INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
Determining the effects of inservice training on the job performance of child welfare workers serving children with developmental disabilities

Rycus, Judith Sonya, Ph.D.
The Ohio State University, 1990
DETERMINING THE EFFECTS OF INSERVICE TRAINING
ON THE JOB PERFORMANCE OF CHILD WELFARE WORKERS
SERVING CHILDREN WITH DEVELOPMENTAL DISABILITIES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy in the Graduate School
of The Ohio State University

By

Judith S. Rycus, M.S.W.

The Ohio State University
1990

Dissertation Committee:                  Approved By:

Henry Leland, Ph.D.
David Hammer, Ph.D.
Johannes Rojhan, Ph.D.
Samuel Dixon, DSW.

Adviser
Department of Psychology
ACKNOWLEDGMENTS

Thanks to Henry Leland, Ph.D., David Hammer, Ph.D. and Johannes Rojhan, Ph.D. for their suggestions and direction during the planning, execution, and review of this research; to Ronald C. Hughes, Ph.D. and Rhonda J. Lilley, Ph.D. for their invaluable assistance in the collection and scoring of the research data; and to David Boninger for his help in conducting the statistical analyses.

Thanks to the administrators and staff of Ohio's county child welfare agencies for their interest and willingness to support and participate in this research.

A very special thanks to Dr. Leland, whose continuing support, guidance, and willingness to be innovative in applying developmental psychology made this research possible.
VITA

September 10, 1946 Born, Detroit, Michigan

1968 B.A. Psychology, Wayne State University, Detroit

1969-1971 Adoption Caseworker, Los Angeles County Department of Adoptions, Los Angeles, California

1972 Master of Social Work Wayne State University, Detroit, Michigan

1973-1977 Supervisor and Trainer, Franklin County Children's Services, Columbus, Ohio

1977 - Present Program Director and Consultant, Institute for Human Services, Columbus, Ohio

PUBLICATIONS


**FIELDS OF STUDY**

Major Field of Study: Psychology

Studies in developmental psychology and developmental disabilities with Henry Leland, Ph.D.

Secondary Field: Social Work

Studies in child welfare, group work, and clinical social work.
# TABLE OF CONTENTS

**ACKNOWLEDGMENTS**  

**VITA**  

**LIST OF TABLES**  

**CHAPTER**  

I. STATEMENT OF THE PROBLEM  

II. REVIEW OF THE LITERATURE  

III. METHODOLOGY  

IV. RESULTS  

V. DISCUSSION  

**APPENDICES**  

A. Developmental Disabilities Quiz  

B. Behavior Inventory Instrument Interview Questionnaire  

C. Behavior Inventory Instrument Scoring Criteria  

D. Subject Data Sheet  

**REFERENCES**
LIST OF TABLES

TABLE 1: Summary of Workshops, 1981-82 64
TABLE 2: Comparison of Subjects in Experimental and Control Groups on Significant Demographic Variables 70
TABLE 3: Comparison Between Groups on Developmental Disabilities Quiz Scores 85
TABLE 4: Comparison Between Groups on Behavioral Criteria 90
TABLE 5: Comparison of Test Scores and Total Points on Behavioral Criteria for All Subjects 92
Chapter I

Statement of the Problem

Legislative and consumer advocacy during the 1970's resulted in the establishment of national standards to assure an improved quality of life for persons with developmental disabilities. Federal legislation mandated equal rights for citizens with disabilities in almost all social systems, including education, transportation, health and medical care, employment, mental health, and child protection. Underlying values of self-determination and normalization led to the development of national policies for deinstitutionalization and community placement in the least restrictive environment. The impact on the social service delivery system has been pervasive.

Yet, despite these changes, many children with developmental disabilities still do not receive the full range of services that are necessary for them to develop to their potential and to assume active participation in society (Schilling, Kirkham, & Schinke, 1986; Hughes & Rycus, 1983; Kurtz, 1979.) Many factors contribute to this problem, including a widespread lack of knowledge and skill regarding disabilities among professional staff in most service systems (Martin & Laidlow, 1980; Falconer, 1982), and a serious lack of coordination between community
social agencies that serve children with developmental disabilities and their families (Richardson, West, Day, & Stuart, 1989).

Persons with developmental disabilities often have diverse and complex service needs that require a variety of specialized services (Salisbury, 1984; Hughes & Rycus, 1983; Kurtz, 1979; Richardson, West, Day, & Stuart, 1989). The administrative functions of the agencies that provide these services are frequently separated at the federal, state, and local levels. These agencies are often structured to meet different legislative and regulatory mandates, they are funded by different public and private funding sources, and their services are most often delivered by different local agencies. The service mission and goals of each agency promote remediation of the "primary condition" the system has been organized to serve. As a result, the service agencies' values and service objectives may be quite different from one another, and at worst, may be contradictory. This makes interagency coordination and collaboration difficult, if not impossible, in a service system that depends upon such coordination to assure the effective delivery of comprehensive services to clients (Richardson, West, Day & Stuart, 1989; Hughes & Rycus, 1983).
Because of this specialization of function, agencies not mandated to serve citizens with disabilities rarely include tasks related to developmental disability services in the job descriptions of their professional social service staff. These staff are, therefore, unprepared to serve clients with disabilities (Martin & Laidlow, 1980). They may not be able to recognize or identify developmental disabilities (Schilling, Kirkham, & Schinke, 1980), they lack skills to provide comprehensive case management (Kurtz, 1979), and they lack skills to engage in professional teamwork to promote community service coordination (Richardson, West, Day, & Stuart, 1989).

A lack of proper staff training has particularly serious implications for clients with developmental disabilities. Savage, Novak and Heal state that the philosophy of normalization and deinstitutionalization "call for the full and complete utilization of generic services by the developmentally disabled individuals and their families," creating reliance upon "a network of community services that is prepared to adapt itself to the full range of human handicapping conditions." They cite a significant correlation between the underutilization of generic services in the community and the number of persons returning to an institution after having been placed in the community. They contend that poor or absent training of staff within community
social service agencies is the most often cited reason for this underutilization of services (Savage, Novak & Heal, 1980). The research of Martin and Laidlow (1980) indicates that in over half the cases of readmission of persons with disabilities to institutions, the primary cause of failure in community living was "a lack of training among professionals," which resulted in staff who "were not used to dealing with the developmentally disabled," Falconer (1982) suggests that health care professionals are often unable to provide adequate services to persons with disabilities because of their lack of understanding of the scope and extent of particular developmental disabilities.

Children with developmental disabilities form a significant subpopulation in the child welfare system. A recent study by the Institute for Child Advocacy reported that approximately 20% of all children in the custody of the Ohio Department of Human Services and placed in out-of-home care were mentally retarded or developmentally disabled (Institute for Child Advocacy, 1987). Richardson, West, Day, and Stuart report findings of a study conducted in Hennepin County, Minnesota, that suggests that as many as 40% of all children receiving child welfare services have developmental problems (Richardson et.al., 1989). Coyne and Brown (1985) conducted a mail survey to 799 agencies in 49 states, 8 Canadian provinces, and Washington D.C. to determine
the status of adoption of children with developmental disabilities. Data from the responding agencies indicated that in a one-year period, 1588 children with developmental disabilities had been placed in adoptive homes by 324 agencies in 46 states, 6 provinces, and Washington D.C. The majority of these children were mentally retarded or had cerebral palsy, and the remainder had a wide variety of other handicapping conditions, many of which were severely disabling. Considering the impact of a permanent adoptive placement on a child's life, this represents a substantial number of children whose futures are dependent upon the skills of child welfare professionals.

The presence of large numbers of children with disabilities in the child welfare system is largely due to the reciprocal relationship between child abuse and neglect and developmental disabilities. Many contributing factors to abuse and neglect are also contributors to developmental disabilities, including; poor diet and nutrition; lack of regular and adequate prenatal and medical care; unsanitary and unhealthy living environment; exposure to harmful and toxic substances; drug and alcohol abuse; poor supervision and subsequent injury of children; and lack of nurturance and stimulation. Physical abuse has been widely identified as a significant causative factor for developmental disabilities. Martin (1972) has conducted numerous studies
regarding the effects of abuse on development. One study of 42 abused children, conducted over a 3-year period, found that 43% had neurological abnormality on follow-up examination; 33% were functioning at a mentally retarded level, with a measured IQ of less than 80; and a large number of the children, including those with normal intelligence, demonstrated absent, minimal, or impaired speech and language. Martin also references Prugh and Harlow (1962), Spitz and Wolf (1946), and Provence and Lipton (1962), all of whom report that environmental deprivation, such as that experienced by institutionalized infants, can result in retardation. Martin also reaffirmed that permanent retardation can result from undernutrition during the first year of life. Helfer, McKinney and Kempe (1976) assert that the "most disturbing and consistent finding in observation of young children who have been abused and neglected is the delay, or arrest, of their development." They identified several developmental problems frequently seen in abused and neglected children, including difficulty in feeding, delay in motor development, poor muscle tone, delay in the development of social skills, absence of social responses such as smiling and vocalization, lack of activity, generalized apathy toward other people, consistent and considerable delay in the acquisition of speech, and an absence of distress when separated from parents, possibly indicating a lack of discrimination of significant others.
The reciprocal relationship between maltreatment and disabilities is further supported by the fact that children with developmental disabilities are at considerably higher risk of maltreatment than are non-disabled children. After many years of providing therapy to abusive parents, Steele (1987) concluded that in parents with the personality disposition to abuse, a common precipitating factor for the abuse was the parent's perception of the child as being somehow "special," "different," or "defective." Children with developmental disabilities very often are different in their physical appearance and limited in their abilities, and they may require very specialized care and treatment. Painful emotional responses to children with disabilities are common for many parents. Caring for a child with a disability can create enormous emotional and environmental stresses for families, leading to both concrete and emotional losses. The parents' confidence and self-esteem may also be threatened. Families of children with severe disabilities often need considerable physical and emotional support. When one considers that most abusive parents have low-self esteem and lack confidence, have limited frustration tolerance and poor impulse control, hold unrealistically high expectations for their children's behavior, and do not turn to other people for help, a child with a disability in such a family is at exceptionally high risk of maltreatment (Steele, 1987).
Finally, the physical and cognitive limitations of children with developmental disabilities increase the likelihood of victimization by adults. In a survey of 87 non-institutionalized adolescent females with mental retardation, one-third of the youth who were mildly retarded and one-fourth of the group who were moderately retarded had been sexually abused (Richardson et al., 1989).

There is considerable evidence to suggest that children with developmental disabilities in the child welfare system are not always recognized, nor are they properly served. Schilling, Kirkham, and Schinke (1986) suggest that the child protection system may be neglecting many children with developmental disabilities by failing to recognize or to document them. Kurtz (1979) indicates that social work professionals are often unaware of handicapping conditions in their clients; and, are less than enthusiastic about trying to prevent these conditions or to promote early intervention services. Richardson, West, Day, and Stuart (1989) surveyed the organization and provision of child welfare services for children with developmental disabilities. Their data indicates that the national child welfare system is not programmatically equipped to serve these children. The authors claim this is largely because child welfare legislation, which has shaped federal and state responses to child protection and out-
of-home placement, does not specifically refer to children with developmental disabilities. They indicated that most states, therefore, must rely heavily on interagency agreements as a means of providing and coordinating services to abused and neglected children with developmental problems. Kadushin (1980) concurs, suggesting that services to children with mental retardation and physical handicaps have historically had a "different frame of reference from child welfare services," and therefore, traditional child welfare agencies have operated in these areas "only tangentially."

A central tenet in providing services to children with developmental disabilities is that early recognition of disabling conditions and timely remedial and supportive services are essential (Kurtz, 1979; Hughes & Rycus, 1983). Early intervention can often correct or compensate for disabling conditions, can prevent the cumulative negative effects of the condition on normal growth, and can help guard against deterioration. Since child welfare workers have regular and frequent contact with a population of children at high risk of developmental disabilities, they are therefore in an excellent position to provide timely and comprehensive early intervention services (Kurtz, 1979; Schilling, Kirkham, & Schinke, 1986; Hughes & Rycus, 1983.)
Despite the absence of legislation to mandate such services, and in spite of Kadushin's (1980) onerous statement that traditional child welfare agencies operate in the area of developmental disabilities "only tangentially," the beginnings of systemic changes are evident. In 1981, the Child Welfare League of America, the national standard setting organization for child welfare services, adopted policy that stresses the responsibility of child welfare agencies to provide comprehensive services to children with developmental disabilities. In the preface to the League's publication, *Child Welfare Services for Children with Developmental Disabilities* (Hughes & Rycus, 1983), Executive Director Edwin F. Watson affirmed:

"It is of the utmost importance that policy makers and program planners reaffirm and reemphasize their commitment to the provision of a continuum of services within the child welfare system that will meet the needs of children with disabilities. Services for this population of children must be clearly articulated, conscientiously planned, and consistently delivered." (Hughes & Rycus, 1983, p.7).
If child welfare agencies are to effectively discharge this responsibility, they must be properly prepared. Local agencies must routinely include services for children with developmental disabilities in their service planning and programming. However, without a legislative mandate to provide such services, most agencies will receive little additional funding to support expansion of their service programs. The challenge is, therefore, to identify how children with developmental disabilities can benefit from existing child welfare services, and to identify ways to adapt or modify current services to better address these children's more specialized needs (Hughes & Rycus, 1983).

Inservice training is one intervention that can potentially increase the ability of child welfare professionals to serve children with disabilities within the context of their normal work activities. These professionals routinely conduct thorough case assessments, identify service needs, perform case management and service linkage, and provide supportive services in situations of child abuse and neglect. These same activities could be readily applied in situations where children have developmental disabilities if: 1) staff had the knowledge to recognize developmental disabilities in children on their caseloads; 2) staff understood the value of early intervention services; 3) staff understood their responsibility to serve
children with disabilities; 4) staff knew intervention strategies that could positively impact these children's lives; and, 5) staff knew how to access and coordinate specialized services in the community for these children.

Between 1980 and 1982, the Institute for Human Services developed and implemented an inservice training program for child welfare caseworkers throughout the state of Ohio to promote the acquisition of knowledge, skills, and attitudes related to developmental disabilities. The three-year program was funded by a grant from the Ohio Developmental Disabilities Planning Council. Three hundred twenty-seven child welfare caseworkers were trained in 19 groups in intensive, two-day training workshops. The curriculum was developed to promote the acquisition by caseworkers of the following knowledge and skills by trainees: 1) they would be knowledgeable regarding the most frequently encountered developmental disabilities; 2) they would be able to recognize signs of developmental disability in children on their caseloads; 3) they would be able to make proper referrals for developmental assessment; 4) they would be able to use assessment information to develop an appropriate service plan; 5) they would be able to work collaboratively with professionals in other agencies to advocate for coordinated service delivery; and, 6) they would be able to provide supportive
casework services to parents, foster parents, and other caregivers parenting children with developmental disabilities.

Training, by itself, cannot guarantee proficient staff performance. Many other variables can contribute to a lack of transfer of the learned knowledge and skills to the job, including system barriers such as lack of time, poor supervision, absence of agency policy and procedure to support good practice, lack of community resources, and a lack of cooperation among agencies; and, personal motivational barriers of staff who either do not value or choose not to perform required job tasks. Training as an intervention can only remediate "knowledge and skill" barriers to effective job performance. However, based on the previously cited literature that suggests a lack of training to be a significant contributor to poor practice, training can potentially be a very effective intervention in improving agency services.

The purpose of this study is to determine the impact of inservice training as an intervention to improve the ability of the child welfare system to meet the needs of children with developmental disabilities, even in the absence of wide scale and formalized system supports. The study is designed to identify whether inservice training leads to a permanent change in trainees' knowledge base, and whether behaviors on the job change
significantly as a result of the training. The study is longitudinal in nature; it will incorporate data gathered during the original training interventions between 1980-1982 and will use follow-up assessments of trainees who are still employed in the child welfare system. These staff will be evaluated to determine the degree to which they have integrated the knowledge and skills taught in the training workshop into their job activities. These staff members will then be compared to a group of matched controls who have not had formal training in developmental disabilities.

The hypotheses of this research are as follows:

1) Caseworkers who attended the two-day training workshop will score significantly higher on follow-up knowledge testing than will workers who have not had training in developmental disabilities.

2) Trained workers will demonstrate significantly higher numbers of criterion behaviors in their casework activities than will caseworkers who did not receive the training.
Chapter II
Review of the Literature

The preponderance of literature related to the evaluation of inservice training dates from the 1970's, when the field of human resource development was undergoing rapid development and expansion. Inservice training programs were being implemented to produce a variety of job performance changes in numerous corporations, businesses, factories, service agencies, and other organizations. A prevalent concern in the literature of this period was that very few of these training programs had been evaluated to determine their success in achieving their stated objectives. Without such evaluative data, it was believed that human resource development professionals could neither demonstrate the effectiveness of training to the organization (Wolfe, 1973; Smith, 1980a; Parker, 1976; Robinson, 1985; Goldstein, 1974), nor could they justify the continued existence of the training department in the face of budget cuts and other systemic pressures (Wehrenberg, 1983a; Alden, 1978; Wolfe, 1973; Smith, 1980a; Robinson, 1985; Parker, 1976; Goldstein, 1974).
However, as early as 1957, Goodacre suggested that while managers rarely made the effort to measure their program results, they expected a return from inservice training. He cautioned them to "take the initiative and evaluate their programs before the day of reckoning arrives" (Goodacre, 1957).

Approximately ten years later, Catalanello and Kirkpatrick conducted a study to determine the types of inservice training evaluations that were being performed in businesses and industry. 110 organizations responded to their survey questionnaire. The authors concluded that very few systematic and objective measurements of training were being performed, with most evaluation activities limited to an assessment of trainee reactions and feelings about the training program. Some organizations used pre and post-tests to measure learning. There were few attempts to measure behavior change, and those that did were "superficial" and "subjective." The authors concluded that the "state of the art" of training evaluation was indeed in its infancy (Catalanello & Kirkpatrick, 1968). Roy and Dolke's (1971) work substantiated that "in spite of many insightful admonitions," the number of controlled training evaluation studies in the literature was very small, and that most evaluation attempts focused on open-ended interviews of participants at the conclusion of each training session. Brown (1980) conducted a
survey to identify whether measurable improvements in employee performance could be attributed to training. He determined that fewer than half of the 285 companies in his study evaluated the results of general knowledge training; and only 12% had evaluated supervisory and management programs. Robinson (1984) suggested that most training departments evaluated their training using a criterion of activity level, including the number of programs conducted and number of trainees who attended. They typically did not measure results. McGehee and Thayer (1961) concluded a similar discussion by saying that "everybody talks about it, but nobody does anything."

Bunker and Cohen (1978) offered several potential reasons why training managers rarely performed evaluations of their training programs. These reasons included: 1) they believed that conducting evaluations was far too complicated, too time consuming and much too expensive; 2) they felt that evaluation of inservice training was impossible to do in the non-controlled environment of the organization; and 3) ultimately, trainers avoided evaluating because they often "didn't want to know" if their training was of poor quality. Wehrenberg (1983a) suggested that many management trainers failed to conduct training evaluations because they believed that "soft skills," which included most management and supervisory skills, could
not be empirically measured. Bell and Kerr (1987) determined that 90% of all surveyed trainers did not conduct evaluations because they were not required to do so by the firms for which they worked. Smith (1980a) concurred that most training professionals "didn't see the need." McGehee and Thayer (1961) suggested that management was typically reluctant to waste time testing something that was already believed to be of value. Wolfe's (1973) survey of training directors found that they doubted their own, and anyone else's, ability to conduct "meaningful" evaluations and doubted that the results from any evaluation within their ability to perform would be worth the time or expense.

Several authors criticize the methodology used for most training evaluations. Bunker and Cohen (1978) stated that evaluators were "frequently remiss in applying appropriate controls to draw reasonably valid conclusions." Alden (1978) believed that even motivated evaluators were not focused enough in their evaluation approach. Wolfe (1973) reviewed 21 of the "better known" training studies that had claimed to have performed rigorous evaluations and concluded, "they often fail to include many of the crucial elements needed to accomplish a valid and reliable evaluation." Smith (1980a) reviewed several reasons why training evaluations failed to meet high standards of quality.
These included: 1) the evaluator lacked evaluation skill, couldn't formulate objectives in precise measurable terms, made errors in choosing methodology, didn't know how to analyze data, and couldn't interpret the data once it was collected; and, 2) that obstacles in the organization, including poor course documentation, frequent changes in course content, and lack of cooperation from trainers and management, made methodologically sound evaluation impossible to achieve.

Authors who are proponents of a systems approach to training, in which the inservice training function is integral to the overall management of the organization, feel that the purpose of training evaluation is primarily to assure that inservice training activities help the organization meet its goals and objectives (Phillips, 1983; Goldstein, 1974; Robinson, 1984; Brethower & Rummler, 1979; Wehrenberg, 1983a; Blumenfeld & Holland, 1971; Rose, 1968; Tosti, 1979; Holoviak, 1982); and if it does not, it is "doomed to failure" (Dubin, Mezack & Neidig, 1974).

Parker (1976) contends that the purpose of training evaluation is to determine how well the training objectives were achieved. Blumenfeld and Holland (1971) concur that prespecification of training objectives is a critical step in assuring the appropriateness of the evaluation. Both authors suggest that the
training evaluation must be planned concurrently with the planning of the training program, not after the training has been completed.

The literature suggests several appropriate goals and objectives for inservice training programs. Blumenfeld and Holland (1971) state that the goal of inservice training is to bring about changes in trainees' attitudes, skill, and knowledge. They also suggest that it can be an effective tool in attaining management goals. Rose (1968) states that the goal of all employee training is "to develop the abilities of the work force so that the functions of the organization are performed expertly and at a minimum cost." Mikesell, Wilson and Lawther (1975) state that the purpose of training is to make major job performance improvements in a reasonably short period of time. Robinson (1984) states that an appropriate function of the human resource development professional is to assess the work environment and identify variables that may be interfering with job performance. Through this approach it can be determined whether the identified problems can be appropriately addressed by training, or whether some other management intervention is more appropriate. This step prevents the expenditure of training resources to solve problems that may not be training problems.
There is a unifying theme to all these discussions; that is, training should ultimately benefit the organization. Inservice training activities should be incorporated into the total strategy for management of the organization. The evaluation of training programs is "only a phase in the cycle" (Dubin, Mezack & Neidig, 1974). Evaluation cannot, therefore, be planned or implemented in isolation if it is to be effective (Brethower & Rummler, 1979). Linking inservice training to the goals and objectives of the organization is a multi-step process that is described extensively in the training and development literature (Goldstein, 1974; Wehrenberg, 1983a; Rose, 1968; Phillips, 1983; Parker, 1976; Holoviak, 1982; Tosti, 1979; Brethower & Rummler, 1979; Dubin, Mezack & Neidig, 1974). Despite differences in terminology, there is relative concurrence regarding the proper sequence of activities. Prior to the development of any training program, the organization's training needs must be assessed. Goldstein (1974) recommends that first one must perform a thorough organizational analysis, which includes an examination by management of both long- and short-term organizational goals and expectations. This step insures that the training accurately reflects changes in the organization and does not inadvertently "train for obsolescence." This
assessment should also determine the availability of resources, both to conduct the training and to support the performance of newly trained skills by staff on their jobs.

The next step is to conduct a task analysis of the trainees' jobs, which will result in a listing of the job-specific behaviors that trainees must be able to demonstrate upon completion of the training. Tosti (1979) refers to a survey that was conducted to assess the relevance of U.S. Army inservice training programs. This study demonstrated that 50% of the critical skills used on many jobs were never formally taught, and worse, 60% of the skills that were taught in training were not necessary for job performance, or could easily have been learned informally on the job. This study is a classic example of how failure to link training content to the identified tasks of the trainees' jobs can result in largely irrelevant training.

Goldstein (1974) suggests conducting a "person analysis", or an assessment of each employee's existing skills to identify staff members who do not have the required job skills. This will determine the "targets," or the specific staff members who should receive the training. Further, the behavioral objectives to be achieved by trainees upon completion of training must be specified. These behavioral objectives are used in designing the
training, and they are also the criteria against which to measure the training's success. The training evaluation is the step that determines the degree to which the objectives were met.

Dubin, Mezack and Neidig (1974) stress the importance of achieving "buy-in" from upper level management, before any training needs assessment is conducted or any training program is planned, to assure that management will support the transfer of learning from training to the job setting.

Holoviak (1982) includes post-training activities in his five-step approach. He recommends that "feedback loops" be created following the training to ensure continued top management support and commitment; and, that follow-up interventions be provided, including supervised practice, coaching, counseling, and performance appraisal, to promote the use of newly taught skills in the work place.

Robinson (1984) stresses that positive training results will only occur as "the product of a learning experience that is 'on-target and well delivered,' and a work environment that supports the new skills. Wehrenberg (1983a) stresses that if training "fails," it is often because training managers failed to conduct an accurate needs assessment and trained "the wrong skill."
As has been suggested, training evaluation has a clearly defined purpose when it is viewed within a systems framework; that is, to systematically measure and determine the degree to which the prespecified objectives of the training program were met. There appears to be a wide variety of opinion, however, regarding what constitutes an appropriate training objective, and therefore, what exactly the training evaluation should measure.

Blumenfeld and Holland (1971) believe that training programs should bring about desired changes in employee attitudes, skill, knowledge, and behavior which, when integrated by trainees on their jobs, will directly promote the achievement of organizational goals and objectives. These changes can, therefore, be considered appropriate training objectives and should be measured by the evaluation. Rose (1968) believes that training evaluation should properly include a determination of "whether or not the objectives and content of training courses are consistent with the mission and current needs of the organization." Smith (1987) defines productivity, quality, and other "work force factors" as appropriate training objectives.

Several studies have attempted to determine the effects of training on increasing the quantity of worker output in a
designated period of time (Latham & Kinne, 1974; Lefkowitz, 1970; Holoviak, 1982); reducing errors or defects in production (Chaney & Teel, 1967); reducing employee turnover (Lefkowitz, 1970; Holoviak, 1982; Latham & Kinne, 1974); reducing absenteeism and injuries (Latham & Kinne, 1974); increasing employee motivation, and reducing grievance rates (Holoviak, 1982).

A wide variety of potential outcomes have been formulated into training objectives. Goldstein (1974) suggests that this is to be expected, as "there are few, if any, single measures that can adequately reflect the complexity of most training programs." The task becomes more difficult when training is designed to meet organizational goals and objectives because of the very complex nature of most organizational goals.

Considerable discussion in the literature has focused on what constitutes appropriate evaluation criteria. Blumenfeld and Holland (1971) believe that "criterion is the single most important facet of any serious study." Improperly formulated criteria are often implicated in the poor quality of many training evaluations. Brethower and Rummler (1979) contend that until people can agree on what they are trying to evaluate, they won't ever be able to agree on how to go about it.
As early as 1957, Goodacre stressed that properly formulated criteria were critical to assure the accuracy of the training evaluation. For Goodacre, criteria were the quantifiable indexes of the changes in behavior that the training program was designed to accomplish. He envisioned that training criteria could potentially exist at several different levels: immediate criteria, which were success outcomes directly following the training; ultimate criteria, which were the final desired outcomes of training for the organization; and, intermediate criteria, which he defined as "something inbetween." He suggested that the ultimate criteria might not be reachable, or measurable. The nebulous "intermediate" criteria were identified as potentially the best criteria to be used to evaluate training (Goodacre, 1957). Despite the lack of clarity in his definitions, Goodacre's contentions that any criteria used in training evaluation should lend itself to quantification, and must be both reliable and relevant, were entirely valid.

Kirkpatrick has published many articles over a 30-year period delineating a four-level framework for classifying training evaluation criteria (Kirkpatrick, 1959; 1978; 1979; 1987). Kirkpatrick's formulation has become the "industry standard." It is the most well-known, most often quoted, and most widely
used system. It has, on occasion, been imitated by other authors (Brethower & Rummler, 1977; Parker, 1973; Jackson & Kulp, 1979; Del Gaizo, 1984; Alden, 1978); however, while these reformulations change terminology or style, they appear retain the same organizing concepts.

Kirkpatrick (1959; 1978; 1979; 1987) defines his four levels as follows: 1) Level I: Reaction - At this level, success criteria include how participants feel about the training program, that is, how satisfied they are with the training they have received; 2) Level II: Learning - Level two criteria include what knowledge and/or skills were acquired by the trainees during the course of the training; 3) Level III: Behavior - Level three criteria consist of the degree to which trainees' on-the-job behavior changes as a result of the training program. 4) Level IV: Results - Level four criteria include the final outcomes for the organization produced by the training. In describing the need for Level IV criteria, Kirkpatrick explains that a change in employee behavior cannot be considered the final objective training was trying to achieve; rather, it is only a means to an end. The end results should include such things as improved productivity, better quality, lower costs, reduced accidents, increased profits, or better services.
Newstrom (1978) explores several prevalent assumptions regarding Kirkpatrick's model. The first is that the criteria are arranged in increasing order of the value of the information to be gained. Thus, behavior-based (Level III) and results-based (Level IV) evaluation data will have far greater validity than measurement of reaction or learning. Level I criteria, participant reactions, are therefore the least valid. The second assumption is that the Level I criterion, reaction, is used most often in evaluating training programs. Newstrom believes this is predictable, since it is the easiest type of measure to obtain. The third assumption is that there should be congruence between the four levels; that is, if trainee satisfaction is high, then level of learning should also be high; trainees should be more able to apply the learning to the job; and, the end results for the organization should be better.

Most authors agree that reaction is the least valid type of evaluation data. Robinson (1985) says that reaction data is "totally subjective" and is often based on whether the trainee enjoyed the training. He claims there are "serious weaknesses in ad hoc (anecdotal) feedback." Rose (1968) states that participant response questionnaires as a single evaluation method, when not used in combination with other methods, has limited value for evaluating the results of training. Wolfe (1973)
strongly prefers measuring both behavior change and "organizationally meaningful operating improvement," which are more consistent with Kirkpatrick's 3rd and 4th levels. Robinson (1984) agrees that accountability for training should focus on the results achieved by training, both on-the-job behavior change and the organizational impact of the training. She suggests that "the HRD profession...is moving toward results-oriented training as a standard." Blumenfeld and Holland (1971) consider participant questionnaires to be "one of the most popular and least meaningful criterion measures" that rate factors such as course structure, content, and quality of instruction. They contend a high rate of use because training managers "find this type of evaluation easy to administer and painless to analyze." And, Bunker and Cohen (1978) state that in order for measurements of the effects of training to be valid, one should focus on job behaviors and organizational consequences.

Kirkpatrick (1979) however, believes that participant response data can be useful, particularly if the assessment instrument is properly formulated and tabulated, and accurate feedback is encouraged. He contends that trainee input can be very useful in rating the quality of the conference leader or trainer, and that immediate feedback from trainees can lead to improvement of the training program itself. Kohn and Parker (1969) agree that even
though measuring reaction relies on data that represent the opinions and judgments of respondents, "there is no reason to believe that participant views are not sound data."

The apparent disagreement regarding the usefulness of participant response evaluations may result from a lack of congruence regarding the objectives for the evaluation. A distinction between formative and summative evaluation criteria, originally proposed by Scriven (1967), and referenced by Thompson (1978) and Goldstein (1974), may help to clarify the issue. A formative evaluation is a process-oriented evaluation that is best used to assess the quality of the training event itself. Criteria for a formative evaluation might include the trainer's teaching ability, the degree to which the training holds the trainees' interest, the trainer's ability to relate the course content to actual practice, the trainer's knowledge of the content, and the level of comfort in the physical training setting. All these variables contribute toward a climate that promotes learning (Knowles, 1970). A summative evaluation is results-oriented that assesses the outcomes of the training. Thompson (1978) suggests that summative evaluations are typically preferred by researchers, who seek data to validate changes as described in the training objectives. When considered in this...
context, Kirkpatrick's Level I – reaction criteria may potentially be the most appropriate approach to use in a formative assessment of a training event.

Thompson (1978) suggests that many evaluators fail to accurately determine to what end they are conducting the evaluation, and that this contributes to their difficulty in formulating proper evaluation criteria. Wolfe (1973) provides data from a survey of training directors indicating that while all concurred regarding the need to evaluate, there was "disagreement over the focus of the evaluation." Newstrom (1978) has listed several potential purposes for evaluating training; 1) to assess whether training objectives have been achieved; 2) to assess the effectiveness of the trainer; 3) to justify the expense of training through a cost-benefit analysis; 4) to improve the program's content or structure; 5) to decide whether other trainees should receive the program; 6) to identify which trainees benefited the most and the least from the program; 7) to reinforce major points for the trainees; and, 8) to create advance expectations in the minds of trainees (by using pre/post test assessments.) He suggests that the evaluation criteria and the evaluation methodology will change depending upon the purpose the evaluation is to serve.
There is greater concurrence regarding Kirkpatrick's second, third, and fourth levels of criteria. There appears to be little difficulty in defining the level 2 criteria, "learning." Learning is the demonstrated difference in knowledge, skills, and attitudes before and after the training (Kirkpatrick, 1978). Learning is also defined as a relatively permanent behavior change that results from the training experience or from practice (Blumenfeld & Holland, 1971; Blumenfeld & Crane, 1973). Several training evaluations have attempted to measure learning. In Catalanello and Kirkpatrick's survey, they noted that of those organizations that had conducted evaluations, a large percentage did so by measuring learning (Catalanello & Kirkpatrick, 1968). The methodology to measure learning is relatively simple; generally, pre and post-test measurements of knowledge and skill can determine, with some degree of accuracy, whether learning has indeed occurred as a result of the training (Roy & Dolke, 1971; Kohn & Parker, 1969).

Kirkpatrick's third level, the measurement of behavior, is the preferred measurement criteria for many authors. Gustafson (1977) suggests that the "bottom line of training effectiveness is reflected in employee proficiency." Robinson (1985) claims that one of the best criteria for valid training assessment is the ability to relate the results to job performance. Cullen, Sawzin,
Sisson and Swanson (1976) used "worker competence" as their desired outcome, which they quantified by identifying four predetermined worker abilities and behaviors. Mikesell, Wilson and Lawther (1975) believe that training objectives should be written to include specific behavioral changes of trainees. Wehrenberg (1983a) states that the goal of training is improved job performance and that even "soft skills" must be delineated into behavioral performance variables; the "acid test" of effective training is a measure of trainee performance. Gallegos and Phelan (1974) state that training objectives should be statements which specify learner behavior, and that the training methodology should be designed to cause pre-specified changes of behavior, or acquisition of new behavior, in the learner. Robinson (1985) says that measurement of behavioral outcomes must take place on the job, not in the training environment, if the evaluation is to determine whether the training had the desired effects on job performance. The required evaluation methodology is not only more complex, but the evaluation should occur several months after completion of the training to assure that adequate time has passed to permit transfer of learning to the job (Kirkpatrick, 1979).

Measurement of results (Level IV) is considered to be the most valid effectiveness criteria. Wolfe (1973) claims that the
"ultimate test of training effectiveness (is) an improvement in an indicator of operational performance." He suggests that "trainers must not be content with behavior change only, for change must be translated into something meaningful to the organization. Behavior change without operational improvement is a barren adventure." However, measurement at Level IV is also believed to be the most difficult methodologically. Results criteria are the most elusive and the most easily affected by uncontrolled variables. The literature supports Goodacre's (1957) contention that this "ultimate" criteria might not be realistically accessible. Kirkpatrick (1978) suggests that "there are so many factors that influence results that it is frequently impossible to prove that the training program caused the desired result, and no way that tangible results can be directly related to the training program." Alden (1980) attests to the inadequacy of "results" because variables other than employee performance can affect the final product, and a product might prove to be acceptable in spite of an imperfect performance. Bell and Kerr (1987) agree that measuring the bottom line effectiveness of training programs is a difficult and elusive task, particularly in organizations where the final result is a service, since the the product is intangible and the assignment of dollar value to the product is difficult.
Mikesell, Wilson and Lawther (1975) agree that evaluation of results is made somewhat easier when dollar amounts can be attached to performance.

Returning to Newstrom's (1978) three assumptions about Kirkpatrick's model, the third assumption was that there is generally congruence between the four levels of criteria. Newstrom asserts that this assumption is clearly false, and his contention is supported in the literature. Critics of Level I criteria, participant reaction, believe its lack of utility is precisely that it does not reliably provide any information regarding any of the other three, more valid, outcome criteria. The "transfer of learning" literature, to be discussed shortly, also stresses that knowledge and skills acquired as a result of training do not routinely generalize to the job setting. The determination of the purpose of the evaluation is therefore critical prior to selecting either the measurement criteria or the evaluation methodology.

Goldstein's (1974) previously mentioned belief that "there are few, if any, single measures that can adequately reflect the complexity of most training programs," summarizes the difficulty faced by most trainers in choosing their evaluation criteria. Because of this complexity, Goldstein believes that thorough
training assessment must be multi-dimensional and must concurrently measure several different criteria. Kirkpatrick (1959) also suggests that evaluation procedures should consider all four levels of criteria. This multi-criteria approach to training evaluation is exemplified by Bell and Kerr (1987), who describe the evaluation of a training program designed to improve formal communication skills (writing, speaking and dictating). The evaluation methodology utilized assessment at all four levels. An end-of-course questionnaire was used to measure trainee reaction; pre and post test case study assessments were used to measure the acquisition of communication skills; a follow-up survey was completed by course participants between 3 and 23 months after completion of the program wherein trainees reported how they had incorporated the new skills on their jobs; and finally, the employee and supervisor jointly completed a survey that identified predetermined organizational outcome measures, all of which were believed to contribute to more cost effective utilization of personnel and resources.

It is understood that developing criteria that meet stringent experimental research requirements is likely beyond the scope of training evaluation (Smith, 1979). The evaluation of inservice training is not, nor should it be, controlled laboratory research (Brethower & Rummler, 1979). However, since the use of "bad" or
"invalid" criteria can render the evaluation meaningless, some attention must be given to the process of criteria formulation.

Goldstein (1974) suggests that measurement criteria must be relevant. He refers to Thorndike's (1949) definition of relevance of criteria for inservice training purposes; the knowledge, skills, and attitudes that are required for a trainee to "succeed" in the training program must be the same as those required to "succeed" at the ultimate task, that is, the job. This is also referred to by some authors as the "content validity" of measurement criteria, particularly when used in reference to test items that measure the acquisition of knowledge and skill from the training (Roy & Dolke, 1971; Kirkpatrick, 1976). Smith (1979) also suggests other types of "relevance" of a criterion including its concurrent validity. He defines concurrent validity as the power of a criterion to discriminate between groups of people known to differ in job proficiency. If a criterion cannot accurately make this differentiation, it should not be used to evaluate the relative effectiveness of training in achieving a desired level of job performance. Goldstein (1974) adds that properly formulated evaluation criteria are also comprehensive; that is, there are few components in the "ultimate" criteria that are not also present in the measurement criteria; nor do extraneous elements
in the measurement criteria (i.e. elements that are not present in the "ultimate" criteria) contaminate the measure. Smith (1979) cautions that criteria must also be free from known biases, which he defines as extraneous, uncontrolled situational variables that may influence criterion performance.

Two types of criterion measures are also discussed in the training literature: norm-referenced measures, which compare the capabilities of an individual to those of other individuals, and criterion-referenced measures, which provide a standard of achievement by using specific behavioral objectives (Goldstein, 1974). Most of the previously discussed literature advocates the use of criterion-referenced measures for the evaluation of inservice training, written as specific behavioral objectives for performance.

There is clear consensus in the literature that it is within the ability of most training departments to conduct methodologically sound evaluations of their training programs. (Brown, 1980; Brethower & Rummler, 1977 & 1979; Wolfe, 1973; Wehrenberg, 1983a & 1983b; Rose, 1968; Fuqua, 1979; Bunker & Cohen, 1978; Smith, 1978a; Kirkpatrick, 1978; Blumenfeld & Holland, 1971). However, the evaluation of inservice training programs does
present some unique problems that may necessitate modifications in both evaluation methodology and in the researcher's expectations for results.

Brethower and Rummler (1979) refer to the evaluation of inservice training as "action" research, because it should take place in the field setting of the organization rather than the laboratory. As a result, the researcher is likely to encounter "all the constraints that are imposed by an organization's going about getting its work done." The authors contend, however, that training evaluation must be conducted within the organization rather than in a laboratory, particularly when Kirkpatrick's 3rd and 4th level criteria, job performance and organizational impact, are the dependent variables. They suggest that the evaluation will probably not be as clear-cut as laboratory research would be, but that this should not suggest that evaluation cannot be valid.

Robinson (1985) concurs that results of training that are measured in the laboratory may not generalize to the workplace; the trainee's behavior in the "low risk" training environment might be more difficult to implement in the job setting, where conditions might be more inhibiting.
Several types of evaluation methodologies are described and discussed in the literature. Thompson (1978) states that an evaluation methodology cannot be chosen based on its own inherent value, but must be selected because it most appropriately achieves the purpose of the evaluation.

Fuqua (1979) and Smith (1987) recommend the use of the equivalent comparison group design as the preferred method of evaluating inservice training. In this design, subjects are randomly assigned to a treatment or a comparison group. The performance of both groups is measured following the program. Because of the random assignment of subjects, if the training is effective, the trained group should out-perform the control group on follow-up measures of learning or behavior change (Smith, 1987). Fuqua (1979) claims that the complete random assignment of subjects to groups in this model provides "the tremendous advantage of assuming that the groups so assigned do not differ significantly from one another prior to the intervention." This controls for several potential threats to validity that are common with a single group pre- and post-test model. Wehrenberg (1983b) concurs that the ideal training evaluation should use both a control group and pre- and post-test measures, with a post-test immediately following the training.
and repeated again several months later. Smith (1978) used methodology that included the random assignments of trainees to experimental and control groups and pre- and post-test measures to train supervisors in management strategies to promote employee attendance.

Blumenfeld and Holland (1971) also consider the use of pre- and post-test measures along with a control group to be the minimally appropriate experimental design for the evaluation of training. They do not specify whether they prefer random assignment of subjects or a matching procedure, but simply state that the groups must be "comparable."

Wolfe (1973) concurs that the equivalent comparison group model is the ideal model for the evaluation of training. However, he recognizes that in an organizational training situation, random assignment of trainees to groups is not always possible or desirable. As an alternative, Wolfe recommends that the controls be matched to the experimental subjects "in every relevant and distinguishable aspect." In his review of 21 of the "better known" training evaluation studies, he claimed that the best ones methodologically featured before and after measures on both control and experimental groups, using standardized measuring devices. Wehrenberg (1983b) suggests that this practice is
particularly important where trainees volunteer to attend training, since the characteristics of the volunteers, including level of interest, willingness to learn, or previous familiarity with the subject matter, might initially be very different from non-volunteers.

Roy and Dolke (1971) also used this methodology in their evaluation of a supervisory training program. A control group, matched for education, experience, age, and related characteristics, was tested along with the experimental group both before and after the training. Post-testing was performed both immediately after the training and nine months later. The difference between the before and after subject matter test scores was considered the measure of learning. Follow-up testing nine months later was considered the measure of retention. Kirkpatrick (1978) suggests that a performance test be used to measure the ability to perform criterion behaviors before and after the training.

Brown (1980) agrees that the use of control groups "does provide a much more convincing demonstration of cause and effect." However, he identifies several liabilities to the control group method, including higher costs, difficulties in matching experimental subjects and controls, and problems holding other
variables constant. Brethower and Rummler (1979) suggest that control groups should be used only when the experimenter has adequate time to ensure the comparability of the two groups. They feel that it is difficult to find two naturally occurring groups in an organization that "just happen to be comparable on relevant dimensions." Even if this is possible, they claim that the group members' experiences outside the organization may be dissimilar enough that training may not be the only variable affecting performance. Brethower and Rummler strongly suggest using methodologies that do not require the use of control groups.

Pre and post-testing also has its proponents and opponents. Kirkpatrick (1978) recommends pre and post-testing as the most appropriate methodology to measure Level 2 criteria, learning. The cautions regarding pre- and post-testing focus on assuring the validity of the test instrument. Roy and Dolke (1971), Kirkpatrick (1976) and Phillips (1983) all stress that the test instrument must have a high level of content validity, that is, it must include a representative sample of the skills, knowledge, or abilities that are presented in the training program. Similarly, the test should also have concurrent validity; it should be able to differentiate accurately between groups with respect to their job competence. Rose (1968) claims that a measure of mastery of information is not always adequate, even if it is valid.
Knowledge, by itself, is not an adequate result when our goal is to change job performance. He suggests that a properly structured written test should require the trainee to apply and use the information, not simply to recall the facts.

Fuqua (1979) raises issue with the use of pre- and post-test instruments and reviews several variables independent of the training intervention that could potentially account for the change in test scores. These include: 1) memory and practice effects might occur at the post-test as a direct result of the pre-test experience; 2) the presence of certain items on the pre-test could sensitize trainees to particular information; and 3) external variables may cause the scores to change naturally over time. Brethower and Rummler (1977) suggest that before and after measures be used only as "a last resort" when measuring behavioral changes on the job, since multiple variables other than the training can account for job performance differences.

Quasi-experimental designs, sometimes referred to "non-equivalent comparison group" designs, are preferred by many authors to the more classic equivalent comparison group designs for the evaluation of inservice training. Smith (1987) claims that while "true" experiments can produce less equivocal data, they are often impractical in field settings where it is more
difficult to conduct random assignment to groups or to withhold training from control groups. Fuqua (1979) cautions that when using non-equivalent comparison group designs, it is essential to eliminate the effects of differences in the training groups that could potentially confound the variables under investigation.

Several authors, among them Brown (1980), Brethower and Rummler (1977), Campbell and Stanley (1963), and Smith (1987) recommend the use of multiple baseline designs. Brown suggests that there are three variations of this methodology that are ideally suited for evaluating training interventions: 1) multiple baseline across behaviors or performance variables, which includes evaluating the effects of a single intervention on two or more distinctly different employee behaviors or performance indices; 2) multiple baseline across subjects or groups, wherein one would determine the effects of an intervention on one performance variable in two or more classifications or groups of trainees; and, 3) multiple baseline across settings, in which one would determine the effects of the intervention on one performance variable in settings which operate independently of one another. Brethower and Rummler's support of multiple baseline designs is based upon the ability of this methodology to control for extraneous environmental factors.
Brown (1980) suggests that "under no circumstances should a reversal design be used to evaluate training." He suggests that there are ethical issues in purposefully withholding or removing training. He also questions whether the training intervention can legitimately be "removed," since the result of training, that is, learning, is defined as a more or less permanent change in knowledge and skill. Brethower and Rummler (1979) cite more pragmatic concerns, that management would likely be unwilling to discontinue an intervention they believe has caused improved performance. They suggest that opportunities to use the reversal design might occur "naturally" in the environment and should be recognized as an opportunity to conduct evaluation research.

There is limited discussion in the literature regarding the use of participant input in the evaluation of criteria other than response. There appear to be mixed opinions regarding the use of "self report" as a valid source of data. Rose (1968) indicates that personal interviews with trainees can produce information that is not available by other methods. He suggests that effectively structured interviews can help trainees freely express their opinions of the training and its value and focus attention on specific problems with the training. Kohn and Parker (1969) used 1 1/2 hour long semi-structured personal interviews with trainees several years after completion of the training to
determine how they had been able to apply the course material on their jobs. The authors believe there is "no a priori reason to believe that the views of participants aren't sound data...the important consideration is whether (the measure) has been carefully devised, and whether it is quantifiable so that its relationship to other variables may be established."

Fast (1974) recommends that trainees be involved in setting the training objectives. Each trainee should identify personal objectives prior to the beginning of the training, and then should measure the achievement of those objectives upon completion of the training. Fast suggests that "the clarification of goals actually helps some participants get more out of the course than they would otherwise."

Reeves and Jensen (1972) suggest that participants may be in the best position to judge the applicability of a program to their own work environment. The authors sent follow-up questionnaires to trainees between 6 months and a year after a management training program to determine whether in the trainee's estimation, they had experienced any significant behavioral change as a result of the training, and whether they were using the techniques they had learned. Kirkpatrick (1976) used structured interviews with trainees 2-3 months following a
supervisory training institute to obtain information regarding on
the job behavior changes that occurred as a result of the
training. They corroborated the trainees' self-assessment data
through interviews with immediate supervisors.

One final issue to be considered in the evaluation of training is
the way in which "transfer of learning" problems can confound the
interpretation of evaluation results. Even in situations where
there were no measurable changes in employees' behavior on the
job as a result of the training, it may be inaccurate to conclude
that the training was ineffective in teaching the desired skills.
Variables other than knowledge and skill are known to affect job
performance, and even a highly successful training program
cannot, by itself, assure that trainees will use their newly
acquired skills on the job.

Kirkpatrick (1976) delineates five prerequisites to competent job
performance: 1) the employee's desire to perform; 2) the ability
to perform (which includes the knowledge and skills that are
transmitted through training); 3) the job climate must not
prevent or discourage implementation of new skills; 4)
assistance must be available to trainees in the work environment
to apply newly acquired skills on the job; and 5) rewards must be
provided to trainees for changing behavior. Several other authors
support Kirkpatrick's contention. Brethower and Rummler (1977) suggest that if a training program "passes" at levels 1 and 2 (response and learning) and "fails" at levels 3 and 4 (behavior and results), the failure is "an indictment of management, not of the training." Roy and Dolke (1971) state that retention of acquired knowledge and skill depends, in large part, upon the trainees being given opportunities to practice on the job what they acquired in the training program. Robinson (1985) suggests that training usually occurs in artificial conditions that cannot accurately reflect the reality of the job situation, and that conditions in the work environment might be inherently more inhibiting than in the training environment.

Wehrenberg (1983 a & b) suggests two possible causes for "failed" training other than the failure of the training program itself: the work environment has thwarted the use of the new skills, and an accurate needs assessment was not conducted prior to planning the training, and therefore, trainees were trained in the "wrong skill." Blumenfeld and Holland (1971) concur that the training may not be effectively transferred to the job if the proper prespecification of training objectives was not determined, even if the training program itself was effective.
Robinson (1984) suggests that the role of the HRD professional includes identifying variables in the work environment that may impede change on the job and communicating this information to managers to promote necessary changes to support the integration of newly acquired skills. Without this intervention, she contends that long-term results from training will be impossible to achieve.

Liefer and Newstrom (1980) identify several steps in the planning and delivery of training that can help to overcome the transfer of learning problem. They propose the following: 1) trainees must first be aware of the work situations in which the new skill is most appropriately used; 2) trainees should enter into a learning contract to acquire specific skills; these then become personal learning objectives; 3) trainees must be provided with the opportunity during the training to plan how they are going to implement the skills on the job, including identifying potential obstacles; 4) trainees must be encouraged, must receive ongoing feedback, and must be positively reinforced on the job while they are trying to implement the skills. This last point underscores the importance of the active involvement of the trainee's direct supervisor in transfer of learning. This concept is strongly supported by Smith (1987 & 1978a), Robinson (1985), Goldstein (1974) and Rose (1968).
A survey of training evaluation studies in the fields of mental retardation/developmental disabilities and in child welfare provide ample evidence that the "general" concepts cited in the training and development literature are equally applicable to the specialized training situations in these fields of practice.

Ziarnik and Bernstein (1981) reviewed the training literature to determine the effects of inservice training on the performance of direct care staff for persons with developmental disabilities. They noted that most of the inservice training programs they reviewed produced equivocal results in effecting behavioral changes in staff, in demonstrating secondary effects on client behavior, and in maintaining or generalizing performance changes. They suggest that the "effectiveness of staff training has yet to be demonstrated because it is often incorrectly applied." They support the systems view of training, stressing that staff are too often referred to training for "performance problems" that may have nothing to do with knowledge and skill deficiency. They suggest that "no amount of inservice training will remediate the problem" if the origin of the problem is not a need for training. They also comment on the "transfer of learning" problem by
suggesting that "no matter how skilled staff are, the environment must support the exhibition of those skills for performance to be satisfactory."

Salisbury (1984) stresses the importance of a "systems" approach to training in her report on training for respite care providers for persons with developmental disabilities. She outlines training strategies that can "effect sustained, generalized performance changes in staff and clients." She outlines a nine-step schema that she recommends for respite care provider training that includes conducting an ecological inventory of the administrative system and job environment; assuring that training content is functional and relevant for job performance; assessing individual trainee performance; formulating individualized training plans based upon individual needs; and monitoring and evaluating trainee performance using multiple objective measures.

Templeman, Fredericks, Bunse, and Moses (1983) described a model to train special education teachers. Their "systems" approach to training is evident in their presentation of "critical, non-negotiable" items to be included in the development of training: 1) the training environment must be designed to most closely replicate the work environment; 2) the training must have specified instructional objectives; 3) trainees must have
defined criterion levels of acceptable performance for each training objective; and, 4) training must be followed with technical assistance, further training, and reinforcement in the trainee's work environment.

Knowles and Landesman (1986) conducted a national survey of state-sponsored training for residential direct-care staff. In their concluding remarks they stress the importance of a "systems perspective" to guarantee that training programs will be maximally beneficial. This includes a study of the total ecology of the residential facility; assuring that training objectives and curriculum are consistent with the goals and operating policies of the residence; assessing the probability that staff members will be able to apply newly acquired knowledge and skills; and determining what natural reinforcers are available in the work environment to help staff maintain positive attitudes and behaviors. Demchak (1987) reported on the results of behavioral staff training in special education settings. She identified the value of training programs that use methodologies to promote both the acquisition and transfer of skills, such as modeling and role play. She stressed, however, that even a successful training program may prove to be "ineffective unless the training is followed by opportunities for reinforcement and feedback that intend to maintain desired
performance." When feedback to trainees is paired with supervisor approval, it appears to be effective in increasing staff performance.

Page, Christian, Iwata, Reid, Crow, and Dorsey (1981) successfully demonstrated the effective use of a multiple baseline training design to evaluate the training of interdisciplinary teams in writing Individual Program Plan (IPP) goals and objectives. The training took place in a residential facility for approximately 90 persons with developmental disabilities. Both baseline and post-training data were collected to enable determination of both "the true extent of deficient areas" prior to the training and the generalization and maintenance of skills after the training.

Inservice training designed to improve job performance has been reported in many areas of child welfare practice, among them the acquisition of values and attitudes conducive to good practice (Ammons, 1979); improvement in listening skills and empathy (Gelfand, Starak, & Nevidon, 1973); improving the behavior management and parenting skills of foster parents (Guerney, 1976); permanency planning skills (Jones & Biesecker, 1980); treatment skills of residential care staff (Kagan, 1983); interagency collaboration skills (Krause, 1977; White, Cornely, &
Gately, 1978); ethnic competence and racial sensitivity (Montalvo, Lasater & Valdez, 1982; Kautz, 1976); stress management and burnout (Shannon & Saleebey, 1980); and basic skills for new child protective service workers (Rosick, 1979).

A review of the above cited studies suggests that the "state of the art" in the evaluation of inservice training in child welfare appears to be consistent with that in the training field in general. There is a high preponderance of Level I - response evaluation, and some attempt to measure learning as the result of training using standardized pre- and post-test measures, with occasional statistical validation of the significance of test scores. The fact that there are no measures at Level 4 is expectable, since "results" in the human services field are almost always defined in terms of positive changes in the quality of life of persons receiving the service; and therefore, are exceptionally difficult to quantify. However, the failure to consistently measure changes in job skills and performance subsequent to training is most likely a failure to properly plan training from a systems perspective and to include evaluation in the original training methodology.

A review of the child welfare literature indicates that a lack of training is often believed to contribute to poor services to
children with developmental disabilities. Coyne and Brown (1985, 1986) suggest that caseworkers' lack of accurate information about developmental disabilities interferes with the proper placement of seriously handicapped children in adoptive families. The authors state that children with disabilities are perceived by their workers as especially hard to place for adoption. This is based on an inaccurate assumption that children with handicaps are extremely hard to parent, are less desirable to potential adoptive parents, and there is a greater likelihood of placement disruption. While this has been discounted by research data, the authors felt that the prevalence of misinformation among staff interfered with the successful adoptive placement of many children with disabilities.

From a review of the literature as well as their own survey findings, Schilling, Kirkham, and Schinke (1986) concluded that the child protection system may be failing to recognize and document children with handicaps in populations of abused and neglected children. They suggest that child welfare workers are "not attuned to the needs of children with handicaps" and don't typically understand "the special risks of developmentally disabled children." The authors claim that the lack of training is a significant contributor, as "few social workers, even those with graduate degrees, have had developmental disabilities courses."
Many workers are believed to lack the knowledge and skills to assess children's handicaps, particularly milder developmental delays.

A survey by Richardson and associates suggested that ignorance on the part of staff members in one agency about how to use services in another agency created barriers to the accessibility of services for children with developmental disabilities (Richardson, West, Day & Stuart, 1989).

Kurtz (1979) strongly suggests that lack of knowledge and skill of child welfare workers is a serious impediment to the identification of children with developmental disabilities and the provision of early intervention services. While he feels that child welfare workers have a general understanding of child development, they lack the specific knowledge and skills to make judgments about early signs of disabling conditions or handicaps. He suggests that this lack of training also contributes to children with disabilities being underrepresented in child welfare statistics.
The literature focuses on three program areas in which knowledge and skills regarding developmental disabilities are particularly important for child welfare professionals: foster care, adoption, and screening and early intervention.

Barsh, Moore, and Hamerlynck (1983) state that a proportionately higher percentage of children with handicapping conditions are placed in foster care than would be expected from their percentage in the general population. Arkava and Mueller (1978) indicate that many specialized foster care systems have been developed to assure that children with handicaps are cared for in the most homelike, least restrictive setting. Richardson and her associates (1989) concur that federal law and regulation have resulted in the creation of community based living alternatives for children to prevent their placement in institutions. This has resulted in an exponential growth in both regular agency and specialized foster care. (Richardson, West, Day, & Stuart, 1989; Hill, 1987.)

These articles identify a number of skills that are needed by foster parents to properly care for children with disabilities. These include: skills to provide direct care of children with handicaps, skills to deal with behavior problems; and skills to promote development (Barsh, Moore, & Hamerlynck, 1983); skills
in observing children's behavior, skills in setting appropriate goals, the ability to teach new skills to children, and the ability to properly reward appropriate behaviors (Dickerson, 1978); and skills to help children overcome their passive orientation and the tendency toward dependency (Castle, 1980). Social workers are also believed to need skills to help them select and train foster caregivers for children with disabilities, who can train such children to become independent (Arkava & Mueller, 1978.)

Tremitiere (1979) and Forsythe & Marshall (1984) promote the use of a model of adoption homestudy that prepares and educates prospective adoptive families for children with disabilities, in addition to evaluating them. This homestudy model requires not only that potential adoptive parents learn about the needs and potential problems of children with disabilities and specialized parenting skills; but, that the caseworkers leading the adoption homestudy groups be equally knowledgeable and skilled in these topic areas.

Kurtz (1979) outlined several specific functions of child welfare professionals in the provision of early intervention services. He stresses the importance of these activities since protective workers serve a high risk population and are in an ideal position to find potentially handicapped youngsters. He suggests that
Caseworkers must 1) be able to recognize potentially handicapping conditions in children; 2) be able to conduct a screening to determine the potential presence of developmental delays; 3) be able to arrange for comprehensive developmental assessment to determine if a disabling condition is present; and, 4) be able to appropriately link the child and family to the proper developmental, remedial, and supportive services.
This research is a longitudinal study to measure whether in-service training increases job related knowledge; the degree to which knowledge gained in an in-service training program is retained by trainees over time; and the degree to which in-service training can effect changes in the on-the-job behaviors of trainees. The initial study data were collected in 1981-82, and the follow-up data were collected eight years later.

In 1981, the principal investigator developed and presented a two-day in-service workshop to train child welfare caseworkers and supervisors to recognize developmental disabilities in children served by their agencies, to refer these children for diagnostic assessment, and to provide and assure appropriate service linkage and delivery. The workshop, entitled "Screening and Identification of Developmental Disabilities," was presented to 24 groups of caseworkers and line supervisors from children's service agencies throughout the state of Ohio. A total of 442 child welfare caseworkers and supervisors attended the training. The content of the training curriculum was standardized, and all 24 of the workshops were conducted by the principal investigator of this research.
The training curriculum was structured to promote the acquisition by caseworkers and supervisors of the following job knowledge and skills: 1) they would be knowledgeable regarding common developmentally disabling conditions, including the principal characteristics and potential developmental outcomes of these conditions; 2) they would be able to recognize early warning signs of developmental disability in children on their caseloads; 3) they would be able to properly refer children for thorough developmental assessment and use this information for treatment planning; 4) they would be able to use assessment information to provide referral to appropriate agencies and special services, and would be able to provide case management and service coordination; 5) they would recognize parents' and caretakers' needs for, and be able to provide, respite and supportive services; and, 6) they would be able to advocate, when necessary, to assure that children with developmental disabilities received appropriate services, and to preserve the children's rights.

The workshop used several training methods, including presentation, audio-visual aids, discussion, and exercises. Trainees were shown 140 slides of normally developing infants and preschool children to strengthen their knowledge of milestones in normal physical, cognitive, social, and emotional developmental domains, and to help them to recognize developmental delays. Presentation was used to communicate
information about many of the primary developmental disabilities, including mental retardation, Down Syndrome, cerebral palsy, epilepsy, autism, spina bifida, PKU, microcephaly, hydrocephalus, Turner Syndrome, Tay Sachs, and Kleinfelter Syndrome. Slides of children with developmental disabilities were used to promote visual recognition of the conditions most commonly encountered by child welfare caseworkers. Extensive descriptions of early signs and symptoms of the conditions were also provided. Slides of a four-year-old with severe mental retardation were used in an exercise wherein trainees practiced their observation and assessment skills to assess the child's level of development. A slide/tape presentation entitled "Handicapism" was used to generate discussion on myths, misconceptions, and attitudes regarding developmental disabilities, to increase trainees' awareness of stereotyping, equal rights legislation, deinstitutionalization, normalization, mainstreaming, and least restrictive environment. Finally, trainees discussed the role of child welfare professionals in early intervention and in assuring the delivery of services to children with disabilities and their families.

The training sessions were held on-site in children's service agencies. All of the workshops were conducted in two successive 6-hour days. The schedule of training workshops, the location of the training, and home agencies of the 442 workshop participants is included in Table I.
### TABLE 1

**SUMMARY OF WORKSHOPS 1981-1982**

<table>
<thead>
<tr>
<th>WS#</th>
<th>Location of Training</th>
<th>Counties Attending</th>
<th>Number Trained</th>
<th>Number Pre/Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Knox Co.</td>
<td>Knox</td>
<td>18</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>Wood Co.</td>
<td>Hancock, Williams, Fulton Sandusky, Ottawa</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>Crawford Co.</td>
<td>Crawford, Huron, Richland</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>4</td>
<td>Summit Co.</td>
<td>Portage, Summit, Stark, Mahoning, Columbiana</td>
<td>21</td>
<td>20</td>
</tr>
<tr>
<td>5</td>
<td>Scioto Co.</td>
<td>Scioto, Gallia, Adams, Jackson</td>
<td>19</td>
<td>12</td>
</tr>
<tr>
<td>6</td>
<td>Athens Co.</td>
<td>Athens, Washington</td>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>7</td>
<td>Cuyahoga Co.</td>
<td>Cuyahoga, Lorain</td>
<td>17</td>
<td>12</td>
</tr>
<tr>
<td>8</td>
<td>Mercer Co.</td>
<td>Mercer, Logan, Van Wert Auglaize, Putnam, Shelby</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>9</td>
<td>Allen Co.</td>
<td>Allen</td>
<td>15</td>
<td>12</td>
</tr>
<tr>
<td>10</td>
<td>Ross Co.</td>
<td>Ross, Madison</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td>Coshocton Co.</td>
<td>Coshocton, Muskingum Harrison, Holmes</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td>12</td>
<td>Adams Co.</td>
<td>Adams, Brown, Hamilton</td>
<td>18</td>
<td>7</td>
</tr>
<tr>
<td>13</td>
<td>Franklin Co.</td>
<td>Franklin</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>Miami Co.</td>
<td>Miami, Montgomery, Shelby</td>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>15</td>
<td>Lucas Co.</td>
<td>Lucas</td>
<td>25</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>Stark Co.</td>
<td>Stark</td>
<td>24</td>
<td>23</td>
</tr>
<tr>
<td>17</td>
<td>Muskingum Co.</td>
<td>Muskingum</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>18</td>
<td>Summit Co.</td>
<td>Summit</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>19</td>
<td>Cuyahoga Co.</td>
<td>Cuyahoga</td>
<td>25</td>
<td>21</td>
</tr>
<tr>
<td>20</td>
<td>Cuyahoga Co.</td>
<td>Cuyahoga</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>21</td>
<td>Summit Co.</td>
<td>Summit</td>
<td>24</td>
<td>17</td>
</tr>
<tr>
<td>22</td>
<td>Cuyahoga Co.</td>
<td>Cuyahoga</td>
<td>28</td>
<td>23</td>
</tr>
<tr>
<td>23</td>
<td>Marion Co.</td>
<td>Marion, Franklin</td>
<td>24</td>
<td>20</td>
</tr>
<tr>
<td>24</td>
<td>Union Co.</td>
<td>Union, Champaign</td>
<td>21</td>
<td>16</td>
</tr>
</tbody>
</table>

**TOTAL** | **442** | **354**
This study is a longitudinal follow-up to determine the effects of having participated in the two-day training workshop on both trainees' knowledge regarding developmental disabilities, and on their job performance subsequent to the training workshop.

The study was designed to evaluate the effects of inservice training at two of Kirkpatrick's four levels: Level Two, learning, and Level Three, changes in behavior on the job (Kirkpatrick, 1959; 1978; 1979; 1987.) The initial portion of the study was designed to measure the degree of change in knowledge and skill, i.e. learning, by 354 trainees as a result of their participation in the training workshop. This was measured by comparing mean pre-test and post-test scores for the trainees in the 24 original training groups. The follow-up portion of the study measured whether the on-the-job behaviors of trainees had changed significantly as a result of having attended the training. This was measured by comparing job behaviors of trainees with job behaviors of subjects who had not been trained.

In none of the study conditions could the study subjects be randomly assigned to the study groups. Therefore, pre-experimental research designs (Campbell & Stanley, 1963) rather than true experimental designs were used to conduct this study.
During the initial phase of the study, pre- and post-tests were administered to each of the subjects attending the 24 two-day training workshops. These training workshops were held in different locations around the state over a period of 1 1/2 years. The methodology used would be considered a one-group pretest-posttest design according to Campbell and Stanley (1963), or a multiple baseline design as defined by Brethower and Rummler (1977, 1979), Brown (1980), and Smith (1987).

A static group comparison design, with features of an "ex post facto" design, was used for the follow-up portion of the study; (Campbell & Stanley, 1963.) Because the study subjects could not be randomly assigned to the experimental and control groups, there was a risk that the groups might differ significantly from one another on critical variables. Fuqua suggests that eliminating group differences on those variables which are likely to confound the variables under investigation can minimize the problems presented by non-equivalent comparison groups (Fuqua, 1979.) In the present study, a control group was formed that matched the experimental group on critical variables that were likely to affect the subjects' level of job knowledge related to developmental disabilities and their performance on the job.
Subjects

The subjects of this research were caseworkers and line supervisors employed in Ohio's public county children's service agencies. In each of Ohio's 88 counties, the children's service agency has mandated responsibility for the provision of protective and placement services to physically abused, sexually abused, neglected, and dependent children and their families. The job responsibilities of child welfare caseworkers include: investigating referrals in which children are suspected of being abused or neglected, and substantiating or dismissing the complaints; assessing the service needs of abused and neglected children and their families; developing case plans and providing case management services; linking children and families with appropriate community services to protect children and rehabilitate families; and providing substitute care in foster, adoptive, or residential placement for children who cannot be assured protection in their own homes. The line supervisor is responsible for case plan review and management, monitoring of worker activities, and technical assistance to assure that caseworkers properly perform their job responsibilities. Many of the children served in children's service agencies have been identified as having developmental disabilities, are significantly delayed in their development, or are at high risk of developmental disability as a result of abuse, neglect, or other maltreatment.
The subjects in the first phase of this research were drawn from the 442 child welfare caseworkers and supervisors who attended one of 24 two-day training sessions on Screening and Identification of Developmental Disabilities in 1981-82. These trainees represented 44 of the 88 county child welfare agencies in the state of Ohio. They worked in a wide range of program areas, including intake and investigation, in-home protective services, foster care placement, and adoption. Only those trainees who had attended the entire two-day session and had completed both the pre- and post-test assessments were included in the subject pool. This group included 354 trainees.

Twenty child welfare caseworkers and supervisors who had attended one of the two-day training sessions in 1981-82, and who were still employed in their children's service agencies, were the subjects in the follow-up experimental group. Six of the original 44 county children's service agencies, in Franklin, Cuyahoga, Marion, Union, Champaign, and Knox counties, were included in the follow-up study. The six agencies were chosen for inclusion in the study to obtain a statewide geographic distribution and a representative sample of subjects from both urban and rural and large and small agencies and communities. The largest of these agencies, in Cuyahoga and Franklin counties, each employs many hundreds of child welfare caseworkers and serves a largely urban population in Cleveland or Columbus. The county children's service agencies in Union and Champaign
counties employ 5-7 child welfare caseworkers each, and they serve primarily rural areas around the small cities of Marysville and Urbana. The agencies in Knox and Marion counties serve the middle-sized cities of Marion and Mount Vernon and the surrounding rural areas, and each agency employs approximately 10-15 caseworkers. The caseworkers and supervisors in the experimental group were selected from the same job classifications as the original 354 trainees, including intake and investigation, provision of in-home protective services, foster care, and adoption services.

A matched control group was also formed, which included twenty child welfare caseworkers and supervisors from the same six children's service agencies as the experimental group, but who had never attended the two-day inservice training workshop. To control the effects of other variables that were likely to influence job performance, the control subjects were matched to the experimental subjects by age, number of years' of practice experience in child welfare, sex, post-secondary educational level, amount of previous training in topics related to developmental disabilities, and level of interest in developmental disabilities. The subjects in the control group represented the same range of job classifications and responsibilities as did the subjects in the experimental group.
### Table 2

**Comparison of Subjects in Experimental and Control Groups on Significant Demographic Variables**

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>CONTROL GROUP MEAN SCORES</th>
<th>EXPERIMENTAL GROUP MEAN SCORES</th>
<th>SIGNIF.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>40.6 Years Range 32-51</td>
<td>40.9 Years Range 30-62</td>
<td>(p = .886)</td>
</tr>
<tr>
<td>Years of Work Experience in Child Welfare</td>
<td>13.7 Years Range 8-24</td>
<td>13.4 Years Range 9-26</td>
<td>(p = .835)</td>
</tr>
<tr>
<td>Years of Post-Secondary Educ. (Possible 2-6 Yrs.)</td>
<td>4.5 Years Range 2-6</td>
<td>4.5 Years Range 2-6</td>
<td>(p = 1.00)</td>
</tr>
<tr>
<td>Level of Previous Training (Possible 1-3)</td>
<td>1.4 Points Range 1-2</td>
<td>1.6 Points Range 1-2</td>
<td>(p = .156)</td>
</tr>
<tr>
<td>Level of Interest In Developmental Disabilities (Possible 1-5)</td>
<td>3.0 Points Range 2-4</td>
<td>3.4 points Range 2-5</td>
<td>(p = .133)</td>
</tr>
<tr>
<td>Sex Distribution</td>
<td>12 Female 8 Male</td>
<td>17 Female 3 Male</td>
<td>(p = .080)</td>
</tr>
</tbody>
</table>
To further control the effects of extraneous variables, several potential subjects were rejected for inclusion in the study groups for one of the following three reasons: 1) they were employed in upper level administrative positions, or in other positions that did not include the direct practice of child welfare, and therefore, they had little direct contact with families and children during the previous 8 years; 2) they had considerable additional education or training in developmental disabilities; or 3) they had worked in units or agencies in which their job responsibilities were to provide specialized services to a caseload of children with developmental disabilities and their families. Subjects who indicated they had had "some previous training" were interviewed to determine the extent of this training. Subjects whose training had included a brief overview of developmental disabilities in a college level child development class were accepted for inclusion in the study. Subjects who had had specific training regarding screening and identification of developmental disabilities, or service provision for this population, were not accepted for inclusion. A profile of the two study groups on the identified critical variables is provided in Table II.
Instruments

Two instruments were developed for use in this research; the pre-test/post test instrument, entitled the "Developmental Disabilities Quiz," and the Behavior Inventory Instrument. Both were developed by the principal investigator.

The pre/post test instrument was originally developed in 1981 to measure the acquisition of knowledge by trainees from attending the two-day training workshop. It was administered to each trainee at the beginning of the 24 two-day workshops and administered again immediately upon completion of the second day of the training. The quiz is a 65-item paper and pencil test that requires approximately 15 minutes to complete. The test items include 30 true/false items to determine the trainee's knowledge of facts, information, and attitudes regarding developmental disabilities; 10 fill-in-the-blank items testing trainees' knowledge of contributing factors to developmental disabilities; 20 items in which trainees are required to determine whether a child described in the quiz item exhibits signs of developmental delay or disability, and to identify and name the condition suggested by the description of the child's symptoms; and 5 items in which trainees are required to properly name the developmental disability exhibited by a child in a photograph. One point was awarded for each correctly answered item. The test instrument was pre-tested for content validity.
prior to the training program by two other trainers, both of whom had been trained to conduct the developmental disabilities workshop for other target groups using the standardized curriculum. The test instrument was also checked for concurrent validity by administering the instrument to three "experts" in developmental disabilities and three social work professionals who had never had training in developmental disabilities. The three experts attained scores of 53, 54, and 61 points respectively, with a mean of 61 points. The three "not trained" professionals attained scores of 17, 34, and 36 points respectively, with a mean of 29 points. This validated the ability of the test instrument to differentiate between "trained" and "not trained" professionals. A copy of the pre/post-test instrument is included in Appendix A.

The Behavior Inventory Instrument was developed for use in conducting structured personal interviews with subjects in the experimental and control groups. The instrument was designed to assess the trainees' performance of six criterion behaviors on their jobs. These behaviors included: 1) recognizing and identifying children on their caseloads who have developmental disabilities; 2) recognizing suspicious early warning signs of developmental disability or delay in children on their caseloads; 3) making appropriate referrals of children for comprehensive developmental assessment and using this information for service planning; 4) providing service referral, service linkage, and case
management to assure that children with developmental disabilities or delays receive appropriate developmental and remedial services; 5) providing supportive and respite services to parents, foster parents, adoptive parents, and other caretakers of children with disabilities; and, 6) providing advocacy for children with developmental disabilities to assure that they receive needed services. A scoring system was developed which awarded zero, one, or two points for each of the six identified criterion behaviors. A score of zero points indicated there was no evidence of performance by the subject of the criterion behavior; 1 point indicated some evidence of performance of the criterion behavior; and 2 points indicated considerable evidence of performance of the criterion behavior. Very specific behavioral indicators were defined to guide the scoring of each category and were included in the Behavior Inventory scoring instructions. A maximum of twelve points could be awarded to each subject on the combined six criterion behaviors. A copy of the Behavior Inventory Instrument and the Scoring Criteria are included as Appendices B and C.

Procedures

During the 1981-82 training program, the Developmental Disabilities Quiz was administered to all trainees at the beginning of each of the original 24 training workshop sessions. Trainees were instructed to answer the questions as honestly as possible to provide an accurate measure of their current level of knowledge. They were
possible to provide an accurate measure of their current level of knowledge. They were instructed to use the "I don't know" column, or to leave an item blank, if they did not know the answer, and were told not to randomly guess at answers. Trainees were reassured that the data would be used to measure the group's level of learning rather than to test them individually. This was done to reduce test anxiety and to promote honesty of responses. At the completion of the two-day workshop, the trainees again took the test, and the same instructions were presented. Pre and post-test data were recorded for 354 trainees in the 24 groups who attended the full workshop session and completed both tests. Mean pre-test and post-test scores were tabulated and compiled for these 354 subjects and are included in Table III-A.

The follow-up portion of the study was conducted in 1990. Caseworkers and supervisors who had attended one of the original 24 training workshops, and who were still employed in the child welfare system, were located by a computer search of personnel training records from the 88 county agencies. A total of 84 of the original 354 trainees were determined to still be employed by their agencies. Trainees in the six targeted children's service agencies were contacted and asked if they would be willing to participate in a follow-up study subsequent to their attendance at the developmental disabilities training in 1981-82. All agreed. Three potential subjects were disqualified for reasons
identified above. Twenty remaining trainees comprised the experimental group. Demographic data, including age, sex, level of post-secondary education, length of experience in child welfare, previous training in developmental disabilities, and level of interest in developmental disabilities, were collected for each subject using a data sheet that was completed by the subject prior to the beginning of the structured interview. This information was then used to construct the matched control group. A copy of the data sheet is included in Appendix D.

Staff development personnel and supervisors in the six county children's service agencies represented by the subjects in the experimental group were asked to help identify caseworkers and supervisors within those same agencies who matched the characteristics needed for inclusion in the control group. These individuals were also asked if they would be willing to participate in the study. All agreed. Seven of the recommended individuals were disqualified from participation for one of the three reasons identified above. Twenty of these individuals comprised the control group.

A structured personal interview was scheduled with each of the 40 subjects on-site in their employing agencies. Two interviewers conducted the interviews with randomly assigned subjects. Immediately prior to the interview, the Developmental Disabilities Quiz was administered to each of subjects. The
instructions for completing the Quiz were identical to those given at the time of the original testing during the training workshops.

An interview of approximately 1/2 hour in length was then conducted with each subject using the Behavior Inventory Instrument. Prior to beginning data collection, the two interviewers had thoroughly discussed the instrument and had standardized the manner in which the questions would be asked to assure consistency between interviewers. The subjects' responses to all six questions were recorded directly on the Behavior Inventory Instrument form, and were also tape recorded.

The Developmental Disabilities Quiz and the data on the Behavior Inventory were scored by a third psychology professional, an expert in both child welfare and developmental disabilities, who had been trained to use the standardized curriculum and had taught the two-day workshop from 1986-1989. The scorer participated in the initial discussions between the interviewers regarding the manner in which the interviews would be conducted. The scorer was trained by the principal investigator in the use of the scoring criteria. The principal investigator gave the rater a copy of the Scoring Criteria data sheet and reviewed the types of behaviors that could be included in each scoring category. The scorer then "practiced" by scoring four sample protocols that had
been previously scored by the Principal Investigator. The scorer's ratings were consistent with the ratings given by the principal investigator on the same sample protocols.

At the beginning of the personal interview, each subject randomly selected a subject number from 1-50, and both the Quiz and the Behavior Inventory Interview data sheet were identified by the subject's assigned number only. The scoring of both protocols was, therefore, performed by a blind rater.

**Measures**

The independent variable in this research was the completion of the two-day inservice training workshop on Screening and Identification of Developmental Disabilities during 1981.

There were three dependent variables. The first was the degree of knowledge gained by trainees from attending the two-day workshop. This was determined by the increase in test scores from the pre-test condition to the post-test condition at the time of the initial workshop. Data was used for all 354 trainees who had completed both the pre and post test assessments at the time they attended the training workshop.

The second measure was the retention by trainees of knowledge regarding developmental disabilities over time. The data for this
measure was gathered by re-administering the post-test to the trainees in the experimental group and comparing their scores with those of the control subjects, who had not received the training.

The third measure was the performance by caseworkers and supervisors on their jobs of the six criterion behaviors on the Behavior Inventory Instrument. The data for this measure was gathered from the personal interviews with each of the 40 subjects in the experimental and control groups.

The effects of other variables that might have affected both the increase in job knowledge and the performance of the behavioral criteria, including the subjects' length of experience in child welfare, age, level of interest in developmental disabilities, employing agency, and previous training, were minimized through the use of the matched control group and appropriate statistical procedures.

Data Analysis

A T-test was performed to measure the level of comparability of the experimental and control groups on the six matching variables of age, sex, level of child welfare experience, level of post secondary education, level of interest, and previous training in developmental disabilities.
The data for the first measure, the increase in knowledge by trainees from attending the training workshop, included the test scores for the 354 trainees who completed both the pre and post-tests at the time of the initial training workshops. The mean pre-test and mean post-test scores were computed and compared for all 354 subjects in the initial 24 training groups. The data was tested for statistical significance using the paired T-test, as the same subjects were used for both measures.

The data for the second measure, the retention of knowledge by trainees over time, were derived from the scores on the follow-up post-test administered to the 20 caseworkers in the experimental group. The mean score for the experimental (trained) group was compared to the mean test score for the 20 subjects in the control group, who had not received the training. This measure was tested for significance using a one way ANOVA.

The data for the third measure, the degree of performance of the six criterion behaviors, were derived from the raw point scores achieved by the 40 subjects in the experimental and control groups for each of the six criterion behaviors on the Behavior Inventory Instrument. The mean scores for the experimental group were compared to the mean scores for the control group on each of the six criterion behaviors individually, and, on the six criterion behaviors together. The data were also tested using a one-way ANOVA.
Pearson Correlation Coefficients were computed to determine the relationships of each of the six matching variables (age, sex, education, experience, interest, and previous training) to the subjects' test scores; to their total point scores on the six behavioral criteria combined; and to each of the six behavioral criteria individually.

A one-way ANOVA was performed to determine the relationship between the subjects' level of post-secondary education, that is Bachelors' or Masters' level, and their performance on each of the six behavioral criteria.

A MANOVA was performed to determine the potential effects of differential learning or practice effects in the subjects' responses between Item #1 and Item #6 on the Behavior Inventory Interview.

Finally, an Analysis of Covariance was performed to determine the degree of the main effect, that is the trained versus non-trained condition, after the effects of the covariates (age, sex, experience, education, interest, and previous training) had been considered.

The SPSS statistical program was used for all data analysis.
The experimental and the control groups were tested using a T-test to determine whether they were statistically matched on the six characteristics believed to exert potential effects on the subjects' performance in the study. These six characteristics were: age, sex, level of post-secondary education, length of experience in child welfare, previous training in developmental disabilities, and level of interest in developmental disabilities. The result of the t-test indicated that sex was the only characteristic on which the groups were not statistically matched. The experimental group was 85% female and 15% male, and the control group was 60% female and 40% male, with t=1.8 and p=.080, indicating that sex was marginally significant. None of the other matching variables was significant, indicating that the two groups were statistically matched on all remaining characteristics.

The first portion of this research was designed to evaluate the effectiveness of the inservice training program in increasing trainees' knowledge and skills, which corresponds to Kirkpatrick's Level II assessment of training, the measure of learning (Kirkpatrick, 1959, 1978, 1979, 1987.) Pre- and post-
test scores on the Developmental Disabilities Quiz (referred to throughout this discussion as "the test") were used as the measure of knowledge and skill. The pre- and post-test scores were used in two analyses.

The first analysis determined the impact of training on learning immediately following the training. The pre-test scores of the 354 trainees in the original 24 training groups were compared with post-test scores of the same trainees upon completion of the training program. The mean pre-test score for the 354 trainees was 27.7 of a possible 65 points. The mean post-test score for the same group was 46.8 of a possible 65 points. The mean difference between the pre- and post-test scores was 19.1 points. The paired T-test indicated a highly significant difference, with $T = 43.49$, and $p < .000$.

The second measure was to determine whether learning is retained by trainees over time. The Developmental Disabilities Quiz was re-administered eight years after the training to the 20 subjects in the experimental group, and was also administered to the 20 subjects in the control group. The mean test score of the experimental group was 45.7 points, and the mean test score of the control group was 32.8 points, of a possible total of 65 points. The mean difference in scores was 12.9 points. This difference was found to be highly significant on a one-way
ANOVA; $f = 20.48 (1,38)$ and $p = .0001$. The subjects in the experimental group had taken the Developmental Disabilities Quiz a total of three times; twice as pre- and post-test measures in 1981, and once during the follow-up condition. The subjects in the matched control group, by comparison, took the Quiz only once, during the 1990 follow-up condition. Tables III-A and III-B show the breakdown of pre- and post-test scores of all study groups.

An analysis of covariance was conducted to determine the effects of the training on the subjects' test performance beyond the effects of group differences. The covariates were the six characteristics on which the experimental and control groups were matched; the subjects' age, sex, level of post-secondary education, years of experience in child welfare, previous training in developmental disabilities, and level of interest in developmental disabilities. When the effects of these variables were covaried out, the main effect of the training condition on the subjects' test performance remained highly significant, with $f = 17.387$ and $p = .000$. Of the six covariates, the subjects' level of post-secondary education significantly affected their test performance, with $f = 8.840$ and $p = .006$, as did the subjects' sex, with $f = 6.281$ and $p = .017$. 
TABLE 3

COMPARISON BETWEEN GROUPS ON DEVELOPMENTAL DISABILITIES QUIZ SCORES

<table>
<thead>
<tr>
<th>Group Description</th>
<th>Pre-Test</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A) Pre and Post-Test Assessments of 354 Trainees in 24 Training Groups, 1981</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Score</td>
<td>27.7 Points</td>
<td>46.8 Points</td>
</tr>
<tr>
<td>N = 354</td>
<td>Possible 65</td>
<td>Possible 65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group Description</th>
<th>Pre-Test</th>
<th>Post-Test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>B) Test Scores of 20 Subjects in Matched Control Group and 20 Subjects in Experimental Group, 1990 Follow-Up.</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean Score</td>
<td>32.8 Points</td>
<td>45.7 Points</td>
</tr>
<tr>
<td>N = 20/Group</td>
<td>Possible 65</td>
<td>Possible 65</td>
</tr>
</tbody>
</table>
The second part of the study was designed to determine the effect of training on subsequent job performance, which corresponds to Kirkpatrick's Level III of training evaluation, the assessment of behavior change (Kirkpatrick, 1959, 1978, 1979, 1987.) The measures of job performance were the points awarded to the subjects on the six behavioral criteria on the Behavior Inventory Instrument. Mean scores were also computed individually for each of the six criterion behaviors. The mean score for the experimental group on the combined six criterion behaviors was 8.9 of a possible 12 points. The mean score for the control group on the combined six criterion behaviors was 5.2 of a possible 12 points. There was a difference of 3.7 points. This difference was found to be highly significant on a one-way ANOVA, with $f = 18.00$ (1,38) and $p < .0001$.

An analysis of covariance was conducted to determine the effects of the training beyond group differences, using the six matching variables as covariates; the subjects' age, sex, level of post-secondary education, years of experience in child welfare, previous training in developmental disabilities, and level of interest in developmental disabilities. When the effects of these variables were covaried out, the main effect of the training condition remained highly significant on the trainees' performance on the six behavioral criteria, with $f = 11.902$ and
Of the six covariates, only the previous training of the subjects proved significant, with $f = 6.335$ and $p = <.017$. Sex was marginally significant, with $f = 2.96$ and $p = <.095$.

Behavior Criterion #1 was the subjects' recognition and identification of children on their caseloads who had developmental disabilities. On this behavior, the mean score for the experimental group was 1.8 of a possible 2 points, and the mean score for the control group was 1.2 of a possible 2 points, a difference of .6 points. The one-way ANOVA results were; $f = 8.34 \ (1,38)$ and $p = <.0064$. On the analysis of covariance, the main effect remained highly significant, with $f = 4.075$, $p = <.052$. Previous training was the only significant covariate, with $f=10.27$ and $p=.<.003$.

Behavior Criterion #2 was the subjects' recognition and identification of suspicious early warning signs of developmental disabilities in children on their caseloads. On this behavior, the mean score for the experimental group was 1.4 of a possible 2 points, and the mean score for the control group was .55 of a possible 2 points, a difference of .85 points. The one-way ANOVA results were; $f = 12.62 \ (1,38)$ and $p = <.0010$. On the analysis of covariance, the main effect remained highly significant, with $f=7.653$ and $p=.<.009$. Again, previous training was the only significant covariate, with $f=4.065$ and $p=.052$. 
Behavior Criterion #3 was the subjects' referral of children suspected of having developmental disabilities for a thorough diagnostic and developmental assessment. On this behavior, the mean score for the experimental group was 1.5 of a possible 2 points, and the mean score for the control group was .85 of a possible 2 points, a difference of .65 points. The one-way ANOVA indicated the difference to be significant, with \( f = 6.82 \) (1,38) and \( p = .0129 \). On the analysis of covariance, the main effect remained marginally significant, with \( f = 3.471 \) and \( p = .072 \). None of the individual covariates was, by itself, significant.

Behavior Criterion #4 was the subjects' referral of children with developmental disabilities to appropriate service agencies and their provision of appropriate case management and case coordination services. On this behavior, the mean score for the experimental group was 1.6 of a possible 2 points, and the mean score for the control group was .75 of a possible 2 points, a difference of .85 points. The one-way ANOVA showed the difference to be significant, with \( f = 13.36 \) (1,38) and \( p = .0008 \). On the analysis of covariance, the main effect remained highly significant, with \( f = 10.81 \) and \( p = .002 \). None of the covariates was individually significant.
Behavior Criterion #5 was the subjects' recognition of caretakers' need for special emotional support and respite, and provision of these services. On this behavior, the mean score for the experimental group was 1.30 of a possible 2 points, and the mean score for the control group was .95 of a possible 2 points, a difference of .35 points. The one-way ANOVA showed this difference not to be significant, with $f = 1.71$ (1,38) and $p = <.1983$. On the analysis of covariance, while the main effect remained not significant, sex, with $f = 3.34$ and $p = <.077$, and previous training, with $f = 3.49$ and $p = <.071$, were both marginally significant.

Behavior Criterion #6 was the subjects' provision of direct advocacy services to children on their caseloads with developmental disabilities. On this behavior, the mean score for the experimental group was 1.30 of a possible 2 points, and the mean score for the control group was .90 of a possible 2 points, a difference of .40 points. The one-way ANOVA indicated this difference was not significant, with $f = 2.17$ (1,38) and $p = <.1488$. On the analysis of covariance, while the main effect remained not significant, sex was significant, with $f = 5.017$ and $p = <.032$.

The distribution of point scores for each of the six behavioral criteria and for the six criteria combined are included in Table IV.
### TABLE 4

**COMPARISON BETWEEN GROUPS ON BEHAVIORAL CRITERIA**

<table>
<thead>
<tr>
<th>CONDITION</th>
<th>CONTROL GROUP (N=20)</th>
<th>EXPERIMENTAL GROUP (N=20)</th>
<th>F-VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Six Behaviors Combined</td>
<td>Mean 5.2 SD 2.6872</td>
<td>Mean 8.9 SD 2.8266</td>
<td>F = &lt;.0001</td>
</tr>
<tr>
<td>Behavior #1 Recognition of DD</td>
<td>Mean 1.20 SD .7678</td>
<td>Mean 1.80 SD .5231</td>
<td>F = &lt;.0064</td>
</tr>
<tr>
<td>Behavior #2 Recognizes Early</td>
<td>Mean .55 SD .8256</td>
<td>Mean 1.40 SD .6806</td>
<td>F = &lt;.0010</td>
</tr>
<tr>
<td>Warning Signs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior #3 Refers: Diagnostic</td>
<td>Mean .85 SD .7452</td>
<td>Mean 1.5 SD .8272</td>
<td>F = &lt;.0129</td>
</tr>
<tr>
<td>Assessment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior #4 Refers: Services</td>
<td>Mean .75 SD .7864</td>
<td>Mean 1.60 SD .6806</td>
<td>F = &lt;.0008</td>
</tr>
<tr>
<td>Behavior #5 Provides Support</td>
<td>Mean .95 SD .8256</td>
<td>Mean 1.30 SD .8645</td>
<td>F = &lt;.1983</td>
</tr>
<tr>
<td>and Respite Svcs.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavior #6 Advocacy</td>
<td>Mean .90 SD .8522</td>
<td>Mean 1.30 SD .8645</td>
<td>F = &lt;.1488</td>
</tr>
</tbody>
</table>

---

Total Mean Score: Maximum 12
Mean Score for Each Behavior: Maximum 2
On the MANOVA, there were no differential learning or practice effects between Question #1 and Question #6 for the experimental (trained) versus the control (not trained) groups. The results were not significant, with $f = .97$ and $p = .436$.

A Pearson Correlation Coefficient was computed to assess the relationship between the test scores of the 40 subjects in the experimental and control groups and their performance on the six behavioral criteria, to determine the degree to which greater knowledge is correlated with improved job performance. The Correlation Coefficient was $r = .41$, $p = .004$. A summary of behavior point scores and test scores for each of the 40 subjects is included in Table V.

Pearson Correlation Coefficients were also computed to determine the relationship between each of the six matching variables and the subjects' performance scores on the combined behavioral criteria. The subjects' previous training in developmental disabilities had the highest correlation with their behavioral performance, and their level of education had the lowest. The correlations were as follows: age $= .25$, $p = .054$; sex $= .26$, $p = .048$; experience $= .23$, $p = .072$; education $= .07$, $p = .315$; previous training $= .30$, $p = .027$; level of interest $= .20$, $p = .100$. 
TABLE 5

COMPARISON OF TEST SCORES AND TOTAL POINTS ON BEHAVIORAL CRITERIA FOR ALL SUBJECTS

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Points</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>11</td>
<td>46</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>51</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>46</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>11</td>
<td>38</td>
</tr>
<tr>
<td>12</td>
<td>12</td>
<td>46</td>
</tr>
<tr>
<td>13</td>
<td>10</td>
<td>56</td>
</tr>
<tr>
<td>17</td>
<td>8</td>
<td>47</td>
</tr>
<tr>
<td>19</td>
<td>12</td>
<td>49</td>
</tr>
<tr>
<td>21</td>
<td>10</td>
<td>37</td>
</tr>
<tr>
<td>23</td>
<td>11</td>
<td>39</td>
</tr>
<tr>
<td>29</td>
<td>9</td>
<td>50</td>
</tr>
<tr>
<td>31</td>
<td>11</td>
<td>45</td>
</tr>
<tr>
<td>37</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td>38</td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td>40</td>
<td>3</td>
<td>39</td>
</tr>
<tr>
<td>42</td>
<td>12</td>
<td>54</td>
</tr>
<tr>
<td>44</td>
<td>7</td>
<td>41</td>
</tr>
<tr>
<td>47</td>
<td>8</td>
<td>54</td>
</tr>
<tr>
<td>48</td>
<td>3</td>
<td>35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject #</th>
<th>Points</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>40</td>
</tr>
<tr>
<td>8</td>
<td>8</td>
<td>45</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>37</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>40</td>
</tr>
<tr>
<td>14</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>15</td>
<td>6</td>
<td>37</td>
</tr>
<tr>
<td>16</td>
<td>6</td>
<td>53</td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>20</td>
<td>6</td>
<td>42</td>
</tr>
<tr>
<td>22</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>24</td>
<td>4</td>
<td>30</td>
</tr>
<tr>
<td>25</td>
<td>8</td>
<td>28</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>29</td>
</tr>
<tr>
<td>27</td>
<td>6</td>
<td>32</td>
</tr>
<tr>
<td>28</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>30</td>
<td>5</td>
<td>46</td>
</tr>
<tr>
<td>32</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>35</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>36</td>
<td>9</td>
<td>12</td>
</tr>
</tbody>
</table>

Mean 8.9 45.7

Mean 5.2 32.8

Maximum Possible Points = 12
Maximum Possible Test Score = 65
Correlations were also computed between the six matching variables and subjects' responses on each of the behavioral criteria. Only eight significant correlations were found. The age variable affected behavior #6, advocacy, with a correlation of .44, p<.002. The sex variable affected behavior #5, supportive services to parents, with a correlation of .29, p<.034. The subjects' level of experience in child welfare only slightly affected behavior #5, supportive services to parents, with a correlation of .24, p<.065; and, level of experience affected behavior #6, advocacy, with a correlation of .28, p<.039. The subjects' level of post-secondary education did not significantly correlate with their performance on any behavior. Their level of interest in developmental disabilities affected behavior #3, their referral of children for developmental assessment, with a correlation of .30, p<.027. The variable of previous training in developmental disabilities had the greatest effect on the subjects' performance. Previous training affected behavior #1, identification of developmental disabilities, with a correlation of .42, p<.003; behavior #2, recognition of early warning signs, with a correlation of .29, p<.032; and behavior #5, supportive services to parents, with a correlation of .29, p<.034.

The Pearson Correlation Coefficient was also computed to determine the relationship between the six matching variables
and the subjects' test performance. A moderate correlation of .40 was found between the subjects' level of post-secondary education and their test scores, with \( p = .005 \). The remaining correlations ranged between .11 and .20, but none was significant.

Finally, the subjects were examined to determine differences in their test performance and on-the-job behaviors as a function of their level of post-secondary education. There were 19 subjects who had a Bachelors' degree, computed at four years of post-secondary education. There were 16 subjects who had a Masters degree, computed at six years of post-secondary education. A one-way ANOVA was computed to determine the differences between the Bachelors' level and Masters' level subjects on their test scores and in their responses on the six behavioral criteria. The mean score of the behavioral criteria for the Bachelors' level group was 6.68 points; the mean score of the behavioral criteria for the Masters' level group was 7.06 points. This difference was found to be not significant, with \( f = .108 \), and \( p = .7443 \). Of the one-way ANOVAs computed to compare the level of post-secondary education with each of the six behavioral criteria individually, none was significant.

The study results appear to strongly support both research hypotheses. The first hypothesis, that caseworkers and supervisors who attended the two-day workshop will score
significantly higher on follow-up knowledge testing than will workers who have not had training in developmental disabilities, was supported by the comparison of test scores between the experimental and the control groups. The experimental group's mean score of 45.7 points was significantly higher than the control group's mean score of 32.8 points; \( p<.0001 \). This score remained highly significant even after the effects of covariates were considered.

The second hypothesis, that trained workers will demonstrate significantly higher numbers of criterion behaviors in their casework activities than will caseworkers who did not receive the training, was also supported by the findings. When the scores on the six criterion behaviors are combined, the experimental group's mean point score of 8.9 points is significantly higher than the control group's mean point score of 5.2. This score also remained highly significant, even after the effects of covariates were considered.

An independent examination of each of the six criterion behaviors showed significant differences between the study groups on four of the six criterion behaviors. This would suggest that the training had a greater effect on some of the worker's on-the-job behaviors than on others. This will be more fully explored in the Discussion section which follows.
CHAPTER V
DISCUSSION

This study was designed to examine the effects of inservice training on the acquisition and retention of job knowledge, and on the performance of specific job behaviors, by staff in public children's service agencies. The study results supported both of the study hypotheses, suggesting that inservice training does increase trainee knowledge and also promotes behavior changes in the work place.

The results of the knowledge testing were highly significant, suggesting that the trainees as a group gained considerable knowledge from the training. The original training group of 354 subjects demonstrated a point increase of 19.1 of 65 possible points between the pre-test condition and the post-test condition. The wide range of scores in the pre-test condition suggests that the child welfare workers came into the training with varying degrees of knowledge about developmental disabilities. However, the test results suggest that despite trainees' initial level of knowledge, they almost universally increased their knowledge from having attended the training.
The comparison of the experimental group's test scores with those of the control group supported the contention that knowledge is retained over time. The mean test score of the experimental group of 20 subjects was 45.7 points of a possible 65; the mean score of the 20 subjects in the control group was 32.8 points of a possible 65, with a mean difference of 12.9 points. The effect of the main condition, the training, on test performance remained highly significant even after the effects of covariates were considered. This suggests that once knowledge is acquired from the training, it is not likely to be forgotten.

A statistical comparison of the test scores of the original post-test group (N=354) and the follow-up experimental group (N=20) would not be valid, considering the differences in the sample sizes. However, the closeness in the point scores of these two groups (46.8 and 45.7 points) would suggest that the experimental subjects were very similar in their knowledge of developmental disabilities to their condition immediately upon completion of the two-day workshop eight years earlier. It should be noted that the subjects in the experimental group had taken the Quiz three times, as compared to the subjects in the Control Group, who took the test only once. However, since the first two test administrations had occurred eight years earlier, it is not likely that the experimental subjects' performance on the follow-up
test would have been significantly affected by their having remembered the questions from the earlier administration of the test.

The data indicated that training also had a strong positive effect on the subjects' performance of the six criterion behaviors. The experimental group achieved a mean score of 8.9 of a possible 12 points, while the control group achieved a mean score of 5.2 of a possible 12 points. This difference was highly significant. The effect of the main condition, the training, remained highly significant even after the effects of covariates were considered. This suggests the potential for inservice training to have a powerful effect on the job performance of staff, regardless of other situational variables.

On the six behavioral conditions, the differences in performance of the experimental and control subjects were found to be significant for four of the behaviors and were found not to be significant for the other two.

A comparison of the groups' responses to behavior item #1 suggested that the control group was less likely than the experimental group to know if children on their caseloads had developmental disabilities, or to properly identify those conditions. The mean score of the control group was 1.2 of a
possible 2 points, while the mean score of the experimental group was 1.8 of a possible 2 points. This difference was found to be significant. Using the behavioral scoring criteria to interpret the study findings, the control group's mean score of 1.2 points suggests that most of the control subjects could describe children with developmental problems without being able to name the disability, or they could identify a child as having a disability without being able to describe the symptoms or characteristics of the child's condition. By comparison, the experimental group's mean score of 1.8 of a possible 2 points indicated that they most often properly identified children on their caseloads who had disabilities, accurately named the children's conditions, and accurately described characteristics of the condition. The responses of the experimental group suggested they had a much better understanding of the nature of the children's conditions. Many of the control subjects failed to differentiate between developmental disabilities and other conditions including general medical, behavioral, or emotional problems of children. Four of the control subjects indicated they had never had a child on their caseload who had a developmental disability. With the preponderance of children in the child welfare field with developmental problems, including a sizeable percentage of children who are mentally retarded at the mild level, it is almost impossible that in 10+ years of direct practice a caseworker
would never have a child with a disability on her caseload. This suggests that non-trained workers may fail to recognize children's problems as developmental disabilities.

The most significant difference between the study groups was found on behavior item #2, the ability of trainees to recognize early warning signs of developmental disabilities in children on their caseloads. The mean score of the trained group was 1.40, compared with a mean score of .55 for the control group. Using the scoring criteria to interpret the study findings, the control group's mean score of .55 of a possible 2 points suggests that about half the controls had never suspected any child on their caseload as having had a developmental disability; and while the remaining controls either recognized when children were "slower than they should be," or recognized more visible delays such as speech or hearing deficits, they did not recognize unusual developmental patterns or combinations of developmental characteristics as being specific to a particular disability. Subjects in the experimental group, for example, were much more likely to recognize a lack of physical mobility and delayed physical development in an infant, combined with muscular rigidity, as predictable symptoms of cerebral palsy.

This finding is consistent with the positions taken by several authors in Chapter II. Schilling, Kirkham, and Schinke (1986)
suggested that the child protection system may be failing to recognize and document children with handicaps in populations of abused and neglected children, and that workers' lack of specific training in developmental disabilities was a significant contributor to this problem. Kurtz (1979) also suggested that lack of knowledge and skill of child welfare workers was a serious impediment to their identification of children with developmental disabilities, as many workers lacked the ability to make judgments about early signs of disabling conditions or handicaps. Our study data suggest that child welfare professionals who have received only 2 days, or 12 hours, of inservice training on the screening and identification of developmental disabilities are almost three times as likely to recognize early warning signs of developmental problems in children on their caseloads than are staff who have not been trained. This supports the potential value of inservice training in resolving the problems suggested by Kurtz, Schilling, Kirkham and Schinke.

A significant difference between groups was also found on behavior item #3, referral of a child for a comprehensive developmental assessment. The mean score on this item for the control group was .85 points, and the mean score for the experimental group was 1.5 points. Using the behavioral scoring criteria to interpret the study findings, the control group's mean
score of .85 of a possible 2 points suggests that the majority of subjects had made referrals for diagnostic assessment, but stated the purpose of such assessments as "getting more information." They could not describe specific referral questions, nor could they elucidate how they intended to use the information once they had received it. Approximately 15% of this group had never made a referral of a child for a diagnostic assessment. By comparison, approximately one half of the experimental group had not only made referrals of children for diagnostic assessment, but had clearly delineated referral questions. They also had a plan to use assessment information to assure that the services provided to the child were appropriate for the child's condition and situation. Their referral questions also typically reflected general service goals of enhancing the child's development, finding the most appropriate placement, learning how to better manage the child's condition, and preparing substitute care parents to meet a child's special needs.

An equally significant difference between groups was found on behavior item # 4, referral of children and their families to special service providers, and coordination of services to meet children's multiple needs. The mean score of the control group on this item was .75 of a possible 2 points, and the mean score of the experimental group was 1.60 points. Using the scoring criteria to interpret the current study findings, the difference in
scores suggests that the experimental group was more than twice as likely to make appropriate service referrals to address the developmentally disabling condition, to understand the multiple service needs of children with developmental disabilities, and to coordinate services. Many of the experimental subjects had identified community resources that specialized in providing case management, education, parent support, and early intervention services for children with developmental disabilities. They had developed collaborative relationships with individual staff members in these agencies, they made frequent referrals of children to these programs, and they generally worked jointly with the specialized agency to assure coordination of services.

By comparison, the subjects in the control group had made referrals, or directed families to seek services, for a single identified problem, such as referring to a speech therapy clinic because a child's speech was delayed; or, they had not referred any children for specialized services. It is possible that this group's general failure to recognize and to understand the complex nature of many developmentally disabling conditions might have contributed to their lower rate or referral for coordinated and comprehensive services.

The experimental and control groups differed in their behaviors on item #5, the provision of supportive and respite services to parents, but this difference was not significant. The mean score
of the control group was .95 of a possible 2 points, and the mean score of the experimental group was 1.30. Both the control subjects and the experimental subjects had made referrals to acquire respite care for parents and caretakers of children with developmental disabilities. The experimental group were more likely to provide other supportive interventions as well, such as doing one-on-one supportive counseling with a parent, engaging the services of a homemaker to go into the home to provide help, or referring parents to support groups. However, the experimental group's mean score if 1.3 points on this item is lower than their mean scores on any of the previous items. The provision of respite care and support services to parents may be more difficult for the caseworker to accomplish than identifying disabling conditions or initiating referrals to service providers. Respite services and parent support groups are specialized services which are not always available or easily accessible. In addition, providing direct emotional support to a parent of a child with a disability may require considerable direct personal contact by the caseworker on an ongoing basis. The high workload of many child welfare caseworkers often prohibits a large expenditure of time with any single family. It appeared that workers were more likely to refer parents to respite care services or to agencies that offered support services than to provide these services themselves. The control group's score of 9.5 on this item suggested that as a group these subjects
typically referred parents for respite care services but did not provide any other supportive interventions. While the difference between the groups was not statistically significant, the trend was consistent with the other behavior items, in that the experimental subjects as a group were more likely to provide parental support. It is possible that these workers were more aware of the parent's needs for supportive services than were the controls, but that situational variables prevented their being able to deliver this service. This might be an appropriate question for further research.

The study groups' responses to behavior item #6, the provision of advocacy services, were also determined to be not significant. The control group's mean score was .90 points, and the experimental group's score was 1.3 points. Using the scoring criteria to interpret the study findings, the control group's mean score of .90 suggests that in selected situations the workers served in an advocate role for a child with a disability, but that no special or extended effort was required. They typically made phone calls to procure equipment or services, or they attended IEP meetings. The experimental subjects were more likely to perform multiple advocacy activities for a single child, or were more willing to engage in conflict with other service agencies and practitioners to assure that the child's needs were met. Some of the experimental subjects described using the legal
system or the court process to acquire services. Some challenged the medical profession until they were able to get a definitive diagnosis and treatment recommendations for a child. Some had been licensed and trained as surrogates and regularly participated in the IEP process. Few of the control subjects described behaviors that advocated at this level of sophistication. Two issues might explain the absence of statistical significance on this behavior, despite a trend toward more advocacy behavior on the part of the experimental subjects. Advocacy work in situations of conflict often requires considerable time of the caseworker, and again, high workloads might prevent workers from being able to engage in advocacy activities, even if they recognized the need. In addition, one of the areas in which many of the subjects in both groups advocated for children with disabilities was in finding them appropriate specialized placements. The identification of appropriate substitute care placements for children with all types of problems is one of the central functions of a child welfare caseworker. The control subjects, therefore, advocate for appropriate placement on a regular basis for children on their caseloads with other types of problems, including severe behavioral handicaps and emotional problems. They would, therefore, be likely to engage in similar behavior for children on their caseloads with developmental disabilities. This might account for their somewhat higher mean score on this behavioral item than on most of the other items.
The data would suggest that the training had a positive impact upon child welfare practice on all six behavioral criteria. However, because of the "quasi-experimental" nature of the study design, it is still possible that other variables might have affected the study. Most of the important variables that were thought to potentially interfere with the validity of the data were controlled through the matching of experimental and control subjects. The two groups were matched statistically on variables of age, years of child welfare experience, previous training in developmental disabilities, level of post-secondary education, and level of interest. The experimental group indicated a not significant but slightly higher level of interest in the topic of developmental disabilities, but one cannot be sure whether this is the cause of, or the result of, greater knowledge and understanding of developmental disabilities. The groups were not statistically matched on the variable of sex. The study subjects in both groups were primarily female, but the control group had 8 male subjects and the experimental group had only 3. Of all the variables under consideration, however, the sex of the subject is probably the least critical in potentially affecting the study outcomes.

There were other variables that were not controlled through the matching process, and that could potentially have affected the
study outcomes. One was the nature and quality of the direct supervision received by the caseworkers; specifically, the supervisors' level of knowledge of developmental disabilities, and their commitment to proper service interventions, including assigning a high priority status to the provision of services to children. A second variable was the degree to which their local agencies that serve clients with developmental disabilities interact and cooperate with the child welfare agency. These factors could include whether compactual agreements existed between these agencies for collaborative service delivery, or whether refusal by the local MR/DD agency to assume responsibility for children who had been identified as abused or neglected precluded easy referral of children into the MR/DD system. The availability of appropriate specialized services within a community might also have affected the subjects' responses on the items related to service referral and coordination.

A third variable was the specific nature of the caseworkers' jobs in the agency. The experimental and control groups both included staff who had worked in all service programs within the child welfare system, including intake and investigation, child protective services, foster care, and adoption. The differences in job responsibilities of staff in these service areas could potentially affect their responses to particular items. For
example, intake workers are generally involved with families for approximately 30 days, long enough to substantiate allegations of maltreatment and develop an initial service plan. They would be likely to recognize early warning signs of disability and would very likely refer, or recommend that the ongoing caseworker refer, a child for assessment or special services. The intake worker would be much less likely to perform activities related to parent support or advocacy, or service follow-up, because of the short length of their involvement with the family. Adoption workers, by contrast, might have extensive involvement in procuring special services for a child with a disability in an adoptive home and in providing support to an adoptive parent. However, they would be less likely to encounter large populations of abused and neglected children, and would, therefore, be less likely to personally identify undiagnosed conditions in these children. Further investigation could more clearly delineate the particular behaviors that are performed by child welfare practitioners in a specific job classification, and could explore the degree to which inservice training can promote the improved practice of these jobs.

The study methodology must also be considered when interpreting the results. First, the study methodologies were "pre-experimental" in nature, and as a result, the validity of the findings may be supported with less assurance than had the study
design been more tightly controlled. For example: the measure of knowledge on the pre- and post-tests of the 354 original trainees demonstrated a significant increase in knowledge as a result of the training program. The methodology, formally a one group pre-test/post-test design, is limited in its ability to control for the effects of variables other than the experimental variable. However, by staging the data collection over a period of time in different parts of the organization, we may reduce the effects of any particular environmental condition on the study outcomes. Brethower and Rummler (1977, 1979) refer to this method as a "multiple baseline" study design. They suggest that "when a program is introduced to different organizational units at different times... factors such as changes in the market place, work loads, business cycles, and personal variables are spread; therefore, one can attribute the change in performance as a result of the training with more certainty than if only one region had shown such improvement." In the current study, the 24 training groups included trainees from different employing agencies, and attended the training at different times over a period of a year in different locations in the state. It is assumed that the data is, therefore, more valid than if all 354 trainees had attended the training at the same time and worked in the same community and agency.
The follow-up portion of the study, in which a static group comparison design and "ex post facto" analysis were used, must also be considered. In an "ex post facto" situation, the experimenter attempts to replicate equivalence between study groups, after the fact, by matching group members on attributes or characteristics that could confound study results. Campbell and Stanley (1963) describe the difficulties in achieving comparability of subjects through this type of matching procedure. They also suggest that there is considerable potential for error in this methodology, even with very careful matching, due to an "uncorrectable source of undermatching," that is, the "fact of self-selection to exposure or nonexposure" to the treatment condition. They suggest that an analysis of covariance be used to reduce the likelihood of error from inherent differences in the composition of the study groups. In the current study, the analysis of covariance did, in fact, indicate that the main effect of the training condition remained highly significant, even when the effects of the covariates were considered. However, the fact that the original 354 trainees may have self-selected to attend the 24 training workshops should be considered. This variable would be particularly important if a high level of motivation to serve children with developmental disabilities prompted the trainees to attend the workshops. This level of motivation would probably also increase their willingness to provide special services to meet these childrens'
needs. By contrast, the control group subjects, who may not have elected to attend the training, might be lacking this high level of motivation.

While this is a consideration, it should be noted that the original 354 trainees may not have self-selected to attend the training. Some may have been registered by their supervisors or other agency managers. Additionally, one must consider that a high level of motivation, without specific knowledge and skill, is not likely to result in appropriately directed job behavior. Therefore, while level of motivation might have affected the trainee's performance by increasing their level of service delivery, the training itself would still have to be considered important in assuring the appropriateness of their interventions. This point is further supported by the higher level of significance of the "previous training" covariate on the behavioral criteria. While none of the matching variables exerted enough influence to significantly change the study outcomes, the subjects' previous training in developmental disabilities had a more significant effect on job behaviors than did the other matching variables.

A final consideration in evaluating the validity of the study was the use of a single rater to code the data. The rater was trained by the principal investigator in the use of the behavioral scoring criteria, and the rater achieved comparable ratings to those of
the principal investigator on the four practice protocols. Because a single rater coded all the data, any potential bias would have likely been comparable throughout the ratings. However, having additional raters score the data and determining the inter-rater reliability would ultimately strengthen the study results.

While the study supports a conclusion that inservice training significantly affects job knowledge and behavior, it must be cautioned that this conclusion cannot be immediately generalized to all inservice training programs. As discussed in Chapter II, the effectiveness of any inservice training program is largely dependent upon its ability to help an organization meet its goals and objectives, and also upon the degree to which it meets the job-related training needs of staff. The two-day workshop, Screening and Identification of Developmental Disabilities, was developed in a manner that was consistent with a systems approach to inservice training. Specifically, the objectives for the training program were derived from national policy and practice standards, which stressed the role of child welfare professionals in recognizing and serving abused and neglected children with developmental disabilities. In developing the training program, the specific job behaviors that would facilitate achievement of these objectives were clearly delineated, and the knowledge and skills necessary to perform these activities were identified. These knowledge and skills then informed the
development of the workshop training curriculum. The content of the training was, therefore, job relevant. Second, the use of a specialized training technology developed for adult education and inservice training assured that the structure of the training and the methods of presentation were appropriate to achieve the workshop goals. Specifically, factual information was presented in a manner that provided both a theoretical framework for practice and specific information that could be easily translated to the job. There was considerable opportunity for experiential learning such as groups discussion and involvement in group and individual exercises that helped trainees assimilate and practice the new skills. Finally, training time was devoted to discussions of ways the trainees could implement these skills on their jobs. The careful planning to assure the job relevance of the training was probably a significant factor in assuring the effectiveness of the training program.

The potential effectiveness of an "inservice training model" as described above in achieving specific job performance outcomes is further strengthened by the fact that no significant differences in job behaviors were found between subjects who had Bachelors' degrees and those who had Masters' degrees. One can conclude that the type of education needed to prepare workers to assume job responsibilities, related to serving children with developmental disabilities, is not provided in Masters programs.
One final point for discussion is the degree to which content knowledge correlates with job performance. While not formally stated in the hypotheses, it is of interest to note whether trainees who had greater content knowledge about developmental disabilities also performed better on their jobs. Let us return briefly to Newstrom's (1978) contention that there is typically not congruence between the four levels of Kirkpatrick's training evaluation criteria; Participant Response, Learning, Behavior Change, and System Outcomes (Kirkpatrick, 1959; 1978; 1979; 1987.) Newstrom's belief is only partially supported by the study. In comparing the experimental and control groups' test scores with their ratings on the behavioral performance variables, the Pearson Correlation Coefficient was .41. This is a moderate correlation that suggests that subjects with increased knowledge about developmental disabilities are more likely to perform better on their jobs. The moderate level of the correlation, however, suggests that variables other than content knowledge affect job performance.

In conclusion, the study results strongly suggest that while increased job knowledge cannot guarantee better job performance, the knowledge and skill transmitted by a well developed inservice training program can significantly affect trainees' performance on their jobs.
APPENDIX A

DEVELOPMENTAL DISABILITIES QUIZ
PRE-TEST/POST-TEST
DEVELOPMENTAL DISABILITIES QUIZ

TRUE FALSE ???

Please mark each of the following statements true, false, or don't know. If you can make an educated guess, do so. If you truly do not know the answer, mark "don't know."

____ __________ A great many of the factors that are known to contribute to developmental disabilities are preventable.

____ __________ Premature birth increases the probability of a child being developmentally disabled.

____ __________ Cerebral palsy is a genetic, i.e. inheritable disorder.

____ __________ Cerebral palsy involves damage to areas of the brain involved in motor functions, which renders muscles incapable of proper functioning.

____ __________ Persons with cerebral palsy are also mentally retarded.

____ __________ Brain damage during the birth process can cause cerebral palsy.

____ __________ Severe child abuse involving injuries to the head can cause cerebral palsy.

____ __________ The symptoms of cerebral palsy are usually apparent before school age.

____ __________ Most developmentally disabled children learn better when placed in a special school or class where their needs can be addressed, rather than having to compete with non-disabled children in a regular classroom setting.
Athetoid, spastic, and ataxic refer to various types of epilepsy.

Autism is a condition almost always diagnosed at birth.

Autism affects boys almost four times as often as girls.

Autism is a psychological mental illness which can usually be traced to psychosis in the family, severely deprived environments, and emotional coldness on the part of parents or caretakers.

Autistic children are usually described by their parents as affable and extremely sociable infants.

If a worker suspects that a child may be delayed or have a disability, he should wait until he is certain a problem exists before referring the child for professional assessment.

Autistic adults are often institutionalized or otherwise viewed as psychotic or mentally retarded.

In spite of the identification of more than 200 specific factors as possible causes of mental retardation, in 80% of the cases, the cause cannot be determined.

Idiot, imbecile, and moron were at one time clinically accepted diagnostic categories referring to various levels of mental retardation.
Almost 90% of persons with mental retardation are minimally retarded, falling into the "mild" or least retarded category.

A developmental disability is usually a permanent condition that will require a lifetime of very specialized care and prevents an individual's full participation in normal activities.

Frequently mental retardation is not diagnosed until a child reaches school age, even though signs may have been present long before.

Down Syndrome is most often a genetic accident, resulting in a chromosomal abnormality that usually occurs during reproductive cell development.

Down Syndrome is much more likely to occur in the children of very young mothers who have not yet reached physical maturity, or in older mothers beyond the age of 36.

Amniocentesis is a test which is used to determine the presence of many fetal abnormalities, including Down Syndrome.

Phenylketonuria (PKU), an inherited metabolic disorder that causes progressive mental retardation, has no effective cure or treatment.

Even with fairly severe disabilities, early intervention can prevent unnecessary deterioration.

Approximately 50% of persons diagnosed with epilepsy can expect to be seizure free with proper medication.
Epilepsy that first appears in early childhood may disappear in later years.

A child with uncontrolled epileptic seizures may develop a degree of mental retardation as a result.

Mongolism is a proper term for an abnormality in the 21st chromosomal group.

There are dozens of known causal and contributing factors that have been associated with developmental disabilities. List ten in the spaces below.

a) ................................................................. f) .................................................................
b) ................................................................. g) .................................................................
c) ................................................................. h) .................................................................
d) ................................................................. i) .................................................................
e) ................................................................. j) .................................................................

Would you be concerned that the children described below might be developmentally disabled? Check yes or no for each child. If you check yes, name the possible disability(ies) in the space to the right.

YES NO

A mother complains to you that she has a hard time separating her baby's legs to diaper him; they are stiff and tight.

A mother says her baby will not relax in her arms, is fretful, and seems to prefer lying alone to being held.

A six-month old child is not yet sitting up unassisted; he wobbles and topples over.

A ten-month old child always finger feeds herself and reaches for things with her right hand, showing a decided preference for her right hand.
A two-year (24 month) old child speaks only single words, never in phrases or sentences.

A mother says her three-year old is a chronic daydreamer; he often stares off into space and doesn't seem to hear her when she talks to him.

A two-year old child still crawls up and down stairs on her hands and knees.

A two-month old baby seems unable to focus his eyes and follow a moving object (tracking.)

An 18-month old child will watch you hide his favorite toy but will not make an effort to go after it, and seems to forget about it once it is gone.

Pictures/Slides

1)
2)
3)
4)
5)
APPENDIX B

BEHAVIOR INVENTORY INSTRUMENT
INTERVIEW QUESTIONNAIRE
SUBJECT INTERVIEW FORM

These questions should all be answered from the perspective of the subjects' jobs as child welfare caseworkers or supervisors, even if they are not currently working in this capacity.

Question #1:

Have you had children on your caseload in the past nine years who had developmental disabilities (or supervised cases on your workers' caseloads in which children had developmental disabilities)?  ____Yes  ____No

(If Yes)

Can you tell me about the children? Use their first names only. Tell me what kind of disability each child had, and describe the nature of the condition, or the child's symptoms. Please name as many children as you can remember.

Question #2.

Have you ever suspected that a child on your caseload (or on your caseworkers' caseloads) might potentially have a developmental disability?  ____Yes  ____No

(If Yes)

What disability did you think the child had? What were the symptoms that made you think the child had a disability? Name as many children as you can remember.
Question #3.

In the past 9 years, have you ever personally initiated referral (or instructed your caseworkers to initiate referral) of a child to a psychologist or other mental health professional for a full developmental assessment because you suspected the child had a developmental disability or delay? ___Yes ___No

(If Yes)

Using first names only, tell me about the child for whom you made the referral. To whom did you refer the child? What exactly did you want the assessment to tell you? What particular questions did you want answered? Be specific. Name as many children as you can remember.

Question #4.

Have you ever personally initiated (or instructed your caseworkers to initiate) a referral of a child with a developmental disability or their family to a community agency for specialized services, such as special education, medical care, recreation, mental health, or other developmental services? ___Yes ___No

(If Yes)

Name the child, using the first name only, describe the child's service needs. Name the agencies to which you referred the child and describe the nature of services you were seeking.

Did you do any follow-up after you made the referral? Explain how you did this, and what you were trying to achieve by following up. What did you hope to accomplish?
Question #5.

Have any of the parents, foster parents, or other caretakers of children with developmental disabilities on your caseload (or your workers' caseloads) ever needed special emotional support or respite services to help them care for their children?  
____Yes  ____No

(If Yes)

What did you do personally in your casework with these parents to help them (Or what did you instruct your caseworkers to do in their direct work with these parents?)

Did you refer them to other providers, including support groups or respite providers, for special support and help? To whom did you refer?

Question #6.

Have you ever had to act as an advocate (or instruct your caseworker to act as an advocate) for a child on your caseload with a developmental disability to make sure that the child's rights were maintained, or to procure special equipment or services, to assure that services were properly delivered, or to improve the child's quality of life?  
____Yes  ____No

(May prompt, if needed, with examples; working with SSI to procure a wheelchair, attending an IEP or IHP meeting to assure proper educational planning, or working with a school system to mainstream a child with a disability.)

(If Yes)

Describe the situation and your advocacy activities. What did you do? What agencies or service providers did you have to work with?

Have you ever had to "take on a whole service system" in order to get your child's needs met? Describe the situation.
APPENDIX C

BEHAVIOR INVENTORY INSTRUMENT
SCORING CRITERIA
BEHAVIOR INVENTORY INTERVIEW AND SCORING CRITERIA

**Criterion Behavior #1:** Accurate recognition of children on the caseworker's caseload who had conditions that can be considered developmental disabilities.

**Question:** "Have you had children on your caseload who had developmental disabilities (or supervised cases on your workers' caseloads in which children had developmental disabilities)?" ___Yes ___No

(If Yes)
"Can you tell me about the children? Use their first names only. Tell me what kind of disability each child had, and describe the nature of the condition, or the child's symptoms. Please name as many children as you can remember."

**Scoring Criteria**

0 Does not remember having had children on the caseload with developmental disabilities; or, names only, or primarily, conditions that cannot be considered developmental disabilities, indicating an inability to discriminate between DD and other types of childhood disorders.

1 Can identify a child or children, and describes developmental problems without being able to name the child's disability or underlying problem; or, names an appropriate condition with no description of the symptoms or characteristics.

2 Can identify a child or children, properly names conditions that are considered developmental disabilities, and accurately describes characteristics of the condition, reflecting a knowledge of the nature and defining characteristics of the condition. (Does not have to name mental retardation if children are below the age of six and are not diagnosed.)
Criterion Behavior #2: Recognition of suspicious early warning signs of developmental disability or developmental delay in children on their caseloads.

Question: "Have you ever suspected that a child on your caseload (or on your caseworkers' caseloads) might potentially have a developmental disability?"  Yes  ____No

(If Yes) "What disability did you think the child had? What were the symptoms that made you think the child had a disability? Name as many children as you can remember."

Scoring Criteria:

0  Cannot remember any child they thought had a disability, or indicates they have not suspected any child had a disability, or describes medical problems, behavior problems or conditions that cannot be considered developmental disabilities.

1  Remembers having identified children with developmental delays or abnormal developmental patterns, and suggests suspicion because the child "was slow," "there was just something wrong, he wasn't developing like he should" or names a single developmental problem (i.e. "he wasn't walking, his speech was delayed."

2  Can accurately describe specific early warning signs that are particular to certain disabilities (i.e. "he was stiff as a board, rigid in his lower body, and wouldn't relax") or describes several developmental characteristics that are consistent with a particular condition (at 8 months he wasn't rolling over, sitting up, or reaching.) May or may not be able to identify or name the child's potential disability.
**Criterion #3:** Appropriate referral of children to other agencies or professionals for developmental/diagnostic assessment.

**Question:** "Have you ever personally initiated a referral of a child to a psychologist, a health or mental health professional, or other professional for a full developmental assessment because you suspected the child had a developmental disability or delay?" ___Yes ___No

(If Yes) "Using first names only, tell me about the child(ren) for whom you made the referral. To whom did you refer the child(ren)? What exactly did you want the assessment to tell you? What particular questions did you want answered? Be specific. Name as many children as you can remember."

**0** Does not remember having been responsible for initiating a diagnostic assessment of a child, or has referred a child with problems other than developmental disabilities for assessment (such as emotional problems or behavior disorders with no other characteristics of DD

**1** Has made referrals, or has directed others to make such referrals, for developmental assessment of children accurately identified as potentially DD through description of the child's conditions; made the referral "to confirm a diagnosis," "to find out if the child had a problem," or "to get more information;" or, describes in detail an appropriate referral process used for children, but cannot remember specifics about the child(ren) for whom the referral was made.

**2** Has made appropriate referrals of specific children, or directed others to make such referrals, and can identify referral questions that indicate the use of assessment information for specific case/treatment planning; "I wanted to know what the best placement for him was," "I wanted to find out what we should be doing to enhance his development." Referral to a specialized clinic (i.e. the Myelomeningocele clinic at Children's Hospital) for a comprehensive "work up" prior to treatment planning is also appropriate.
Criterion # 4: Appropriate service linkage to specialized community based services to meet service needs of children with developmental disabilities and their families, and service management/coordination.

Question: "Have you ever personally initiated (or instructed your caseworkers to initiate) a referral of a child with a developmental disability or their family to a community agency for specialized services, such as special education, medical care, recreation, mental health, or other developmental services?"  ___Yes  ___No

(If Yes)
"Name the child, using the first name only, and describe the child's service needs. Name the agencies to which you referred the child, and describe the nature of services you were seeking. Did you do any follow-up after you made the referral? Explain how you did this, and what you were trying to achieve by following up. What did you hope to accomplish?"

Scoring Criteria:

0  Has not referred children with DD or their families for specialized services for DD-related needs; or, has suggested to parents that they seek services for their children with no follow-up or direct involvement in the referral.

1  Has made a referral, or directed others to make a referral, of children with DD or their families to an appropriate service provider or agency "to get them help" or "to get services for the problem;" or, has made a referral for a single identified problem (i.e. to the speech clinic for speech therapy). Did minimal follow up, and does not demonstrate case management or coordination of services.
Has referred, or has directed others to refer, children and families to appropriate service agencies for comprehensive services. Indicates an awareness of children's multiple service needs. Has followed-up referrals to evaluate services or to assure that they were being provided. Includes referral to an agency that provides comprehensive case management services, such as a County Board/or Early Intervention Program.

**Criterion #5:** Provides or arranges respite care, special education, or emotional support services to parents, foster parents, adoptive parents, or other caretakers of children with developmental disabilities.

**Question:** "Have any of the parents, foster parents, or other caretakers of children with developmental disabilities on your caseload (or your workers' caseloads) ever needed special emotional support or respite services to help them care for their children?"  

(If Yes)  
"What did you do personally in your casework with these parents (or instruct your caseworkers to do) to help them? Did you ever refer them to support groups or respite providers? To whom did you refer?"

**Scoring Criteria:**

0 Has not had parents or caretakers who needed special or supportive services, or suggests "they just needed someone to talk to now and then."

1 Has made referrals for respite care, or has suggested to parents that they seek respite care, with no other supportive interventions; or, describes appropriate supportive services they have provided without being able to remember the specifics of the case situation.
2 Can describe supportive interventions provided to parents, including arrangement of homemaker to go into the home, direct supportive counseling specific to helping the parent understand how to meet the child's needs; referring the parent to educational programs to learn how to care for the child; has provided respite for the parent by taking the parent out for a while; or has combined a referral for respite services with other supportive interventions.

Criteria #6: Has advocated for children with developmental disabilities and their families with other community service providers to assure that services needs were met, services were delivered, and rights were not violated.

Question: "Have you ever had to act as an advocate (or instruct your caseworker to act as an advocate) for a child with a developmental disability to make sure that the child's rights were maintained, to procure special equipment or services, to assure that services were properly delivered, or to improve the child's quality of life?" ___Yes ___No

(If Yes)
"Describe the situation and your advocacy activities. What did you do? What agencies or service providers did you have to work with? Have you ever had to 'take on a whole service system' in order to get your child's needs met? Describe the situation."

Scoring Criteria:

0 Has not provided advocacy for a child on the caseload with developmental disabilities, or describes single efforts to provide linkage to services that would be considered a normal part of the caseworker's job.

1 Can describe situations in which the worker assumed the role of an advocate for the child, but for which no special effort was required. For example: attending an IEP meeting at school; making telephone calls to arrange for procurement of equipment or services.
Describes situations in which multiple advocacy activities were performed for a single child; or, in which a lengthy amount of time, "creative" effort, or "stick-to-it-ive-ness" was required; or, describes conflictual situations in which the worker had to advocate to protect the child's rights or to assure the provision of services; or, referred the child to an agency that provides comprehensive advocacy services for persons with MR/DD (such as Ohio Legal Rights.)
Subject Data Sheet

The following information will be used only for matching subjects in experimental and control groups, and will remain confidential.

Name ______________________ Assign Subject No. ______
Agency ____________________ Current Position ______ How Long? ______
Total # Years Of Experience in Child Welfare _______________
Level of Education: ______ High School ______ Bachelors' ______ Masters ______ Other

In what field of study is your degree? ____________
What is your age? ______

Have you ever been employed by an agency or in a unit designed to serve persons with developmental disabilities, such as a County Board of MR/DD, a Developmental Center, a special education classroom, or a crippled children's center? Yes ______ No ______
If yes, please name the agency and briefly describe your job.

__________________________________________________________

Do you have an immediate family member or close friend who has a developmental disability? If so, what is the nature of the disability, and what is your relationship to the person who has the condition?

__________________________________________________________
How much training have you had in developmental disabilities? Do not include the 2-day inservice training workshop on Screening and Identification of Developmental Disabilities that was offered by Ms. Rycus. (Check one)

____ Little or No Training

____ Some Training (such as, a brief overview in a college class, like child development, or an inservice training workshop.) Please describe the type of training and the content.

________________________________________________________________________

____ A Lot (A college specialization in a field related to developmental, disabilities, such as special education, abnormal/developmental psychology, or handicapped children; or, multiple inservice training workshops.) Please describe the type of training and the content.

________________________________________________________________________

________________________________________________________________________

How would you characterize your level of interest in the topic of developmental disabilities? (Circle One)

No Real 
A Little 
Average 
Very 
Extremely
Interest 
Interest 
Interest 
Interested 
Interested
References


