to any training, assistants met with EIRI professional staff to define tasks assigned within the research areas designated in this project. Assistants used this information to formulate goals and objectives while proposing strategies for meeting THEIR ITP objectives. Two training modes are described below.

DCHP interdisciplinary training. As noted above, all graduate students who are employed by the DCHP formulated an interdisciplinary training plan. Many opportunities for other training or internships were provides as a result of EIRI's association with various disciplines. For instance, a student working on meta-analysis could design a 1-3 credit practicum in learning about the Portage model or in functioning as a member of an interdisciplinary assessment team. Such a practicum would be listed in the person's ITP and supervised by EIRI or DCHP senior staff. EIRI also extended practica opportunities to students outside the project staff to learn specific procedures for data collection and analysis (e.g., through the workshops on meta-analysis and cost-effectiveness).

EIRI project-related training. Each aspect of the three research thrusts and major project activities (dissemination, evaluation, and performance management system), provided research assistants with training opportunities. Formal training sessions on meta-analysis and cost-effectiveness were made available to all staff. Activities of a particular research thrust gave students experience in data collection and analysis, reporting and disseminating findings, developing materials; and conducting training. Many other secondary skills were also developed by participation (e.g., working as a member of an interdisciplinary team, constructing questionnaires, planning the logistics and content for advisory committee meetings, dealing with political problems in the field, interviewing, programming, and generating personal management reports on computer).

Students Affiliated with EIRI

Name	Status	Dates Associated	Current Position/Title
Susan Watkins	Research Associate	1982-83	Director of Research, SKI*HI Project
Debra Cochran	Graduate Assistant	1982-83	Project Director, Nisonger Center, Columbus, Ohio
David Bush	Graduate Assistant	1982-83	Psychologist, USAF
Kay Walker	Graduate Assistant	1982-83	Director of Special Education, Fallon, Nevada
Gary Goodrich	Graduate Assistant	1982-85	Graduate Student
Dennis Clarkson	Graduate Assistant	1982-85	Director of Special Education, Browning, Montana
Larry Wilcox	Graduate Assistant	1982-84	Psychologist, California Child Clinic
Tish Cavalieri	Graduate Assistant	1982-84	Assistant Professor of Instructional Technology
Tom Mills	Graduate Assistant	1982-84	
James Pezzino	Post Doc.	1982-83	Consultant, Milwaukee, Wisconsin
Becky Richards	Graduate Assistant	1982-84	Graduate Student
Maryam Mehran	Graduate Assistant	1982-84	Graduate Student
Karen Seibel	Graduate Assistant	1983-84	Graduate Student
Colette Escobar	Graduate Assistant	1982-84	Research Associate, EIRI
Faith Fecteau	Graduate Assistant	1983-84	Special Education Teacher, Ogden, Utah
Helal Mobasher	Graduate Assistant	1983-87	Research Associate, EIRI
Duane Reeder	Graduate Assistant	1982-84	Professor of Psychology, Arizona
Margo Mastropieri	Post Doc.	1983-86	Assistant Professor of Special Education, Purdue
Janet Millard	Graduate Assistant	1986-87	Psychologist, Private Practice
Teri Wingate-Corey	Graduate Assistant	1985-87	Research Associate, EIRI
Bernard Wazlavek	Graduate Assistant	1983-87	Research Associate, EIRI
Glenn Goodwin	Graduate Assistant	1985-86	Psychologist, USAF
Laura Gaynard	Post Doc.	1985-86	Coordinator, Infant Research
Steve Cook	Graduate Assistant	1983-85	Graduate Student, Phoenix, Arizona
Vanessa Moss	Graduate Assistant	1984-87	Graduate Student
Bill Lowry	Graduate Assistant	1983-84	Graduate Student
Richard Elghammer	Graduate Assistant	1984-87	Research Associate, EIRI
Matt Taylor	Graduate Assistant	1985-87	Graduate Student
Todd Braeger	Graduate Assistant	1986-87	Graduate Student
Robert Bailey	Graduate Assistant	1985-87	Graduate Student
Bill Corey	Graduate Assistant	1986-87	Graduate Student
Carl Summers	Graduate Assistant	1986-87	Graduate Student

Student Theses/Dissertations

Richard Elghammer	-- Correlates of intraventricular hemorrhage in infants
Gary Goodrich	-- Stability of Infant IQ
Glenn Goodwin	-- A Neuropsychological Approach For Differentiating the Residual Effects of Neonatal Intraventricular Hemorrhage
Janet Millard	-- The Effect of an Early Sensorimotor Intervention Program on the Development of Infants with Perinatal Intraventricular Hemorrhage
Beverly Myette	-- The effectiveness of various approaches for remediating language impairment: A meta-analysis.
Teri Wingate-Corey	-- A Neuropsychological Follow-up of Low Birthweight Infants With and Without Neonatal Intraventricular Hemorrhage at Preschool Age
William Corey	-- The Effects of Ventriculoperetoneal Shunts on Neurodevelopmental Outcome Among Low Birthweight Infants
Jack Shamaly	-- The Effects of Teacher and Parent Behavioral Interventions Upon the Aggressive Behaviors of Preschoolers
Dennis Clarkson	-- The Usefulness of the DIAL-R as a Screening Tool to Identify Native American Preschoolers in Need of Special Services
Robin Bradley	-- A Study of Parental Involvement with Behavior Disordered and Developmentally Delayed Preschoolers

Personnel

During the five years of the institute, the following personnel were associated with the institute. They are listed below together with the years they were affiliated with EIRI.

EIRI Personnel

Title	Name	Years
Project Director	Glendon Casto, Ph.D.	1982-87
Project Co-Director	Karl White, Ph.D.	1982-87
Director, Economic Analysis	Steve Barnett, Ph.D.	1984-87
Principal Investigator	Cie Taylor, Ph.D.	1982-84
Principal Investigator	David Shearer, M.S.	1982-84
Principal Investigator	Ann Austin, Ph.D.	1982-83
Post Doctoral Fellow/Principal Investigator	James Pezzino, Ph.D.	1982-85
Post Doctoral Fellow/Principal Investigator	Margo Mastropieri, Ph.D.	1983-86
Principal Investigator	Terry Glover, Ph.D.	1983-84
Principal Investigator	John Keith, Ph.D.	1983-84
Principal Investigator	Thomas Scruggs, Ph.D.	1983-86
Principal Investigator	Karen Arnold, Ph.D.	1985-86
Post Doctoral Fellow	Laura Gaynard, Ph.D.	1985-86
Post Doctoral Fellow	Janet Millard Ph.D.	1986-87
Principal Investigator	Carol Tingey, Ph.D.	1986-87

Advisory Personnel

Name	Position	Years Served
Gene Glass, Ph.D.	Professor of Education, University of Colorado	1982-87
Henry Levin, Ph.D.	Professor of Economics & Education, Stanford University	1982-87
Craig Ramey, Ph.D.	Professor of Psychology, University of North Carolina	1982-87
Merle Karnes, Ph.D.	Professor of Education, University of Illinois	1982-87
Philip Strain, Ph.D.	Director, Pittsburgh Early Intervention Research Institute	1982-87
Rune Simeonsson, Ph.D.	Investigator, Carolina Institute for Early Intervention Research	1982-87
Sharon Hixon, M.S.	Classroom Teacher, Early Intervention	1982-87
Jackie Walker, Ph.D.	Yakima Washington, Tribal School	1982-87
Jessica Strout	Parent	1982-87
Marsha Shearer, M.S.	Classroom Teacher, Early Intervention	1982-84
Peter Fanning, Ed.D.	Director, Special Education, Colorado	1982-85
Tal Black, Ed.D.	President CEC Division of Early Childhood	1982-85
Amy Toole, M.S.	President CEC Division of Early Childhood	1985-87
Beverly Osteen, M.S.	National Association of State Directors of Special Education	1985-86
Irving Lazar, Ph.D.	Professor, Cornell University	1984-87

In addition to advisory group members, the institute utilized a group of 50 researchers, practitioners, administrators, and parents as a field advisory group. This group contributed by reviewing instruments, and responding to questionnaires. Their names follow.

Name	Affiliation	State
Nicholas Anastasiow	Hunter College	New York
Maddie Appell	St. Luke's Roosevelt Hospital	New York
Victor Baldwin	Teaching Research	Oregon
Jo Bunce	Division of Special Education Virginia Department of Education	Virginia
Sue Chappel	Van Asselt School	Washington
Jack Cole	New Mexico State University	New Mexico
Carl Dunst	Department of Human Resources Western Carolina Center	North Carolina
Rebecca Fewell	University of Washington	Washington
Pam Frakes	Tennessee Children's Services Commission	Tennessee
David Franks	University of Wisconsin - Eau Claire	Wisconsin
Corrine Garland	Williamsburg-James City County Public Schools	Virginia
Linda Gilkerson	Wheelock College	Massachusetts
Bea Gold	Children Youth & Families, Inc.	California
Michael Guralnick	Ohio State University	Ohio
Alice Hayden	University of Washington	Washington
Diane Holland	Parent	Michigan
Patricia Hutinger	Western Illinois University	Illinois
Beverly Johnson	Detroit City School District	Michigan
Merle Karnes	University of Illinois	Illinois
Robert Kibler	Peabody College - Vanderbilt University	Tennessee
Karlene Knebel	Frank Porter Graham	North Carolina
Arleen Lewis	University of Arkansas	Arkansas
Jeanette Walker-McCallum	University of Illinois	Illinois
Katie McCartan	Iowa State University	Iowa
Jeanne McCarthy	University of Arizona	Arizona
Jim McLean	University of Kansas	Kansas
Brian McNulty	Colorado Department of Education	Colorado
Kris Montgomery	Peoria 0-3 Outreach Program	Illinois
Rick Offne	University of Montana	Montana
Lucille Paden	Dept. of Public Instruction	Kansas
Nancy Peterson	University of Kansas	Kansas
Kenneth Reavis	Utah State Office of Education	Utah
Mary Tom Riley	Texas Tech University	Texas
Caryn Robbins	University of Kansas	Kansas
Pat Robinson	Director, Severely Handicapped Programs	Wyoming
Donna Rokicki	Parent	Illinois
Earl Schaefer	Frank Porter Graham Child Development Center	North Carolina
Marsha Shearer	Educational Service District No. 121	Washington
Phillip Sipos	Department of Education	Louisiana
Jessica Strout	Parent	Utah
William Swan	University of Georgia	Georgia
Paul Swatensbarg	Division of Community Rehabilitation	Idaho
Amy Toole	BOCES 2, Special Education Department	New York
Ted Tjossem	National Institute of Child Health and Human Development	Maryland
Jamie Tucker	Institute for Child & Family Studies	Texas
Ruth Turner	Dallas Independent School District	Texas
Kay Tymer	University of Southern California Santa Barbara	California
Warren Umansky	University of Georgia	Georgia
Lizbeth Vincent	University of Wisconsin	Wisconsin
Ann Rogers-Warren	Peabody/Vanderbilt	Tennessee

INSTITUTE IMPACT

The major impact of the activities of the institute may be summarized under three categories. These categories include: improving the overall quality of efficacy research, influencing early childhood legislation, and making practitioners and researchers more aware of the importance of collecting both cost and effects data in the early intervention field. The institute's impact in each of these areas will be briefly discussed in the section to follow.

Improving the Quality of Efficacy Research

One of the institute's first tasks was to conduct an integrative review of the literature related to efficacy. The overall conclusion reached by previous reviewers regarding the efficacy of early intervention was supported by the EIRI integrative review. That is, early intervention does have an immediate positive benefit across an array of handicapping conditions and across several develompental domains.

When certain variables which many previous reviewers had thought were crucial to intervention success were examined, however, a major discrepancy was found between the EIRI findings and the conclusions of previous reviews. For variables such as degree of parental involvement, the age at which intervention should start, etc., the EIRI integrative review could not find objective data to support, for example, previous reviewers' conclusions that parental involvement is crucial to intervention success or that the earlier an intervention starts, the more effective is is likely to be. EIRI researchers concluded that the data which exists does not support either of these conclusions, and that a definitive answer to either question awaits further, more tightly controlled research.

The reporting of the findings of the integrative review provoked a series of healthy debates in the literature. Some researchers questioned the findings (Strain & Smith, 1987), while others questioned the meta-analysis methodology (Dunst & Snyder, 1987).

The integrative review also produced evidence that much of the early intervention research done to date has been so marred with methodological difficulties that the findings are, in many cases, uninterpretable. In addition, the narrow array of outcome measured used in previous research (40% used IQ as a major outcome measure) inhibits our ability to determine the impact of early intervention across other important develomental domains. An important impact of the EIRI research program is that the publications done by project staff, and the final research projects conducted, have demonstrated the importance and feasibility of conducting randomized trials research in early intervention.

Such randomized trials research include random assignment to groups, verification of treatment implementation, impartial data collection and the selection of outcome measures which accurately capture the intended outcomes of the intervention program. These principles are discussed briefly below as they relate to EIRI's early intervention research.

Random Assignment

In the field of early childhood special education, the subject populations are diverse, and the potential sample sizes are usually so small that both stratification procedures and random assignment need to be addressed to ensure comparability of groups.

EIRI researchers recommend that stratification in selecting experimental and control groups begins by grouping or ranking the study subjects on what may be considered the critical or the most important variable or variables. Severity of handicapping condition or type of handicapping condition are examples of critical variables. For example, within a hearing impaired sample, the variable might be degree of hearing loss. If degree of hearing loss is used as the critical variable, then all subjects would be rank-ordered on degree of hearing loss, and randomly sorted (odd/even) into two groups (e.g., 1st into Group 1, 2nd into Group b, 3rd into Group A, etc.). If these groups were not sufficiently comparable on certain important variables (e.g., age, SES, maternal, IQ, etc.), then pairs could be switched from one group to the other (i.e., the places of members of a pair would be switched) until equality was achieved. This would hold the critical variable more or less constant. The two groups thus formed would be finally designated treatment or control at random; for example, by the flip of a fair coin.

The primary reason for using matching procedures prior to random assignment is, of course, to reduce variability within pairs and to help insure comparability of gorups. Although random assignment to groups, in and of itself, may accomplish these aims, stratification prior to assignment reduced the probability of producing groups which are not comparable.

The number of variables to use in stratifying groups is dependent on subject characteristics and sample size. As the sample size increases, it may be possible to match on more variables, but matching becomes less necessary with large samples. In general, it becomes increasingly difficult to obtain a high percentage of successful matches from a limited subject pool as the number of matching variables increases. However, matching on at least two variables with available subject pools is usually feasible. EIRI researchers have found that using at least two matching variables has the promise of reducing the variability within subject pools significantly.

Treatment Implementation

Rresearchers attempting to design experimental research may start out with an ideal design, including random assignment of subjects and clear-cut distinctions between experimental and control group in terms of the independent variable, and yet fail to find differences between the two groups at posttesting. That finding may mean that the treatment made no difference. However, frequently the treatment made a difference, but extraneous, uncontrolled variables, related to the independent variables, influenced the results. Perhaps in a more common scenario in early childhood special education, the treatment was implemented differentially or not at all.

In the EIRI meta-analysis of the early intervention efficacy research, Casto and Mastropieri (1986) found that verification of treatment implementation was one of the most neglected aspects in efficacy studies. Research reports typically included inadequate descriptions of the treatment to be offered and provided almost no data on the degree of treatment implementation. There are procedures which can be utilized to verify treatment implementation which add little to the costs of the research. EIRI has field tested, and is making available to the field, treatment verification materials.

Impartial Data Collection

"Love is blind" and examiners should be. This terse statement summarizes the necessity of using data collectors who are both uninformed and unaware of the purpose of a given research study and who are unaware of the purpose of a given research study and who are unaware of the group membership of the subjects from whom they collect data. Casto and Mastropieri (1986) found that only about 20% of the findings in their review of efficacy research came from studies where "blind" examiners were utilized. Since procedures to ensure impartial data collection are easy to implement, and in most cases add nothing to the costs of doing the research, the use of "blind" examiners should be given more emphasis in early childhood special education research.

Outcome Measures

EIRI has made several recommendations to the field in an attempt to have broader based outcome measures utilized in efficacy research. When tests are being considered, guidelines for their selection have been developed by EIRI staff and disseminated. The research team has also tested the feasibility of using questionnaires, direct observation procedures, interviews, and unobtrusive measures to document the efficacy of interventions. Institute staff have also researched and made available to the field annotated bibliographies of both child and family measures which might be used to document efficacy. EIRI staff have also disseminated the notion that the best advice for the researcher choosing tests is to begin with an operational definition of the dependent variables in the research and then carefully select instruments using the following questions:

- What is the rationale for selecting this instrument?
- Will this instrument give the best information regarding the varaibles under study?
- Is this instrument valid and reliable for the study sample?

Influencing Legislation Related to Early Childhood Special Education

When the Early Intervention Research Institute began in 1982, there was a great deal of interest in expanding the availability of early intervention programs for handicapped and at-risk children. In fact, there is some evidence that the decision to fund an institute to investigate the efficacy and cost-benefit of early intervention was largely in response to calls from both state and national legislators for more evidence supporting the need for early intervention programs. During the last five years, the legislative activity related to early intervention programs intensified, culminating in the fall of 1986 with the passage of P.L. 99-457, which provides substantial incentives and sanctions to encourage states to develop comprehensive programs of early intervention for all handicapped birth through five-year-old infants. The research and other activities of staff at the Early Intervention Research Institute have contributed substantially to legislation developed during this time period at both the state and national level.

The link between research and policy is tenuous at best. Legislators often question why researchers do not provide them with more useful data, and researchers wonder why legislators ignore important information as they develop laws and regulations. One of the difficulties with the link between research and policy, is that research has often been used as a political weapon rather than a scientific tool. As a coordinator of a large early intervention program commented to EIRI staff as he was approached about becoming involved in one of our studies, "I use research like a drunk uses a lamp pole--I use it for support, not illumination." This not uncommon approach to research has resulted in many early intervention programs which are based more on, personal prejudices, or prevailing practices, than they are on objective evidence of what works best for which group of children.

It is also clear that legislators must make decisions about whether or not to provide early intervention programs and about what type of programs to provide based on a variety of factors other than research. Constituency support, historical factors, financial conditions, and principles or moral values, all play an important role in decisions about all social service programs of which early intervention is a part. Research can assist us in determining what types of program are best for which children. But research alone should not be used to decide whether handicapped children deserve our help. Questions about whether or not to provide early intervention services can be informed, but will never be decided on the basis of research.

Our experience in analyzing and integrating the results from literally hundreds of past research studies, as well as our experience in conducting dozens of our own research studies on the efficacy and cost-efficiency of early intervention, convinces us that there are at least two major areas frequently addressed in early intervention efficacy research. The first is represented by the question, "**Should we provide early intervention?**" Even though many people have appealed to research data to answer this question, it is a question which research alone is not particularly well-suited to answer. For example, the situation for handicapped children in 1982 was very similar to the situation 20 years earlier regarding cancer. At that time, hundreds of thousands of people were suffering from the effects of cancer. We knew that people who contracted cancer almost always died from the effects of the disease. Research was not used to decide whether cancer patients needed help. Instead, massive research programs were undertaken to determine which treatments where best for which people. So far, no cure has been discovered, by survival rates have been dramatically increased. That was not done by randomly assigning cancer patients to treatment vs. no treatment groups; it was done by comparing the effects of surgery vs. diet, vs. chemotherapy, vs. radiation, etc.

The parallel between cancer research and early intervention programs leads us to the second category of research which is summarized by the question, "**What types of programs are best for which children?**" We know that children who are identified as moderately to severely handicapped during the preschool years, will continue to exhibit reduced level of functioning throughout their lives. We do not need research to decide whether such children need help. In fact, it is just as indefensible to simply do nothing for these children, as it would be to randomly assign cancer patients to treatment vs. no treatment groups. The questions that need to be addressed in early intervention research, are questions about what type of intervention, at what age, with which children, can be most cost-effective in ameliorating the effects of these handicapping positions. Research which addresses the cost-effectiveness of alternative forms of intervention which vary along dimensions of intensity, duration, comprehensiveness, age at start, and type of family involvement, are appropriate and needed. Research which addresses questions of treatment vs. absolutely no treatment, are not needed.

This empirical approach has guided the efforts of the Early Intervention Research staff over the last five years and has had a substantial impact on guiding legislation at state and national levels. EIRI staff have been called upon to provide written and verbal testimony to state and national policy makers, and have emphasized the important role that research can play in systematically examining variation in the types of services which are offered. At the same time, we have emphasized the need for early intervention and the importance of the problems that are being addressed.

The impact of the institute's activities on legislative initiatives is difficult to measure, but is indicated by activities of institute staff such as the following:

- o Service on a national task force convened by the Assistant Secretary of Education to provide recommendations about the future of early childhood special education.
- o Distribution of hundreds of articles and papers to state and national policy makers about the research on the efficacy of early intervention.
- o Invited presentations to national forums, including the National Governors' Task Force, United States Congress Research Forum, State Legislative Groups, and Professional organizations.
- o During a sabbatical leave from the institute, during which time his salary was paid by Utah State University, one of the co-directors of EIRI served on the staff of the United States Senate Subcommittee on the Handicapped and provided significant input into the drafting of the Senate version of what eventually became P.L. 99-457.

The substantive impact of the EIRI activities on legislation can be summarized in four areas. First, EIRI research demonstrates that early intervention must be viewed as a multi-faceted undertaking, and that the term is often used in such a wide variety of situations with such dramatically groups of children, that the results can be, and sometimes are, misleading. Children with whom intervention programs are conducted range from low birthweight infants with no discernable delays, to profoundly retarded deaf/blind infants and preschoolers who heretofore have spent their lives in custodial institutions. Interventions range from a few seconds of vestibular stimulation for children with cerebral palsy, to eight or more hours per day, five days per week, from birth on of interdisciplinary educational, psychological, and medical intervention for more profoundly retarded children. The annual cost per child of early intervention programs ranges from a few hundred dollars to tens of thousands of dollars. The work of the institute in integrating past literature has emphasized the importance of not "lumping" early intervention together as if it were a unidimensional construct that is equally applied in all situations.

The second major result of the institute's work has been to encourage systematic variation as we move to expand more comprehensive and universally available early intervention programs for handicapped children. Past research makes it clear that we do not yet know which types of programs are most cost-efficient for which children. For example, even though everybody talks about the importance of parent involvement in early intervention programs, EIRI's research on parent involvement has demonstrated that parent involvement can range from the mandatory type of involvement required in IEP meetings, to programs where parents are responsible for planning, implementing, and monitoring most of the intervention. Parent involvement varies from those programs where parents are used to provide therapy to their children, to programs which are designed to provide support to families themselves. The importance of implementing programs in such a way that the effects of these very different dimensions of parent involvement can be examined and understood cannot be over emphasized.

Third, the institutes activities have pointed out the need for rigorous and on-going evaluation of the type of systematic variation referred to above. Simply varying the type of programs which are implemented will not lead to improvement unless it is accompanied by objective, carefully documented information about the effects of different types of early intervention programs. Too much of existing early intervention research has been conducted in such a way that the information was not useful for deciding which programs were most cost-effective. The work of the institute is having a substantial effect as state legislators move to incorporate systematic plans for evaluation in their programs for expanding early intervention programs on P.L. 99-457.

Finally, the institute's work demonstrates the importance of incorporating information about both costs and effects in early intervention research, particularly for making policy decisions. It is critical to have the information necessary to decide which program is most cost-efficient for which group of children, in addition to knowing what effects can be attributed to a program.

The research of the last five years conducted by EIRI has raised a great many questions about what types of programs to provide to which children, under which circumstances. However, a careful consideration of all of the available evidence still leads to a very straightforward answer to the very complex question of whether we should provide early intervention to handicapped children. That answer is definitely yes. It is a response based on not only the available research, which is somewhat limited, but also on the professional judgment of thousands of researchers, practitioners, and administrators, as well as our values as to what obligation society has towards handicapped children. Even though as scientists, we cannot yet offer conclusive proof of the immediate and long-term benefits of early intervention for handicapped children, the consequences of not intervening are too great, and everything we know about human development suggests that there is no benefit to waiting. The real value of research lies in better understanding of which early intervention programs are most cost-efficient, rather than in deciding whether or not to offer programs.

Importance of Collecting Costs and Effects Data

Economic evaluation is relatively new in early intervention. Prior to the efforts of economic researchers at the Early Intervention Research Institute, few economic evaluations have been conducted in studies of young handicapped children. Most of the ones which have been conducted were methodologically problematic (Barnett & Escobar, in press). For example, of the 21 studies identified by Barnett and Escobar in their reviews, only four found that present evidence regarding the economics of early intervention for at-risk and handicapped children based on data collected substantailly beyond the intervention period (Barnett, 1985a; b; Seitz, Rosenbaum, & Apfel, 1985; Skeels, 1966; Weber, Foster, & Weikart, 1978; Weiss, 1981).

The rationale for EIRI's work in this area is that the evaluation of both the efficacy and costs of an early intervention program are necessary to consider the value of the program. The most effective program may not be the most "cost-effective." Likewise, the least expensive program may not be the most "cost-effective." Economic analysis allows us to evaluate costs and effects simultaneously, providing a more complete set of information for selecting the "best" program. Moreover, failure to account for the economic consequences of an intervention may not simply result in an inefficient program. It may actually lead to the failure of that program.

Economic analyses in early intervention research have generally been of two types: cost-benefit and cost-effectiveness analysis. Those two procedures differ primarily in their treatment of outcomes and in the types of studies for which they are most useful.

Cost-effectiveness analysis is a way to study the relationship between program outcomes and program costs. It is most useful in considering alternative strategies to address the same problem. Programs can be compared on how much they accomplish with each dollar invested in them.

Cost-benefit analysis is a way to compare the dollar value of a program's advantages (benefits) to the dollar value of its disadvantages (costs). It requires a comprehensive measurement of program effects and the estimation of the economic value of those effects. Often cost-benefit analysis is only partially accomplished, with the researchers recognizing that some important program effects could not adequately be represented in terms of dollars.

Cost-effectiveness analysis procedures. The comparison of costs and effects differs between cost-effectiveness (CE) and cost-benefit analysis (CBA). CE analysis uses a series of matrices that display the costs and effects of each intervention. A hypothetical cost-effectiveness matrix is given in Table 70. Such a matrix displays

The rationale for EIRI's work in this area is that the evaluation of both the efficacy and costs of an early intervention program are necessary to consider the value of the program. The most effective program may not be the most "cost-effective." Likewise, the least expensive program may not be the most "cost-effective." Economic analysis allows us to evaluate costs and effects simultaneously, providing a more complete set of information for selecting the "best" program. Moreover, failure to account for the economic consequences of an intervention may not simply result in an inefficient program. It may actually lead to the failure of that program.

Ec nomic analyses in early intervention research have generally been of two types: cost-benefit and cost-effectiveness analysis. Those two procedures differ primarily in their treatment of outcomes and in the types of stud.es for which they are most useful.

Cost-effectiveness analysis is a way to study the relationship between program outcomes and program costs. It is most useful in considering alternative strategies to address the same pr Programs can be compared on how much they accomplish with each dollar invest them.

Cost-benefit analysis is a way to compare the dollar value of a program's advantages (benefits) to the dollar value of its disadvantages (costs). It requires a comprehensive measurement of program effects and the estimation of the economic value of those effects. Often cost-benefit analysis is only partially accomplished, with the researchers recognizing that some important program effects could not adequately be represented in terms of dollars.

Cost-effectiveness analysis procedures. The comparison of costs and effects differs between cost-effectiveness (CE) and cost-benefit analysis (CBA). CE analysis uses a series of matrices that display the costs and effects of each intervention. A hypothetical cost-effectiveness matrix is given in Table . Such a matrix displays

Table 70

Hypothetical CE Matrix for DQ, Motor Skills, and Positive Responses Across Three Interventions (A, B, C)

	Cost Per Child			Effects			
	Total	Parents	Project	DQ[a]	Skills[b]	Responses[c]	Attitude[d]
A	1,050	550	500	3	12	15	4
B	1,750	1,400	350	9	5	4	5
C	1,800	600	1,300	0	20	17	9

[a]Mean gain in DQ
[b]Mean number of skills mastered
[c]Mean number of positive responses in one trial
[d]Mean attitude-toward-child score on a 10-point scale where 10 is positive and 1 is negative

the relative strengths and weaknesses of each of the interventions in an easily read format. Program C, for example, is associated with more motor skills and positive responses than are programs A or B. However, Program C has higher costs and lower developmental quotients (DQs). The matrix approach allows several different comparisons to be made on program costs and effects. For example, costs can be separated by the groups bearing the expense of the resource, or effects can be displayed according to the type of handicap, severity of handicap, or age served.

This analysis and display procedure is used instead of the direct computation of simple cost-effectiveness ratios for several reasons. First, it may be inappropriate for the evaluators to decide which cost breakdowns and effects are the most important. For instance, some persons may value parent satisfaction more than DQ while others may have the opposite priority. In another instance, a CE comparison disregarding parent time may be desired (if one wants to know what is feasible based on public school resources, for example). The ultimate cost-effectiveness comparisons must be left to the decision-making body Second, this format displays the distribution of the intervention costs and effects. For example, in Table 70, the parents in Program B bear more of the costs than do parents in Program A or C. However, the parents in Program C benefit more from better attitudes than parents in Program A or B. This disaggregation provides decision-makers with valuable information about political and social impacts of the program and potential disincentives or incentives to parent participation. Third, the matrices are easily comprehended by readers without an economics background. Thus, the data are available to a wide audience, increasing the usefulness of the cost-effectiveness data. Fourth, cost-effectiveness ratios do not provide a reliable ranking of programs in terms of economic efficiency (Barnett, 1986).

Cost-benefit Analysis Procedures. Cost-benefit analysis (CBA) is most important when the economic implications of outcomes are readily estimated. For example, a program may reduce special education costs or need for therapy, and the economic benefits of this to society can at least be roughly estimated. Because the process of estimating the dollar value of intervention outcomes is almost always incomplete, it yields a conservative estimate of the net economic return to society. However, such analyses can be accomplished with early intervention studies to a much greater extent than non-economists often suppose, as demonstrated by the economic evaluation of the Perry Preschool Project (Barnett, 1985a; b). For the institute's analyses, three types of measures can be used to quantify the benefits of early intervention.

> Savings in costs of care and education. One measure of benefits is the cost savings which are generated by increasing the capacities of handicapped preschoolers, or improving the efficiency of the service delivery system. These cost savings may derive from: organizational, procedural, or staff changes that reduce intervention costs; a reduction in the intensity or duration of later special services; or an intervention that provides a better transition to later services and so increases productivity or reduces cost. For example, the Perry Preschool Study analyzed cost savings in education and social services (Barnett, 1985a; Berrueta-Clement et al., 1984). Significant cost differences were observed as early as two and three years after the intervention. Seitz, Rosenbaum, and Apfel (1985) found similar kinds of educational savings from an intervention program that focused on families and began at birth.

Cost savings to households. Families with handicapped children have substantially higher child-related expenses of time and money than do families without handicapped children. This applies to many ordinary activities as well as to special activities not required for non-handicapped childr. One way that we can measure cost savings is to compare time use and out-of-pocket expenditures for sample families participating in interventions.

Willingness-to-pay by households. The most complete benefit estimation procedures estimate the value of an intervention program and its effects to families beyond the cost savings discussed above. The techniques used to produce more complete estimates of benefits are generally classed as either (1) "hedonic" approaches or (2) direct measures of willingness-to-pay. The hedonic approach involves the estimation of a "household production function" based on expenditures of money and time by household members on various goods and services (Lancaster, 1966), or the identification of differing prices or wages accepted in order to participate in the activity. Estimation of a household production function can involve difficult theoretical and empirical problems and requires relatively large amounts of detailed data collection by household (Barnett, 1977; Barnett, 1983; Muellbauer, 1974; Pollack & Wachter, 1975).

The second approach to valuation, direct elicitation of willingness-to-pay through "bidding games," might also be successfully applied to early intervention programs and their effects. However, strategic and other biases which are often suspected in hypothetical responses may be a problem. Also, it is sometimes difficult to elicit responses from individuals in cases where very detailed descriptions of the "game" must be used; this would be the case for valuing specific treatment variations in intervention components. The economic analysis staff have developed possible solutions to these problems, however, and have had some success in using this approach. For example, Escobar, Barnett, and Keith (1987) were able to obtain reasonable estimates of parents' valuation of a preschool program for handicapped children. These estimates were highly consistent with predictions based upon economic theory. We have been experimenting with the form of survey used to collect data in several "pilot" sites.

References

Abidin, R. R. (1983). Parenting stress index--Manual. Charlottesville, VA: Pediatric Psychology Press.

Ahmann, P. A., Lazzara, A., Dykes, F. D., Brann, A. W., & Schwartz, J. F. (1980). Intraventricular hemorrhage in the high-risk preterm infant: Incidence and outcome. Annals of Neurology, 7, 118-124.

Alonzo, L., Moor, P. M., & Raynor, S. (1979). Mainstreaming preschoolers: Children with visual handicaps. (DHEW Publications No. OHDS 78-3112). Washington, DC: Head Start Bureau.

Andrews, S. R., Blumenthal, J. B., Johnson, D. L., Kahn, A. J., Ferguson, C. J., Lasater, T. M., Malone, P. D., & Wallace, D. B. (1982). The skills ofmothering: A study of parent child development centers. Monographs of the Society for Research in Child Development, 47, Serial No. 198(6).

Arnold, K., Myette, B., & Casto, G. (1985). Relationships of language intervention efficacy to selected pretreatment subject characteristics in mentally retarded and language disordered children: A meta-analysis. Unpublished manuscript, Utah State University, Logan.

Bailey, E., & Bricker, D. (1984). The efficacy of early intervention for severely handicapped infants and young children. Topics in Early Childhood Special Education, 4(3), 30-51.

Balow, I. H., & Brill, R. G. (1975). An evaluation of reading and academic achievement levels of 16 graduating classes of the California School for the Deaf, Riverside. Volta Review, 77, 255=266.

Barnett, W. S. (1985a). Benefit-cost analysis of the Perry Preschool Program and its long-term effects. Educational Evaluation and Policy Analysis, 7, 333-342.

Barnett, W. S. (1985b). The Perry Preschool Program and its long-term effects: A benefit cost analysis. High/Scope Early Childhood Policy Papers (No. 2). Ypsilanti, MI: High/Scope.

Barnett, W. S. (1986a). Definition and classification of mental retardation: A reply to Zigler, Balla, and Hodapp. American Journal of Mental Deficiency, 91, 111-116.

Barnett, W. S. (1986b). Methodological issues in economic evaluation of early intervention programs. Early Childhood Research Quarterly, 1, 249-268.

Barnett, W. S., & Escobar, C. M. (1987). Economic evaluation of early education intervention. Unpublished manuscript, Utah State University, Logan, UT.

Barrera, M. E., Rosenbaum, P. L., & Cunningham, C. E. (1986). Early home intervention with low birth weight infants and their parents. Child Development, 57, 20-33.

Bates, E. (1976). Language and context: The acquisition of pragmatics. New York: Academic Press.

Bedger, J. E. (1974). Cost analysis in day care and head start. Child Welfare, 3, 514-523.

Bennett, C. W. (1982). Language facilitation techniques for parents and teachers. Paper presented at the annual meeting of the American Speech and Hearing Association, Toronto.

Bereiter, C., & Englemann, S. (1966). Teaching disadvantaged children in the preschool. Englewood Cliffs, NJ: Prentice- Hall.

Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). Changed lives: The effects of the Perry Preschool Program on youths through age 19. Ypsilanti, MI: High/Scope.

Black, T., & Hutinger, D. (1981). Making it work in rural communities. HCEEP Rural Network.

Blackman, J. A. (Ed.) (1981). Medical aspects of developmental disabilities in children birth to three: A resource for special service provers in the educational setting.

Bloom, L., & Lahey, M. (1978). Language development and language disorders. New York: Wiley & Sons.

Bowerman, R. A., Donn, S. M., Silver, T. M., & Jaffe, M. H. (1984). Natural history of neonatal periventricular/ intraventricular hemorrhage and its complications: Sonographic observations. American Journal of Roentgenography, 143, 1041-1052.

Box, G. E. P., & Jenkins, G. M. (1976). Time series analysis: Forecasting and control. San Francisco: Holden-Day.

Boyd, R. D. (1974). The Boyd Developmental Progress Scale. San Bernadino, CA: Inland Counties Regional Center.

Broen, P. (1972). The verbal environment of the language learning child. ASHA Monographs, 17. Washington, DC: American Speech and Hearing Association.

Bronfenbrenner, U. (1974). Is early intervention effective? Teachers College Record, 76, 279-303.

Bush, D. W., & White, K. R. (April, 1983). The efficacy of early intervention: What can be learned from previous reviews of the literature? Paper presented at the annual meeting of the Rocky Mountain Psychological Association, Snow Bird, UT.

California State Department of Education (1982). The efficacy and cost-effectiveness of early education for handicapped infants and preschool children. Sacramento: Author.

Campbell, D. T., & Stanley, J. (1966). Experimental and quasi-experimental designs in research. Chicago, IL: Rand McNally.

Caputo, D. V., & Mandell, W. (1970). Consequences of low birthweight. Developmental Psychology, 3, 363-383.

Carey, W. B., McDevitt, S. C. (1978). Revision of the Infant Temperament Questionnaire. Pediatrics, 61, 735-739.

Carlberg, C. G., Johnson, D. W., Johnson, R., Maruyama, G., Kavale, K., Kulik, C., Kulik, J. A., Lysakowski, R. S., Pflaum, S. W., & Walberg, H. J. (1984). Meta-analysis in education: A reply to Slavin. Educational Researcher, 13, 16-23.

Carlberg, C., & Kavale, K. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. Journal of Special Education, 14, 295-309.

Carran, N. (1983). Early childhood continuum of assessment, programming, evaluation, and resources. Sioux City, IA: Western Hills Education Agency #12.

Casto, G. (Ed.) (1979). Curriculum and monitoring system: Motor program. New York: Walker Publishing.

Casto, G. (1980). Multi-Agency Project for Preschoolers (Validation Document). Washington, DC: National Diffusion Network.

Casto, G., & Mastropieri, M. A. (1986). The efficacy of early intervention programs for handicapped children: A meta-analysis. Exceptional Children, 52, 417-424.

Casto, G., & Tolfa, D. (1981). Multi-Agency Project for Preschoolers. In T. Black & P. Hutinger (Eds.), Cost-effective delivery strategies in rural areas: Programs for young handicapped children, (Vol. 1, pp. 11-17). HCEEP Rural Network Monograph.

Casto, G., White, K. R., & Taylor, C. (1983). An Early Intervention Research Institute: Efficacy and cost studies in early intervention. Journal of the Division for Early Childhood, 7, 5-17.

Catto-Smith, A. G., Yu, V. Y. H., Bajuk, B., Orgill, A. A., & Astbury, J. (1985). Effect of neonatal periventricular hemorrhage on neurodevelopmental outcome. Archives of Disease in Childhood, 60, 8-11.

Center, B. A., Skiba, R. J., & Casey, A. (1985-1986). A methodology for the quantitative synthesis of intra-subject design research. Journal of Special Education, 19, 387-400.

Clark, C., & Watkins, S. (1978). Programming for hearing impaired infants through amplification and home intervention. Logan, UT: Utah State University Printing.

Clark, T. C. (1979). Language development through home intervention for infant hearing impaired children. Unpublished doctoral dissertation, University of North Carolina.

Clark T. C. (1981). Efficacy of early home intervention for promoting language development in hearing impaired infants and very young children. Paper presented at HCEEP-DEC Conference, Washington, DC.

Clarke, A. M., & Clarke, A. D. B. (1976). Early experience: Myth and evidence. London, England: Open Books.

Cohen, J. (1977). Statistical power analysis for the behavioral sciences. New York: Academic Press.

Comptroller General (1979). Early childhood and family development programs improve the quality of life for low income families. Report to the Congress.

Conley, R. W. (1973). The economics of mental retardation. Baltimore, MD: Johns Hopkins University Press.

Cook, T. D., & Campbell, D. T. (1979). Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

Cook, T. D., & Leviton, L. C. (1980). Reviewing the literature: A comparison of traditional methods with meta-analysis. Journal of Personality, 4, 449-472.

Cook, S., Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1985-1986). Handicapped students as tutors. Journal of Special Education, 19, 483-492.

Covert, R. W., & Umila, B. J. (1982). National SKI*HI Outreach Site Summary Report. Unpublished manuscript, University of Virginia.

Craig, W. N. (1964). Effects of preschool training on the development of reading and lipreading skills of deaf children. American Annals of the Deaff, 199, 280-296,

Cronbach, L. J. (1981). Designing evaluations of educational and social programs. San Francisco: Josey-Bass.

Cross, T. (1978). Mothers' speech and its association with rate of syntactic acquisition in young children. In N. Waterson & C. Snow (Eds.), The development of communication: Social and pragmatic factors in language acquisition. New York: John Wiley & Sons.

Datta, L. (1977). The external implications of an internal reviewof effectiveness: The DHEW Education Division's Joint Dissemination Review Panel. In Education Division's Joint Dissemination Review Panel: Three Papers. New York: EPIE Institute. (ERIC Document Reproduction SErvice No. ED 156 217).

DeProspero, A., & Cohen, S. (1979). Inconsistent visual analysis of variance model for the intrasubject replication design. Journal of Applied Behavior Analysis, 12, 573-570.

Duncan, D. B. (1955). Multiple range and multiple F tests. Biometrics, 11, 1-42.

Dunn, L. M., & Dunn, L. M. (1981). Peabody Picture Vocabulary Test: Manual, forms L and M. Circle Pines, MN: American Guidance Service.

Dunst, C. J., & Rheingrover, R. M. (1981). An analysis of the efficacy of infant intervention programs with organically handicapped children. Evaluation and Program Planning, 4, 287-323.

Edmiaston, R. K., & Mowder, B. A. (1985). Early intervention for handicapped children: Efficacy issues and data for school psychologists. Psychology in the Schools, 22, 171-177.

Educational Research Service (1980). Class-size research: A critique of recent meta-analysis. Kappan, 239-241.

Escalona, S. K. (1982). Babies at double hazard: Early development of infants at biologic and social risk. Pediatrics, 70, 670-675.

Ewing, A. W. G., & Ewing, E. C. (1964). Teaching deaf children to talk. Washington, DC: Volta Bureau.

Eysenck, J. J. (1978). An exercise in mega-silliness. American Psychologist, 33, 517.

Fang, W. L. (1981). The Joint Dissemination Review Panel: Can appoved submittals be distinguished from rejected ones on the basis of presented evidence of effectiveness related to cognitive objectives? Doctoral dissertation, University of Virginia.

Fenichel, G. M., Webster, D. L., & Wong, W. K. T. (1984). Intracranial hemorrhage in the term newborn. Archives of Neurology, 41, 30-34.

Field, T. (1978). The perinatal anxieties and attitudes scale. Unpublished manuscript, University of Miami Medical School, FL.

Field, T. (1981). Early development of the preterm offspring of teenage mothers. In K. Scott, T. Field, & E. Robinson (Eds.), Teenage parents and their offspring (pp. 145-175). New York: Grune & Stratton.

Field, T. M., Widmayer, S. M., Stringer, S., & Ignatoff, E. (1980). Teenage, lower-class, black mothers and their preterm infants: An intervention and developmental follow-up. Child Development, 51, 426-436.

Filskov, S. B., & Boll, T. J. (Eds.) (1981). Handbook of clinical neuropsychology. New York: Wiley & Sons.

Fluharty, N. (1974). Fluharty Screening Test for Preschool Children. Journal of Speech and Hearing Disorders, 39, 75-88.

Frakes, P. J. (1981). Early lifestyle program. In T. Black & P. Hutinger (Eds.), Making it work in rural communities. HCEEP Rural Network.

Frohreich, L. F. (1973). Costing programs for exceptional children: Dimensions and indices. Exceptional Children, 517-524.

Fudala, J. B. (1974). Arizona Articulation Proficiency Scale. Los Angeles: Western Psychological Services.

Fudala, J. B. (1981). Arizona Articulation Proficiency Scale--Revised Manual. Los Angeles: Western Psychological Services.

Gage, N. C., & Berliner, D. C. (1979). Educational psychology. Chicago: Rand McNally.

Gaiter, J. L. (1982). The effects of intraventricular hemorrhage on Bayley developmental performance in preterm infants. Seminars in Perinatology, 6, 305-316.

Gallo, P. S. (1978). Meta-analysis--a mixed meta-phor. American Psychologist, 33, 515-517.

Garland, C., Swanson, J., Stone, N. W., & Woodruff, G. (Eds.) (1981). Early intervention for children with special needs and their families: Findings and recommendations. Seattle, WA: Washington University. (ERIC Document Reproduction SErvice No. ED 207 278)

Gentile, J. R., Roden, A. H., & Klein, R. D. (1974). An analysis of variance model for the intrasubject replication design. Journal of Applied Behavior Analysis, 5, 193-198.

Glass, G. V (1976). Primary, secondary, and meta-analysis of research. Educational Researcher, 5, 3-8.

Glass, G. V (1977). Integrating findings: The meta-analysis of research. Review of Research in Education, 5, 351-379.

Glass, G. V, McGaw, B., & Smith, M. L. (1981). Meta-analysis in social research. Beverly Hills: Sage.

Glass, G. V, Smith, M. L., & Barton, M. A. (1979). Methods of integrative analysis (Annual report, NIE Grant No. G-78-0148). Boulder, CO: Laboratory of Educational Research, Univrsity of Colorado.

Goldman, J., Englestein, C., & Guerry, S. (Eds.) (1983). Psychological methods of child assessment. NY: Brunner/Mazel Publishers.

Goodson, B. D., & Hess, R. D. (1975). Parents as teachers of young children: An evaluative review of some contemporary concepts and programs. Stanford: Stanford University, Department of Psychology.

Goodstein, H. A. (1974). The prediction of elementary school failure among high risk children. Storrs, CT: Connecticut University. (ERIC Document Reproduction Service No. ED 108 749)

Gordon, I. J. (1969) Stimulation via parent education. Children, 16, 57-58.

Gottfriend, N. W. (1973). Effects of early intervention programs. In K. S. Miller & R. M. Oreger (Eds.), Comparative studies of Blacks and Whites in the United States.

Gottman, J. M., & Glass, G. V (1978). Analysis of time-series experiments. In T. R. Kratochwill (Ed.). Single subject research: Strategies for evaluation change (pp. 197-235). New York: Academic Press.

Gramlich, E. M. (1981). Benefit-cost analysis of government programs. Englewood Cliffs, NJ: Prentice-Hall.

Grants, E. G., Borts, F. T., & Schellinger, D. (1981). Real-time ultrasound of neonatal intraventricular hemorrhage and comparison with computed tomography. Radiology, 139, 687-691.

Gray, S. W. (1977). Home-based programs for mothers of young children. In P. Mittler (Ed.), Research to practice in mental retardation: Care and intervention (Vol. 1). Baltimore, MD: University Park Press.

Gray, S. W., & Klaus, R. A. (1970). The early training project: A seventh-year report. Child Development, 41, 909-924.

Gray, S. W., Klaus, R. A., Miller, J. O., & Forrester, B. (1966). Before first grade. New York: Teachers College Press.

Green, P. E. (1978). Analyzing multivariate data. Hinsdale, IL: Dryden.

Hack, M., Merkatz, I. R., McGrath, S. K., Jones, P. K., & Fanaroff, A. A. (1984). Catch up growth in very-low-birth-weight infants. American Journal of Diseases in Children, 138, 370-375.

Halpern, R. (1984). Lack of effects for home-based early intervention? Some possible explanations. American Journal of Orthopsychiatry, 54, 33-42.

Hambleton, G., & Wigglesworth, J. S. (1976). Origin of intraventricular hemorrhage in the preterm infant. Archives of the Disabled Child, 51, 651-659.

Haring, N. G., & Hayden, A. H. (1981). Effectiveness and cost-efficiency of early intervention with handicapped children. Testimony before the Subcommittee on the Handicapped, of the Senate Committee on Labor and Human Development, U.S. Senate, Washington, DC.

Harris, D. B. (1963). Children's drawings as measures of intellectual maturity: A revision and extension of the Goodenough Draw-A-Person Test. New York: Harcourt, Brace, and World.

Hayden, A. H., & Smith, R. K. (1979). Mainstreaming preschoolers: Children with learning disabilities. (DHEW Publication No. OHDS 73-31117). Washington, DC: Head Start Bureau.

Hewett, K. D. (1977). Partners with parents: The Home Start experience with preschoolers and their families (DHEW Publication No. DHEW OHDS 78-31106). Washington, DC: U.S. Department of Health, Education, and Welfare.

Hill, A., & Rozdilsky, B. (1984). Congenital hydrocephalus secondary to intra-uterine germinal matrix/intraventricular hemorrhage. Developmental Medicine and Child Neurology, 26, 524-527.

Hodges, W. L., & Sheehan, R. (1978). Follow through as ten year of experimentation: What have we learned? Young Children, 34, 4-14.

Hoge, R. (1985). The validity of direct observation measures of pupil classroom behavior. Review of Educational Research, 55, 469-483.

Holland, A. L. (1980). Communication abilities in daily living. Baltimore, MD: University Park Pressl.

Hollingshead, A. B., & Redlich, F. C. (1957). Two factor index of social position. Unpublished manuscript.

Hopkins, K. (1969). Regression and the matching fallacy in quasi-experimental research. Journal of Special Education, 3, 329-336.

Horowitz, F. D., & Sullivan, J. W. (1981). Mother-child interaction issues. In R. L. Schiefelbusch, & D. D. Bricker (Eds.), Early language: Acquisition and intervention, (pp. 251-256). Baltimore: University Park Press.

Horst, D. P., Tallmadge, G. K., & Wood, C. T. (1975). A practical guide to measuring project impact on student achievement (No. 1, Stock No. 017-080-01460-2). Washington, DC: U.S. Government Printing Office.

Horton, K. B. (1976). Early intervention for hearing impaired infants and young children. In T. D. Tjossem (Ed.), Intervention strategies for high risk infants and young childrenf. Baltimore, MD: University Park Press.

Hubbell, R. (1983). A review of Head Start research since 1970 (Contract No. 105-81-C-026). Washington, DC: Administration for Children, Youth, and Families.

Huitema, B. E. (1985). Auto-correlation in applied behavior analysis: A myth. Behavioral Assessment, 7, 109-120.

Hunt, J. M. (1980). Implications of plasticity and hierarchical achievements for the assessment of development and risk of mental retardation. In D. Sawin, R. Hawkins, L. Walker, & J. Penticuff (Eds.), Exceptional Infant: Psychosocial risks in infant-environment transactions (Vol. 4, pp. 7-54). New York: Brunner/Mazel.

Huron Institute (1973). Short terms effects of Project Head Start: A preliminary report on the second year of planned variation--1970-1971. Cambridge, MA: Author.

Hynd, G. W., Hartlage, L. C., & Noonan, M (1984). Intracranial hemorrhage in neonates: Data on cognitive development. The International Journal of Clinical Neuropsychology, 6, 111-114.

Ireton, H., & Thwing, E. (1974). Minnesota child development inventory. Minneapolis, MN: Interpreting Scoring Systems.

Jackson, G. B. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.

Jason, L. (1975). Early secondary prevention with disadvantaged preschool children. American Journal of Community Psychology, 3, 33-46.

Jones, F. Green, V., & Kraus, D. (1980). Maternal responsiveness of primiparous mothers during the post partum period: Age differences. Pediatrics, 65, 579-583.

Jones, R. R., Weinrott, M. R., & Vaught, R. S. (1978). Effects of serial dependency on the agreement between visual and statistical inference. Journal of Applied Behavior Analysis, 11, 272-283.

Juster, F. T., & Stafford, F. P. (Eds.) (1985). Time, goods, and well-being. Ann Arbor, MI: University of Michigan, Institute for Social Research.

Kakalik, J. S., Furry, W. S., Thomas, M. A., & Carney, M. F. (1981). The cost of special education. Santa Monica, CA: Rand.

Karnes, M. B., Teska, J. A., & Hodgins, A. S. (1970). The effects of four programs of classroom intervention on the intellectual and language development of 4-year-old disadvantaged children. American Journal of Orthopsychiatry, 40, f40, 58-76.

Kavale, K. A. (1980). Auditory-visual integration and its relationship to reading achievement: A meta-analysis. Perceptual and Motor Skills, 51, 947-955.

Kavale, K. A. (1981). The relationship between auditory perceptual skills and reading ability: A meta-analysis. Journal of Learning Disabilities, 14 539-546.

Kavale, K. A. (1982). The efficacy of stimulant drug treatment for hyperactivity: A meta-analysis. Journal of Learning Disabilities, 15, 280-289.

Kavale, K. A., & Forness, S. R. (1983). Hyperactivity and diet treatment: A meta-analysis of the Feingold hypothesis. Journal of Learning Disabilities, 16, 324-330.

Kavale, K. A., & Forness, S. R. (1984). A meta-analysis assessing the validity of Wechsler Scale profiles and recategorizations: Patterns or parodies? Learning Disability Quarterly, 7, 136-156.

Kavale, K. A., & Forness, S. R. (1985). The science of learning disabilities. San Diego: College-Hill.

Kavale, K. A., & Nye, C. (1985-1986). Parameters of learning disabilities in achievement linguistics of neuropsychological and social/behavioral domains. Journal of Special Education, 19, 443-458.

Kazdin, A. E. (1976). Statistical analysis for single-case experimental designs. In M. Hersen & D. H. Barlow, Single-case experimental designs: Strategies for studying behavior change (pp. 265-316). New York: Pergamon Press.

Kazdin, A. E. (1978). Methodological and interpretive problems of single-case experimental designs. Journal of Consulting and Clinical Psychology, 46, 629-642.

Kenny, L., Lee, L. F., Maddalla, G., & Trost, R. (1979). Returns to college education: An investigation of self-selection bias based on the Project Talent data. International Economic Review, 20(3), 775-789.

Kerlinger, F. N. (1979). Behavioral research: A conceptual approach. New York: Holt, Rinehart and Winston.

Kirk, S. A., & Gallagher, J. J. (1983). Educating exceptional children. Boston: Houghton Mifflin.

Klaus, R. A., & Gray, S. S. (1968). They early training project for disadvantaged children: A report after five years. Monographs of the Society for Research in Child Development, 33,(4 Serial No. 120).

Kohlberg, L., & Zigler, E. (1967). The impact of cognitive maturity on the developent of sex-role attitudes in the years four to eight. Genetic Psychology Monographs, 75, 89-165.

Lahey, M. (1978). Readings in childhood language disorders. New York: John Wiley & Sons.

Landry, S. H., Fletcher, J. M., Zarling, C. L., Chapieski, L., Francis, D. J., & Denson, S. (1984). Differential outcomes associated with early medical complications in premature infants. Journal of Pediatric Psychology, 9, 385-401.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education: A report from the consortium for longitudinal studies. Monographs of the Society for Research in Child Development, 47(2-3), Serial No. 195.

Levin, H. (1981). Cost analysis. In N. Smith (Ed.), New techniques in evaluationf. London: Sage.

Levin, H. (1983). Cost-effectiveness: A primer. Beverly Hills, CA: Sage.

Levin, J. R., Marascuilo, L., & Hubert, L. J. (1978). N = nonparametric randomization tests. In T. R. Kratochwill (Ed.), Single-subject research: Strategies for evaluating change (pp. 167-196). New York: Academic Press.

Liberman, A., Barnes, M. J., Ho, S. E., Cuellar, I., & Little, T. (1979). The economic impact of child development services on families of retarded children. Mental Retardation, 17, 158-159.

Light, R. J., & Smith, P. V. (1971). Accumulating evidence: Procedures for resolving contradictions among different studies. Harvard Educational Review, 41, 429-471.

Lindfors, J. W. (1980). Children's language and learning. Englewood Cliffs, NJ: Prentice-Hall.

Littlejohn, R., & Associates (1982). An analysis of the impact of HCEEP: A report to SEP. Washington, DC: LRA, Inc.

Longstreth, L. E. (1981). Revisiting Skeels' final study: A critique. Developmental Psychology, 17, 620-625.

Lowell, E. L. (1967). Home teaching for parents of young deaf children. Final Report. Los Angeles, CA: John Tracy Clinic.

Macy, D. J., & Carter, J. L. (1980). Triple T infant Consortium follow-up study. Dallas: Macy Research Associates.

Magidson, J., & Sorbom, D. (1982). Adjusting for confounding factors in quasi-experiments: Another reanalysis of the Westinghouse Head Start evaluation. Educational Evaluation and Policy Analysis, 4, 321-329.

Marquis, P. (1976). Cognitive stimulation. fAmerican Journal of Diseases of Children, 130, 410-415.

Mansfield, R. S., & Bussey, T. V. (1977). Meta-analysis of research: A rejoinder to Glass. Educational Research, 6, 3.

Mastropieri, M. A., & Scruggs, T. E. (1985-1986). Early intervention for socially withdrawn children. Journal of Special Education, 19, 429-441.

Mastropieri, M. A., Scruggs, T. E., & Casto, G. (1985). Early intervention for behaviorally disordered children: An integrative review. Monographs in Behavior Disorders, 8, 86-104.

McCain, L. J., & McCleary, R. (1979). The statistical analysis of the simple interrupted time-series quasi-experiment. In T. D. Cook & D. T. Campbell (Eds.), Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally.

McCarthy, D. (1972). Manual for the McCarthy Scales of Children's Abilities. New York: The Psychological Corporation.

McDaniels, G. (1977). Successful programs for young handicapped preschool children. Educational Horizons, 56, 26-33

McGee. G. A. (1972). An evaluation of the effects of the Besell-Palomares human program on five year olds in an Appalachian Head Start class. Dissertation Abstracted International, 32(8-a):4329.

McNulty, B. A., Smith, D. B., & Soper, E. W. (1983). Effectiveness of early special education for handicapped children. Denver: Colorado Department of Education.

Meisels, S. J. (1985). The efficacy of early intervention: Why are we still asking this question? Topics in Early Childhood Special Education, 5(2), 1-11.

Ment, L. R., Scott, D. T., & Rothman, S. G. (1978). Prospective long-term follow-up of prematures with subependymal/intraventricular hemorrhage. Pediatric Research, 15, 711.

Mercer, J. R. (1973). Labeling the mentally retarded. Berkeley: University of California Press.

Meyen, E. L. (1982). Exceptional children and youth. Denver, CO: Love Publishing Co.

Miller, D. C. (1983). Social status. In D. C. Miller (Ed.), Handbook of research design and social measurement (pp. 275-282). New York: Longman, Inc.

Miller, L. B., Bizzell, R. P. (1984). Long term effects of four preschool programs: Ninth and tenth grade results. Child Development, 55, 1570-1535.

Miller, L. B., & Dyer, J. L. (1975). Four preschool progr Their dimensions and effects. Monographs of the Society for Research in Child Development, 40 (Serial No. 162).

Mischel, W. (1968). Personality and assessment. New York: Wing.

Mishan, E. J. (1976). Cost-benefit analysis. New York: Praeger.

Moerk, E. L. (1972). Verbal interactions between children and their mothers during the preschool years. Developmental Psychology, 11, 788-794.

Moerk, E. L. (1977). Pragmatic and semantic aspects of early language development. Baltimore: University Park Press.

Moog, J. S., & Geers, A. V. (1975). Scales for early communication skills for hearing impaired children. St. Louis, MO: CID Publication.

Moores, D. F., Weiss, K. L., & Goodwin, M. W. (1978). Early education programs for hearing impaired children: Major findings. American Annals of the Deaf, 123, 925-964.

Mulliken, R. K., & Buckley, J. J. (1983). Assessment of multihandicapped and developmentally disabled children. Rockville, MD: Aspen Systems Corporation.

Mundy, M. J. (1973). An analysis of an academically structured Head Start program for: (1) geographic, (2) academically structured, and (3) high-low subjectability variables.

Murphy, T. G., Nichter, C. A., & Liden, C. B. (1982). Developmental outcome of the high-risk infant: A review of methodological issues. Seminars in Perinatology, 6, 353-364.

National Diffusion Network (1983). Educational Programs that work. San Francisco: Far West Laboratory for Educational Research and Development.

Newborg, J., Stock, J., & Wnek, L. (1984). Battelle developmental inventory. Allen TX: DLM Teaching Resources.

Newport, E. (1976). Motherese: The speech of mothers to young children. In J. J. Castellan, D. B. Pisoni, & G. R. Potts (Eds.), Cognitive theory, (Vol. 2). Hillsdale, NJ: Erlbaum.

Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. (Eds.) (1975). Statistical package for the social sciences: (2nd ed.). New York: McGraw-Hill.

North, A. F. (1979). Health services in Head Start. In E. Zigler & J. Valentine (Eds.), Project Head Start: A leacy of the war on poverty. New York: Free Press.

Northern, J. L., & Downs, M. P. (1974). Hearing in children. Baltimore, MD: Williams and Wilkins.

Odom, S. L., & Fewell, R. R. (1983). Program evaluation in early childhood special education: A meta-evaluation. Educational Evaluation and Policy Analysis, 5, 445-460.

Papile, L. A., Munsick-Bruno, O. T. R., & Schaefer, A. (1983). Relationship of cerebral intraventricular hemorrhage and early childhood neurologic handicaps. The Journal of Pediatrics, 103, 273-277.

Papile, L. A., Munsick, G., Weaver, N., & Pecha, S. (1979). Cerebral intraventricular hemorrhage in infants less than 1500 grams: Developmental follow-up at one year. Pediatric Research, 13, 528.

Papile, L., Burstein, J., Burstein, R., & Koffler, H. (1978). Incidence and evolution of subependymal and intraventricular hemorrhage: A study of infants with birth weights less than 1500 gm. The Journal of Pediatrics, 92, 529-534.

Papile, L. A., Munsick-Bruno, G., & Schaefer, A. (1983). Relationship of cerebral intraventricular hemorrhage and early childhood neurologic handicaps. The Journal of Pediatrics, 103, 273-277.

Parsonson, B. S., & Baer, D. M. (1978). The analysis and presentation of graphic data. In T. R. Kratochwill (Ed.), Single subject research: Strategies for evaluating change (pp. 101-166). New York: Academic Press.

Pasternak, J. F., Groothuis, D. R., Fischer, J. M., & Fischer, D. I. (1983). Regional blood flow in the beagle puppy model of neonatal intraventricular hemorrhage: Studies during systemic hypecterism. Neurology, 33, 559-566.

Pasternak, J. F., Mantovani, L. F., & Volpe, J. J. (1980). Porencephaly from periventricular intracerebral hemorrhage in the premature infant. American Journal of the Disabled Child, 134, 673-675.

Philips, W. P. (1963). Influence of preschool training on achievement in language arts, arithmetic concepts, and socialization of young deaf children. Unpublished doctoral dissertation, Teachers College, Columbia.

Pillemer, D. B., & Light, R. J. (1980). Synthesizing outcomes: How to use research evidence from many studies. Harvard Educational Review, 50, 176-195.

Powers, S., Slaughter, H., & Helmick, C. (1983). A test of the equipercentile hypothesis of the TIERS norm-referenced model. Journal of Educational Measurement, 20, 299-302.

Ragozin, A. S., Basham, R. B., Crnic, K. A., Greenberg, M T., & Robinson, N. M. (1982). Effects of maternal age on parenting role. Developmental Psychology, 18, 627-634.

Ramey, C. T., & Campbell, F. A. (1979). A compensatory education for disadvantaged children. School Review, 82, 171-189.

Ramey, C. T., Bryant, D., & Suarez, T. (1985). Preschool compensatory education and the modification of intelligence: A critical review. In D. Detterman (Ed.), Current topics in human intelligence (pp. 247-296). Norwood, NJ: Albex.

Ramey, C. T., Bryant, D., & Suarez, T. (1985). Preventing developmental retardation: A general systems model. In L. A. Bond & J. M. Joffe (Eds.), Facilitating infant and early childhood development (pp. 345-401). Hanover, NH: University Press of New England.

Ramey, C. T., & Haskins, R. (1981). The modification of intelligence through early experience. Intelligence, 4, 5-19.

Ravens, R. G. (Ed.) (1977). Raven's progressive matrices handbook. New York: Educational Manuals.

Revsky, S. H. (1967). Some statistical treatments compatible with individual organism methodology. Journal of the Experimental Analysis of Behavior, 10, 319-330.

Robinson, J. P. (1985). Testing the validity and reliability of diaries versus alternative time use measures. In E. T. Juster & F. P. Stafford (Eds.), Time, goods, and well-being. Ann Arbor: University of Michigan, Institute for Social Research.

Rosenthal, R. (1978). Combining results of independent studies. Psychological Bulletin, 85, 185-193.

Rosenthal, R. (1984). Meta-analytic procedures for social research. Beverly Hills, CA: Sage.

Ross, G., Schechner, S., Frayer, W., & Auld, P. (1982). Perinatal and neurobehavioral predictors of one-year outcome in infants less than 1500 grams. Seminars in Perinatology, 6, 317-326.

Rourke, B. P., Bakker, D. J., Fisk, J. L., & Strang, A. (Eds.) (1983). Child neuropsychology: An introduction to theory, research, and clinical practice. New York: Guilford Press.

Rule, S., Stowitschek, J., Innocenti, M., Striefel, S., Killoran, J., Swezey, K., & Boswell, C. (1987). The social integration program: An analysis of the effects of mainstreaming handicapped children into daycare centers. Education and Treatment of Children, 10, 175-192.

Ruopp, R., Travers, J., Glantz, F., & Coelen, C. (1979). Children at the center: Summary findings and policy implications of the National Daycare Study. Cambridge, MA: Abt Associates.

Rutherford, R. B., & Nelson, C. M. (1982). Analysis of the response contingent time-out literature with behaviorally disordered students in classroom settings. Monographs in Behavior Disorders, 5, 79-105.

Rutherford, R. B., Jr., & Nelson, C. M. (in press). Applied behavior analysis in education: Generalization and maintenance. In J. C. Witt, S. N. Eliott, & F. M. Greshman (Eds.), Handbook of behavior therapy in education. New York: Plenum.

Rutter, M. (1982). Syndromes attributed to "minimal brain dysfunction" in childhood. American Journal of Psychiatry, 138, 1533-1544.

Sachs, J., Brown, R., & Salerno, R. (1972). Adult speech for children. Paper given at the International Symposium on First Language Acquisition, Florence, Italy.

Sameroff, A. J. (1975). Early influences on development: Fact or fancy? Merrill-Palmer Quarterly, 21, 267-294.

Schub, H. S., Ahmann, P. A., Dykes, F. D., Lazzara, A., & Blumenstein, B. (1981). Prospective long-term follow-up of prematures with subependymal/intraventricular hemorrhage (SEH/IVH). Pediatric Research, 15, 711.

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (in press). Early intervention for behaviorally disordered children: The synthesis of single-subject research. Monographs in Behavior Disorders.

Scruggs, T. E., Mastropieri, M. A., & Cook, S. B. (1986). Preschool language interventions: A quantitative synthesis of single-subject research. Logan, UT: Early Intervention Research Institute.

Scruggs, T. E., Mastropieri, M. A., Cook, S. B., & Escobar, C. (in press). Early intervention for children with conduct disorders: A quantitative synthesis of single-subject research. Behavioral Disorders.

Scruggs, T. E., White, K. R., Bennion, K. (in press). Improving achievement test scores in the elementary grades by coaching: A meta-analysis. Elementary School Journal.

Seitz, V., Apfel, N. H., & Rosenbaum, L. (1981). Projects Head Start and Follow Through: A longitudinal evaluation of adolescents. In M. J. Begab, H. Garber, & H. C. Haywood (Eds.), Psychosocial influences in retarded performance, (Vol. 2, pp. 17-28). Baltimore, MD: University Park Pres.

Select Committee on Children, Youth and Families (1985). Opportunities for success: Cost-effective programs for children (staff report). Ninety-ninth Congress, First Session, Washington, DC: USGPO.

Shaver, J. P. (1979). The usefulness of educational research in curricular/ instructional decision-making in social studies. Theory and Research in Social Educationf, 7, 21-46.

Shearer, D. E., & Shearer, M. S. (1976). The Portage Project: A model for early childhood intervention. T. D. Tjossem (Ed.), Intervention strategies for high risk infants and young children. Baltimore, MD: University Park Press.

Sheehan, R., & Keogh, B. (1982). Design and analysis in the evaluation of early childhood special education programs. Topics in Early Childhood Special Education, 1, 81-88.

Siegel, S. (1956). Nonparametr'c Statistics. New York: McGraw-Hill.

Siegel, L. S. (1982). Reproductive, perinatal, and environmental variables as predictors of development of preterm (< 1501 grams) and fullterm children at 5 years. Seminars in Perinatology, 6, 274-279.

Simeonsson, R. J., cooper, D. H., & Scheiner, A. P. (1982). A review and analysis of the effectiveness of early intervention programs. Pediatrics, 69, 635.

Simpson, S. N. (1979). Comment on "meta-analysis of research on class size and achievement. London, England: Institute of Education.

Sindelar, P. T., & Wilson, R. J. (1984). The potential effects of meta-analysis on special education practice. Journal of Special Education, 18, 81-92.

Skeels, H. M. (1966). Adult status of children with contrasting early life experiences. Monographs of the Society for Research in Child Development (Serial No. 105), 31(3).

Skiba, R. J., Casey, A., & Center, B. A. (1985-1986). Nonaversive procedures in the treatment of classroom behavior problems. Journal of Special Education, 19, 459-481.

Slavin, R. E. (1984). Meta-analysis in education: How has it been used? Educational Researcher, 13(8), 6-15.

Slosson, R. L. (1984). The Slosson Intelligence Test. Ease Aurora, NY: Slosson Educational Publications, Inc.

Smith, L., Somner, F. F., & von Tetzchner, S. (1982). A longitudinal study of low birthweight children: reproductive, perinatal, and environmental precursors of developmental status at three years of age. Seminars in Perinatology, 6, 29 -304.

Snider, J. N., Sullivan, W. G., & Manning, D. (1974). Industrial engineering participation in a special education program. Tennessee Engineer, 1, 21-23.

Stewart, A. (1983). Severe perinatal hazards. In M. Rutter (Ed.), Developmental neuropsychiatry (pp. 15-31). New York: Guilford Press.

Stewart, A. L., Reynolds, E. O. R., & Lipscomb, A. P. (1981). Outcome for infants of very low birthweight: Survey of world literature. Lancet, 1038-1041.

Stile, S. W., & Thompson, J. (1982, April). Intervention efficiency, subsequent placements, and cost-effectiveness of two preschool programs for handicapped children. Paper presented at the 60th International Conference of the Council for Exceptional Children, Houston.

Stock, J. R., Wnek, L. L., Newborg, J. A., Schenck, E. A., Gabel, J. R., Spurgeon, M. S., & Ray, H. W. (1976). Evaluation of the Handicapped Children's Early Education Program. Columbus, OH: Battelle Memorial Institute.

Strain, P., & Smith, B. (1984). The argument for early intervention. ERIC Digest. Reston, VA: Clearinghouse on Handicapped and Gifted Children.

Suppes, P., & Morningstar, M. (1969a). Computer-assisted instruction. Science, 166, 343-350.

Suppes, P., & Morningstar, M. (1969b). Evaluation of three computer-assisted instruction programs (Technical Report No. 142, Psychology Series). Stanford, CA: Institute for Mathematical Studies in the Socal Sciences.

Tallmadge, G. K. (1977). Ideabook: The joint dissemination review panel. Washington, DC: U.S. Office of Education.

Tarby, T. J., & Volpe, J. J. (1982). Intraventricular hemorrhage in the premature infant. Pediatric Clinics of North America, 29, 1077-1104.

Tawney, J. W., & Gast, D. L. (1984). Single-subject research in special education. Columbus, OH: Merrill.

Taylor, C., White, K. R., & Pezzino, J. (1984). Cost-effectiveness analysis of full-day versus half-day intervention programs for handicapped preschoolers. Journal of the Division for Early Childhood, 9, 76-85.

Tekolste, K. A., Bennett, F. C., & Mack, L. A. (1985). Follow-up of infants receiving cranial ultrasound for intracranial hemorrhage. American Journal of Diseases in Childhood, 139, 299-303.

Thompson, M. A. (1980). Benefit-cost analysis for program evaluation. Beverly Hills, CA: Sage.

Tizard, B. (1981). Language at home and at school. In C. B. Cazden (Ed.), Language in early childhood education. Washington, DC: NAEYC.

U.S. Department of Education (1985). Seventh annual report to Congress on the implementation of the Education of the Handicapped Act. Washington, DC: Office of Special Education and Rehabilitation Services, Division of Special Education Programs.

U.S. Department of Education. (1986). Eighth annual report to Congress on the implementation of the Education of the Handicapped Act. Volume I. Washington, DC: Office of Special Education and Rehabilitation Services, Division of Special Education Programs.

United States Department of Labor, Bureau of Labor Statistics (1986). Employment and Earnings, 33(5), 81. Washington, DC: Author.

Vernon, M., & Koh, S. D. (1970). Effects of early manual communication in achievement of deaf children. American Annals of the Deaf, 115, 527-536.

Volpe, J. J. (1981). Neonatal intraventricular hemorrhage. The New England Journal of Medicine, 304, 886-891.

Volpe, J. J. (1987). Intracranial hemorrhage: Periventricular intraventricular hemorrhage of the premature infant. In J. J. Volpe (Ed.), Neuropsychology of the newborn. Philadelphia, PA: Saunders.

Volpe, J. J., Herscovitch, P., Perlman, J. M., & Raichle, M. E. (1983). Positron emission tomography in the newborn: Extensive impairment of regional cerebral blood flow with intraventricular hemorrhage hemorrhagic intracerebral involvement. Pediatrics, 72, 589-601.

Walker, K. (1981). Project Sunrise. In T. Black & P. Hutinger (Eds.), Cost-effective delivery strategies in rural areas: Programs for young handicapped children (Vol. 1, pp. 28-35). HCEEP Rural Network Monograph.

Wallace, I. F., Escalona. S. K., McCarton-Daum, C., & Vaughan, H. G. (1982). Neonatal precursors of cognitive development in low birthweight children. Seminars in Perinatology, 6, 327-333.

Watkins, S. (1971). Guidelines for a model hearing impaired infant program. Unpublished master's thesis, Utah State University, Logan.

Weber, C. U., Foster, P. W., & Weikart, D. P. (1978). An economic analysis of the Ypsilanti Perry Preschool Project. Ypsilanti, MI: High/Scope.

Weikart, D. P., Bond, J. T., & McNeil, J. T. (1978). The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade. Ypsilanti, MI: High/Scope.

Weikart, D. P., Epstein, A. S., Schweinhart, L., & Bond, J. T. (1978). The Ypsilanti Preschool Curriculum Demonstration Project (Monograph Series No. 4). Ypsilanti, MI: High/Scope Educational Research Foundation.

Weikart, D. P., Kamii, C. K., & Radin, N. L. (1967). Perry Preschool Project progress report. In D. P. Weikart (Ed.), Preschool intervention: A preliminary report of the Perry Preschool Project (pp. 1-61). Ann Arbor, MI: Campus.

Weiss, R. (1981). INREAL intervention for language handicapped and bilingual children. Journal of the Division for Early Childhood, 4, 40-51.

Weiss, R., & Heublein, E. (1983). Colorado school district INREAL experimental study. In B. McNulty, D. Smith, and E. Soper (Eds.), Appendix to "Effectiveness of early childhood special education for handicapped children." Denver, CO: Colorado Department of Education.

Weiss, S., & Jurs, S. (1984, October). Cost-effectiveness of early childhood education for handicapped children. Paper presented at the annual meeting of the Atlantic Economic Society, Montreal.

White, K. R. (1982). An integrative review of the relationship between socio-economic status and academic achievement using meta-analysis techniques. Psychological Bulletin, 91, 461-481.

White, K. R., & Myette, B. (1982, March). Drug treatment of hyperactivity: A meta-analysis. Paper presented at the annual meeting of the American Educational Research Association, New York.

White, K. R., & Taylor, C. (1982, March). Typical interaction research: An example from the literature with suggestions for future practice. Paper presented at the annual meeting of the American Educational Research Association, New York.

White, K., Taylor, C., Carcelli, L., & Eldred, N. (1981). State refinements to the ESEA Title I Evaluation and Reporting system: Utah 1979-80 Project. Final report. Salt Lake City, UT: Utah State Office of Education.

Williams, M. L., & Scarr, S. (1971). Effects of short-term intervention on performance in low-birth-weight, disadvantaged children. Pediatrics, 47, 289-298.

Williamson, W. D., Desmond, M. M., Wilson, G. S., Andrew, L., & Garcia-Prats, J. A. (1982). Early developmental outcome of low birth weight infants surviving neonatal intraventricular hemorrhage. The Journal of Perinatal Medicine, 10, 34-41.

Wood, M. E. (1981). Costs of intervention programs. In C. Garland, N. W. Stone, J. Swanson, & G. Woodruff (Eds.), Early intervention for children with special needs and their families: Findings and recommendations (pp. 15-25). WESTAR Series paper No. 11, Seattle: University of Washington.

Wooden, H. (1976). Volunteers, Head Start children, and development. Academic Therapy, 11, 449-454.

Zigler, E. (1966). Research on personality structure in the retardate. In N. R. Ellis (Ed.), International review of research in mental retardation, (Vol. 1). New York: Academic Press.

Zigler, E. (1969). Developmental versus difference theories of mental retardation and the problem of motivation. American Journal of Mental Deficiency, 73, 536-556.

Zigler, E., Abelson, W. D., & Seitz, V. (1973). Motivational factors in the performance of economically disadvantaged children on the Peabody Picture Vocabulary Test. child Development, 44, 294-303.

Zigler, E., & Balla, D. (1982). Selecting outcome variables in evaluations of early childhood special education programs. Topics in Early Childhood Special Education, 1, 11-22.

Zigler, E., Balla, D., & Hodapp, R. (1984). On the definition and classification of mental retardation. American Journal of Mental Deficiency, 89, 215-230.

Zigler, E., & Trickett, P. K. (1978). IQ, social competence, and evaluation of early childhood intervention programs. American Psychologist, 33, 789-798.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (1979). Preschool Language Scale--Revised Edition. Columbus, OH: Merrill.