IDENTIFYING AT-RISK EARLY PRIMARY STUDENTS: GLOBAL, ACADEMIC AND SPECIFIC SKILLS ASSESSMENTS

by

ROSALYN FRANCES FRANCIS

B.Ed., The University of British Columbia, 1970

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ARTS

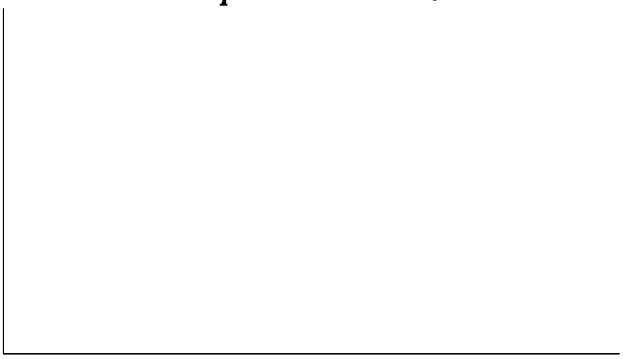
in

THE FACULTY OF GRADUATE STUDIES

Department of Educational Psychology and Special Education

We accept this thesis as conforming

to the required standard



THE UNIVERSITY OF BRITISH COLUMBIA

March, 1993

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

(Signature) _____

Department of Educational Psychology & Special Education

The University of British Columbia
Vancouver, Canada

Date April 7, 1993

DE-6 (2/88)

ABSTRACT

The primary purpose of this study is to investigate the extent to which teacher ratings are consistent in classifying kindergarten students when using measuring instruments that range from a broad global classification to a specific skills rating system. A secondary purpose of this study is to validate a Teacher-made Checklist of skills intended for kindergarten or early primary children. The problem that the study addressed was, if teachers apply their own standards in making global assessments of children's performance, to what extent are these ratings consistent with subsequent specific ratings using several global items or several specific ratings of skills and abilities?

The study was conducted over three phases. In the first phase, the teachers were asked to categorize the students in one of their kindergarten classes as high, average or low. Secondly, teachers were asked to rate their students using the Academic Competence Questionnaire subtest from the Social Skills Rating System (Gresham & Elliott, 1990). In the third and final phase teachers were asked to rate their students using a skill specific Teacher-made Checklist. A total of 17 kindergarten teachers participated in the study. They rated 336 kindergarten students.

A series of correlational analyses were computed in order to establish validity of the Teacher-made Checklist. As well, a procedure using Hierarchical Cluster Analysis was utilized in order to form homogeneous groups of students and to assist in making predictions of at-risk and low-risk groups. A series of

one-way ANOVA's were computed to test if significant differences existed between the clusters and post hoc analysis using Tukey procedure was used to test for the extent of the difference between the groups. Cross-tabular analyses were performed to investigate differences in grouping patterns among the three phases.

Although teachers initially categorized the students into three groups, four groups emerged using Cluster Analysis in Phases 2 and 3. The Teacher Checklist using the subscales appeared to identify the students into four groups most clearly. The data and analyses indicated that teachers are uniform in classifying students, however teachers need guidelines to assist them in identifying those students who may be at-risk for school learning problems. A follow-up study to further validate the Teacher Checklist is needed.

William T. McKee, Ph.D.
Research Supervisor

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENT

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

First, I wish to express my sincere appreciation to Dr. William McKee, my research supervisor, for his unwavering support and his thoughtful advice throughout the course of this research. His zeal for the pursuit of academic excellence was modelled throughout our association. I am also especially grateful to Dr. Nand Kishor for his efficiency, his valuable feedback and for his persistent challenge to "get it done". I gratefully acknowledge the contribution of Dr. Jacquelyn Baker-Sennet for her support and for the insightful feedback on my drafts.

My thanks goes to the Abbotsford Kindergarten teachers who so willingly took the time to participate in this study, and to my colleagues and friends who constantly encouraged me to complete the study.

An expression of thanks goes to my daughter Nicole who assisted in verifying data and who later supported from afar, to my son-in-law Tom Kompas for his thoughtful advice, interest and encouragement. I gratefully acknowledge my daughter Johanna Kompas for her unwavering support and inspiration throughout the stages of writing to the finished product.

Finally, I give special thanks to my husband Roy for his love and encouragement. I appreciate his patience during my long hours of study as well as the many times he filled in for me so that I could pursue part-time studies.

CHAPTER 1

INTRODUCTION

In recent years there has been an increasing emphasis on the early identification of and intervention with academically at-risk primary school-age children. American Federal Policy in 1986 mandated early childhood intervention (Meisels, Harbin, Modigliani & Olson, 1988). As well, research findings point to the effectiveness of intervention programs for developmentally handicapped children and young children who may be at-risk for school learning problems (Dunst, Snyder & Mankinen, 1987; Castro & Mastropieri, 1986; Lazar, Darlington, Murray, Royce, Snipper, 1982). Effectiveness of intervention is related to the early provision of appropriate programs.

The major concern in providing early intervention programs is reliably identifying those children in need of individualized programs or special education in the first few years of school (Diamond, 1990). Reliability in screening can be increased by providing staff development and training, by correlating the screening battery with the type of program being offered, and by informing parents as to the purpose of the screening program.

One problem with the use (or abuse) of standardized tests with young children is that they are often used to label children who have not yet had an opportunity to learn (Bredekamp & Shepard, 1989). Despite this potential difficulty, there is evidence to indicate that early identification combined with

effective remedial intervention can help to reduce the risk of academic failure (Simner, 1983; Becker & Gersten, 1982; Darlington, 1981; Lazar, Darlington, Murray, Royce, Snipper, 1982). Risk of academic failure is not a condition but rather a circumstance to indicate an increased probability that a disorder will occur.

Furthermore, children come to school with varying background experiences, interests and aptitudes. Although they do not come with a label,

> they do not 'catch' a disability as one would catch a biophysiological disorder. Rather, they bring with them as yet unspecified cognitive, behavioral, linguistic, and affective characteristics developed over time, which interact with the academic and social ecology. Out of this interaction comes school success or, in about 10% of children, failure and referral for special education (U.S. Department of Education, 1988 in Cooper & Speece, 1990).

Typically, learning problems are not identified until children have experienced repeated failures in school (Diamond, 1990; Heward & Orlansky, 1988), yet learning disabilities can be detected with reasonable accuracy early in development (Siegel, 1988). The emotional trauma associated with academic failure compounds the learning problem. Those children who are not detected "are often the high-incidence low severity problems such as learning disabilities, behaviour difficulties, and mild delays in intellectual functioning that are economically and physiologically costly but may be ameliorated if remediation is provided early in development" (Siegel, 1988, p.115). Therefore, valid developmental screening tests are needed to identify those students in need of further diagnosis for potential learning, health, or developmental problems

(Meisels, 1985 in Bredekamp, & Shephard). Standardized tests play a vital role in identification and further diagnosis (Bredekamp & Shephard, 1989). Screening for the purpose of intervention has been endorsed (NAEYC, 1988), however, considerable criticism has surrounded the extensive use of standardized testing of young children notably when the purpose is either to retain or to exclude children from programs (Meisels, 1987; Shepard & Smith, 1986). Extensive use of standardized measures can be a drain on resources as well as time; conversely, standardized tests can be useful if guidelines, as well as standards for educational and psychological testing are adhered to (NAEYC, 1988; Bredekamp & Shepard, 1989). Standardized tests used for screening for readiness are useful if the purpose is to diagnose children's needs and to plan programs. In addition to standardized tests, assessment methods, such as anecdotal observations, teacher-developed checklists, and systematic records provide useful information (Bredekamp & Shepard, 1989; NAEYC Statement, 1988).

Because of changes in American educational policy, according to Bracken,

> there will be a significant increase in the number of preschool children who undergo psychoeducational assessment in the future. Concomitant with the increase in preschool assessments, there needs to be increased professional attention paid to the quality of instruments used in school assessments (Bracken, 1987, p.326).

Problems with preschool screening instruments exist. These include the psychometric properties, (Meisels, 1985; Diamond, 1990) the predictive validity, and the accuracy with which they classify children as to whether or not they

are at-risk for significant problems in development, learning or behaviour (Meisels, 1985; Diamond, 1990). If screening for early identification is to have positive utility, we must have high quality instruments with demonstrated reliability and validity. For example, the Revised Denver Developmental Screening Test did not identify as at-risk 55% of the children who were later identified by school personnel as at-risk and who were in need of help (Ysseldyke, Thurlow, Graden, Wesson, Algozzine & Deno, 1983). Because young children develop at different rates, it is difficult to reliably identify all children who may be at-risk. Furthermore, it is difficult for trained psychologists to differentiate between learning disabled (LD) and low achievers as there are no reliable psychometric differences in test profile scores. Often small quantitative differences are found but not qualitative (Ysseldyke, et al. 1983).

The problem with attempting to classify each child is that some children who do not have learning problems are misclassified (Diamond, 1990). However, if the purpose of identification is to provide a stimulating intervention program, there is no evidence that harm is done if some children without learning problems are misclassified (Cornell & Gottfried, 1976; Mercer, Algozzine & Trifiletti, 1988). Many researchers argue the negative consequences of a false positive decision do not seem to be a major problem (Siegel, 1988, p.120).

As well as a need for better quality instrumentation, there is a need for astute practitioners who can discern high quality instruments and practices

(Bracken, 1987, p.326). According to studies conducted by the Minnesota Institute on Learning Disabilities, many tests that have been used for psychoeducational decision-making were technically inadequate in identification of the LD student (Ysseldyke, et al. 1983).

Teacher judgment in identifying those children who require early intervention plays a vital role in the identification and decision making process. As noted by Cadwell and Jenkins (1986), teacher judgment can be directly related to the teacher's ability to organize semantic information. That is, when teachers are asked to rate students, they may do so based on memory. The use of memory-based reports is influenced by the rater's organization of their semantic memory. As a result, similar descriptions of different behaviours may be recalled because of their closeness in meaning and not actually because of the occurrence of the behaviour (Cadwell & Jenkins, 1986; Kishor, 1992). Also there is some evidence to indicate that teacher judgment can be biased (Cadwell & Jenkins, 1986; Algozzine & Curran, 1978), or that the current placement system considers teachers to be suspect judges of academic performance in classrooms (Gresham, Reschly, & Carey, 1987; Gerber & Semmel, 1984).

Other studies have indicated teacher judgments of classroom performance in classifying students as LD or non-LD were highly accurate (Gresham, Rechly, & Carey, 1987). There is also agreement among teachers as to desirable student attributes and teachers were consistently able to define the

characteristics of the "idealized teachable" students. It was noted that teachers have perceptions of the ideal student attributes and use these perceptions as an internal standard by which to base judgements about school success or failure (Kornblau, 1982). Furthermore, "teacher's ratings reflect more than a simple halo effect and, indeed, do discriminate among different components of the child's functioning" (Feshbach, Adelman & Fuller, 1974, p.51). In addition, teachers ratings are "most effective in identifying those children in need of intervention and those not likely to need special programming" (Mercer, Algozzine & Trifiletti, 1988, p. 185) and according to Gresham, Reschly and Carey (1987), "teacher judgments served as a major criterion against which Binet's test was validated" (p.543).

A. Purpose of the study

The primary purpose of this study is to investigate the extent to which teacher ratings are consistent in classifying kindergarten students when using measuring instruments that range from a broad global classification to a specific skills rating system. This study will examine teacher ratings based on the following: the teachers' broad categorization, a standardized questionnaire, and a Teacher-made Checklist to determine if teacher ratings are consistent when guidelines for rating become more specific. The problem this study will address is, "If teachers apply their own standards in making global assessments of children's performance, to what extent are these ratings consistent with

subsequent specific ratings using several global items or several specific ratings of skills and abilities?"

A secondary purpose of this study is to validate a Teacher-made Checklist of skills intended for kindergarten or early primary children.

Questions that this study will attempt to address are:

(1) How does the Checklist compare to acceptable reliability standards?

(2) To what extent do each of these measures: Teacher Rating, Academic Competence Questionnaire, and Teacher Checklist identify children as being at-risk for learning problems?

(3) What is the relationship between the Academic Competence Questionnaire and the Teacher Checklist at the subscale and full scale levels?

(4) What is the evidence of the validity for teachers' ratings of at-risk students?

B. Justification of the study

Research indicates that there is an optimum time for intervention in readiness skills (Oliver,1990; Simner, 1983; Diamond, 1990; Mercer, Algozzine, Trifiletti, 1988). Since teachers initially identify those students who may be at-risk, and who require an indepth psychoeducational assessment, it is essential that teachers are equipped to effectively screen children. Because psychoeducational assessment is costly and time consuming, it may be possible

for teachers to identify children who are at-risk and provide appropriate pre-referral intervention for at least some students without extensive formal assessment. Based on these factors, there is a movement towards more authentic assessment (Ysseldyke, et al. 1983). Effective warning signs of school failure include: in class attention span, distractibility or memory span, in class verbal fluency, interest and participation, letter or number identification skills and printing errors (Simner, 1983). These warning signs noted by Simner's review of several studies, may best be initially identified by teacher observation. Also, with the implementation of the New Primary Program, there is a growing need for the development of accurate benchmarks. It is intended that this study will provide information that will assist teachers in identifying those children who are in need of intervention and/or further assessment.

CHAPTER 2

REVIEW OF THE LITERATURE

A. Introduction

This chapter provides a review of the literature on the historical development of early identification and intervention of students who may be classified as at-risk for learning difficulties in school. Predictors of children at-risk, advantages, and issues in early identification will also be reviewed. In addition, literature pertaining to teachers' judgment as well as recommendations for future identification initiatives will also be presented.

B. History of Early Identification

Early screening and identification has drawn increasing interest since the mid 1960's (Lee, Unruh, Ford & Mace, 1991). Historical roots can be traced to projects such as Head Start, The Early and Periodic Screening and Diagnosis Treatment (EPSDT) Program, (Lee, et al. 1991), as well as the passing of P.L. 94-142 (Reynolds, Gutkin, Elliott & Witt, 1984) which sanctioned the right of all handicapped children in the United States to a free public education including related services to meet their needs. The passage and implementation of P.L. 99-457, the education of the Handicapped Act Amendments of 1986, which included the rights and protection of all infants, toddlers and preschool children also generated interest in early identification. The passing of these laws

increased the need for reliable instruments to identify young children who are at-risk for developmental, cognitive, and academic delays (Bracken, 1987; Harbin, 1988; Harrington & Ford, 1988; Lee, et al. 1991).

There has been a growing concern over the identification and treatment of learning problems in order to facilitate learning (Horn, & Packard, 1985) and to prevent future emotional and behavioral problems (Black, 1974; DeHirsh, Jansky & Langford, 1966; Fincham, 1977; Horn & O'Donnell, 1984). Many academics have assumed that behavioral and emotional problems associated with learning problems are the results of negative learning experiences in school (Douglas & Peters, 1979) and 'acting out' behaviour may be the result of failure and frustrating experiences in school (Cunningham & Barkley, 1978).

Early identification refers to the implementation of a formal plan for identifying a disability as early as possible in a child's life (Reynolds, Gutkin, Elliot & Will, 1984). Its purpose is to identify children who are likely to have difficulty in the education system and to provide individualized instruction (Stone, Cundick, Swanson, 1988). Information received from screening is critical in helping educators promote children's successful learning (Judy, 1986).

C. Predictors of children at-risk

Because early identification and intervention efforts are necessarily limited by the reliability and accuracy of measures predicting children at-risk of future learning problems, much of the empirical literature has been concerned

with identifying early predictors of later school success and failure (Horn & Packard, 1985). In a meta-analysis conducted by Horn and Packard, one of the best predictors of early school achievement was attention-distractibility. Other predictors of achievement during the elementary school years were internalizing behaviour problems, tests of general cognitive functioning, as well as language variables (Horn & Packard, 1985; Simner, 1983). It is widely assumed that the ability to focus attention and maintain one's attention for long periods of time is a necessary prerequisite for success in school (Barkley, 1981; Feshbach, Adelman & Fuller, 1974; Horn & Packard, 1985). Other effective warning signs of school failure are: in class interest and participation, letter or number identification skills, and printing errors (Simner, 1983).

Based on several correlations of kindergarten measures with later academic achievement, Simner found that several warning signs, which are frequently used by educators are not as effective in predicting school failure. These are: gross and fine motor co-ordination, peer acceptance, adult co-operation, basic language skills, as well as drawing and copying errors (Simner, 1983).

It was also found that some specific classroom environments were significantly associated with school failure (Cooper & Speece, 1990) and relatively little emphasis had been placed on the context of instruction before children's disabilities were formally identified (Messick, 1984). There is also variation in the manner first grade teachers react to, evaluate, and interpret

predictive student characteristics. It has been suggested that first-grade teachers require, prefer, tolerate, and react individually and differently to student attributes and thus a student could be considered to be at-risk in one first-grade classroom and not at-risk in another first-grade classroom (Fedoruk, & Norman, 1991).

D. Advantages of Early Identification

The behaviour of young children is more susceptible to change than that of older children, therefore, the effectiveness of early intervention efforts is enhanced by early identification (Hayden, 1974; Stone, Cundick, & Swanson, 1988). Proponents of early intervention claim that there are sensitive periods in the early development of the central nervous system and a greater neural plasticity in younger organisms, thus early positive specific experiences tend to be beneficial, (Anastasiow, 1990; Guralnick, 1991) and have lasting influence (Bowlby, 1969).

A screening battery provides information about a child's area of strengths as well as weaknesses. It provides information that the teacher can use to help focus on the child's area of strength and work through the strengths to help develop the underdeveloped modalities (Judy, 1986).

Based on a meta-analysis of 52 studies, Castro and Mastropieri (1986) found that early intervention programs are generally effective. The four most commonly cited variables in the studies were: parent involvement, age at start

of intervention, degree of structure, and duration/intensity of the intervention. Castro and Mastropieri conclude that even though more parental involvement is better and parents can be effective interveners, they are not essential to intervention success. A majority of the studies indicated earlier intervention was better, however there was a significant trend to indicate that an optimum time for intervention existed. That is, some studies indicated that there is some evidence that for the handicapped population, children who start later do better. The authors also suggest that highly structured programs, with longer and more intense programs were more effective (Castro, & Mastropieri, 1986).

Robert Book (1980) conducted a four year follow-up study of a group of at-risk children. He found that the "individual student's group test performance did not change significantly over the first four years", that the at-risk group performed consistently below the low risk students and that the at-risk group required continuous intervention and support throughout their education (Book, 1980). However, other studies indicated that the rates of progress differ with the severity of the impairment (Dunst, Snyder & Mankinen, 1987).

E. Issues Surrounding Early Identification

Young children's changing mental, physical, and emotional conditions tend to produce discrepant results for the same test given at different times (Judy, 1986). In addition, measured intelligence is affected by environmental conditions and "there is instability associated with preschool psychoeducational

instrumentation and assessment conditions" (Bracken, 1987, p.313).

Test results do not link identification with intervention and do not provide the teacher with information that is useful for planning effective instructional programs (Howell, 1986; Stone, Cundick, & Swanson, 1988; Tucker, 1985). "Unless the end result of the screening is a more effective education program for the child, even the minimal costs involved do not justify its use as a means of identifying children" (Book, 1980, p. 158).

Because of the changing economy, funding for education is becoming more difficult, therefore educational practices such as individual identification and classification which requires extensive time especially that of the school psychologist are being reduced (Lacayo, Sherwood & Morris, 1981; Stone, Cundick, Swanson, 1988). Critics of standardized assessment claim that time and resources spent in teaching and instructing would be used more appropriately than time that is spent in testing and labelling (NAEYC, 1987; Ysseldyke, Thurlow, Wessons, Algozzine & Deno, 1983).

Screening tests are not accurate enough to identify children for placement into special programs (Shepard & Smith, 1988). The tests are commonly selected on the basis of inadequate data, face validity, testimonial evidence and frequency of use by others (Lichtenstein, 1981). Only reliable and valid tests that comply with the joint committee's AERA (American Educational Research Association), APA (American Psychological Association) and NCME (National Council on Measurement in Education), 1985 technical standards for

test construction and evaluation should be used (Bracken, 1987; NAEYC, 1988). The intended use of the test should be considered carefully when selecting instruments for preschoolers (Bracken, 1987). "Computations of sensitivity and specificity are the primary means of evaluating a screening test's capacity to correctly identify individuals" (Diamond, 1990, p. 153). Sensitivity is measured by the proportion of individuals at-risk who are correctly identified by the test, and specificity refers to the proportion of individuals not at risk who are excluded from further assessment (Kochanek, Kabacoff, & Lipsitt, 1988).

Although standardized tests play an important role in providing assessment information, they are only one of multiple sources of information that should be used when making decisions as to what is best for a child's education (NAEYC, 1988).

A child's learning deficiencies may be identified early, however appropriate teaching programs may not be available (Lerner, 1981) consequently the optimum time period for intervening with the child, which is before a child experiences failure may be missed (Strag, 1972).

F. Teacher's Judgment in Identification

In 1947, Kottmeyer suggested that teachers are able to predict reading problems better than the Metropolitan Reading Test (Mercer, Algozzine, Trifiletti, 1988). In their 1974 study, Feshback, Adelman & Fuller found that teachers are able to discriminate among different aspects of a child's

functioning. They found that kindergarten teachers' rating "can predict first-grade reading achievement at least as efficiently as a psychometric battery that has been designed for this purpose" (Feshbach, Adelman & Fuller, 1974).

Gresham, Reschly, and Carey (1987) evaluated the accuracy of teacher judgment in identifying students as learning disabled or non-handicapped and then compared these teacher judgments with standardized tests of intelligence and achievement. This study indicated that "regular classroom teachers are defendable 'tests' of student achievement and perhaps should be used as one of the criteria by which psychoeducational tests are validated" (Gresham, Reschly, Carey, 1987, pp. 551-552). It was pointed out by Kornblau that teachers "probably have a priori perceptions of the characteristics of 'idealized teachable' pupils" and it is likely that these perceptions are used as a model when rating students (Kornblau, 1982, p. 173). Other research indicates that there is agreement among teachers regarding their perceptions of desirable pupil attributes (Kornblau, 1982; Solomon & Kendall, 1977).

Other studies indicate that teacher bias exists when human judgments are used for rating (Greenwood & Morton, 1989; Salvia & Ysseldyke, 1981). Bias in rating children has been reported to be influenced by such rater characteristics as: age, sex, professional role, and whether or not the rater had children (Greenwood & Morton, 1989; Marsh, Stoughton & Williams, 1985).

Classroom teachers with different levels of tolerance for selected behaviours were asked to rate a hypothetical child who exhibited disturbing

behaviours in a study conducted by Algozzine and Curran (1978). Results of this study indicated that students with various negative behaviours (i.e., social defiance like disobedience or disruptiveness) are more tolerated by some teachers than by others, which consequently generates differing student-teacher interactions. It was suggested that an effective educational strategy would be to match individual children with teachers based on the teacher's tolerance level for certain behaviours (Algozzine & Curran, 1978).

When teachers are asked to rate students on their ability, academic performance, and classroom behaviour or effort, each teacher "serves as the 'instrument' that transforms the information obtained through numerous interactions with students into a set of ratings" (Cadwell & Jenkins, 1986, p.460). Based on their study of teacher ratings, Cadwell and Jenkins concluded that "rating is a schema-based process constrained by the rater's information-processing abilities" (Cadwell & Jenkins, 1986, p.471) and teacher ratings contain bias. It was also concluded that although teachers can provide information about a student's behaviour, this information should not be interpreted as if teachers "were direct measures of student behaviour" (Cadwell & Jenkins, 1986, p.472).

Research conducted by Kishor, also indicated that "teachers' judgment of students' performance may depend on how teachers use performance information" (Kishor, 1992). This study was based on Kelly's (1967) model of causal judgment. Kelly proposed that a person uses some combination of

consensus, consistency, and distinctiveness information. Considering these terms in relation to a regular class, consensus information reveals how a student performs in comparison to the rest of the class. Consistency information refers to the stability of performance over time, whereas distinctiveness information refers to how the student performs in one subject in comparison to the other subjects. For example, if the teacher was rating the child in reading, consensus information would deal with how the student performed in reading in relation to the rest of the class. Consistency information would reveal how a student performed over several reading tests and distinctiveness information would reveal how the student's performance in reading compared with his performance in other school subjects. Based on Kishor's study, teachers paid more attention to consensus data than they did to consistency and distinctiveness data, indicating that teachers' strategy of information use is not the best. This would suggest that the individual student differences would not likely be considered when a student departed from the norm (Kishor, 1992).

G. Recommendations for Early Identification Initiatives

Early identification measures must be developed that are multivariate (child and family focused) and that account for various risk factors over time (Kochanek, Kabacoff, Lipsitt, 1990; Siegel, 1988). Identifying, defining and assessing these risk factors over time is an enormous challenge that will require collaboration of specialists in various fields. Measures relating to a child's

developmental competence between four and seven years of age have the greatest predictive value (Kochanek, et al. 1990). As well, the screening process must be examined and a wide array of services related to the screening results, to meet the child and family's needs, should be made available (Kochanek, et al. 1990).

> Decisions that have a major impact on children such as enrolment, retention, or assignment to remedial or special classes should be based on multiple scores of information and should never be based on a single test score (NAEYC, 1987).

Teacher observations, samples of work, as well as anecdotal records are also important sources of information. Considering the scarcity of resources, potential for measurement error as well as test bias, screening tests should only be used when information contributes to improved instruction for the child and should also be considered as only one source of information (NAEYC, 1987).

According to Guralnick, the future direction in research should focus on children's social competence as well as on the "need to develop new intervention strategies in collaboration with biobehavioral scientists and child development specialists" (Guralnick, 1991, p.175). The curriculum content, where individualized intervention strategies will be developed, based on the translation and integration of emerging developmental and biobehavioral knowledge is considered to be one of the major future challenges (Guralnick, 1991). There has been a shift away from activity-based toward needs-based approaches to early intervention (Duran, 1989; Dunst, Snyder & Mankinen, 1987). "A focus on children's social competence may also prove to have long-

term benefits for children and families" (Guralnick, 1991, p.181). Finally, authentic assessment that serves to assist teachers and resource specialists in the everyday instruction of pupils would serve a more effective role (Duran, 1989).

H. Summary

Interest in early identification and intervention has been sustained over the past 30 years. Research studies indicate strong support for the importance of early identification and the need for effective intervention. Specific predictors of children at-risk have been verified through the literature. Behaviour of young children is more susceptible to change, thus early intervention has proven to be more effective (Hayden, 1974; Stone, Cundick, & Swanson, 1988).

Issues surrounding early identification point to the misuse of tests with young children. Many researchers have questioned the reliability of tests to accurately identify children who may be at-risk for learning in school. Tests should meet the technical standards set out by the joint committee.

Some research indicates that teachers are able to accurately predict later achievement, whereas other studies indicate that there was bias in teacher judgment and that accuracy of teachers' predictions depends on their ability to process performance information.

Recommendations for future early identification initiatives are that they

must be multivariate and must be designed to meet the needs of the child as well as the family. Focus of identification should be towards effective program planning and intervention strategies. Integration of developmental and biobehavioral knowledge in order to develop information that will meet individual needs is one of the future challenges. Future trends point towards authentic assessment that provides information for the instruction of students.

In view of the preceding literature survey, this study addresses the problem of early identification by examining teacher judgment in identifying kindergarten children who may have a greater-than-average chance of developing a disability. A series of three classification procedures ranging from general to specific are used to find out if teachers' judgments are consistent throughout the three phases. A skill specific Teacher-made Checklist, teachers' own system of rating, as well as a standardized checklist are utilized in the present study. It is hoped that this study may add information towards developing a cost-efficient identification system that would assist in effective instructional planning to meet the needs of young children.

CHAPTER 3

METHODOLOGY

The purpose of this chapter is to describe the methodology used to answer the research questions. The chapter begins with a brief description of the study. The three phases in the study are described, as well a review of the instruments used is given. The research design used to help answer the research question is presented and the sample selected for the study is described. Methods for data analysis are also presented.

A. NATURE OF THE STUDY

This is an ex post facto design with a primary focus on establishing the predictive validity of an experimental Teacher-made Checklist. Ex post facto research is a description of research that is done "after the fact." That is, in ex post facto research one can not manipulate variables because the variables have already occurred (Kerlinger, 1973).

Briefly, the study was conducted in three phases. These phases range from a broad, general categorization of children to a rating of children using a specific skills checklist. Seventeen kindergarten teachers volunteered to participate in the study. In the first phase they were asked to categorize their students as high (low-risk), average, or low (at-risk). Secondly, the teachers were asked to rate each of their students using the Academic Competence Questionnaire from the Social Skills Rating System (Gresham & Elliott, 1990).

In the final phase the teachers were asked to complete the Experimental Teacher Checklist for each of their students. A Likert-type scoring system was used on both of the checklists. Three hundred and thirty-six kindergarten students were rated in this study.

The study focused on teacher accuracy and consistency in identifying students who might be at-risk for school failure.

B. PHASES IN THE STUDY

The researcher met with each teacher individually to explain the requirements and nature of the study. During this time teachers were asked to complete an information sheet that asked for: the number of years of university training, the number of years of teaching experience, the number of years of teaching at the kindergarten level, the number of divisions in the school, and the school population (see Appendix A). Each phase was conducted separately, that is, the teachers received the new envelope containing the material for the subsequent phase when they completed and returned the previous phase.

1. Phase 1 - Early Identification

Using their own early identification methods, teachers were asked to rank the students in their kindergarten class as low (at-risk), average or high (low-risk) in the first phase. This phase was incorporated to partially address the problem, if teachers apply their own standards in making global

assessments of children's performance, to what extent are these ratings consistent with subsequent specific ratings using several global items or several specific ratings of skills and abilities? For those students who were at-risk, the teachers were asked to indicate the area of difficulty, for instance: behaviour, ability, language, or immaturity (Appendix B). In scoring, the students who were high (low risk) were given a score of 1; those who were average were awarded a 2; and the low (at-risk) students were scored as 3. The classification form and teacher information sheet were sealed in an envelope and returned to the researcher. The teachers then completed the second phase of the study.

2. Phase 2 - Academic Competence Questionnaire

The Academic Competence Questionnaire from the Social Skills Rating System (SSRS) was used in Phase 2 (Appendix C). Each teacher was asked to rate her students using the above questionnaire. The following reliability and validity information was presented in the Social Skills Rating Scale Manual. The Social Skills Rating System is a nationally standardized series of questionnaires that obtain information on the social behaviours of children and adolescents from teachers, parents, and the students themselves. Academic competence is also assessed using a teacher questionnaire, as it provides an index of academic functioning. The SSRS is available in three forms, assessing children between the ages of 3 and 18 (preschool through grade 12). Teacher and Parent forms are available for the three developmental levels: preschool,

grades 3 through 6, and grades 7 through 12. Teacher and student self-rating forms are available for the latter two levels. For the purpose of this study the Academic Competence Scale from the Elementary Level (grades K-6; Teacher Form) was used.

Raw scores were converted to standard scores (M = 100, SD = 15) and National percentile ranks. Separate tables were given for boys and girls.

During test development various approaches were used to establish reliability and validity. Coefficient Alpha reliability measuring Internal Consistency on the Academic Competence Questionnaire was .95. The internal consistency estimate for the Elementary Female was .96 and .95 for the Elementary Male.

Test-retest reliability was measured by re-rating the students four weeks after they were initially rated. Coefficient Alpha reliability of .93 was evidence of temporal stability.

Three criterion-related validity studies have been conducted using the SSRS. One study used the Social Behaviour Assessment (SBA) (Stephens, 1978). This assessment used 136 social behaviours that were organized into subdomains of Environmental Behaviours, Interpersonal Behaviours, Self-Related Behaviours, and Task-Related Behaviours. Correlations between the SSRS Academic Competence (N = 71) were: Environmental -.37, Interpersonal -.66, Self-Related -.57, Task-Related -.72, and Total -.67. The Subscales were negatively related to Academic Competence.

Another study used the Child Behaviour Checklist - Teacher Report Form (CBCL-TRF) (Achenbach & Edelbrock 1983). Correlations between Academic Competence and CBCL-TRF were: Externalizing Behaviours -.60, Internalizing Behaviours -.26, and Total Behaviour Problems -.59 (N = 93). The negative correlations were predicted as there is a strong relationship between social behaviours and academic competence (Gresham & Elliott, 1990).

The third validity study was conducted using the Harter Teacher Rating Scale (TRS) (Harter, 1985). Validity coefficient for Academic Competence was .63 (N = 243).

To establish internal consistency, intercorrelations between subscales and Total Scale Raw Scores were used. Coefficient alpha median reliability was .95 (N = 937).

Factor Analyses were conducted within domains. All factors were rotated obliquely using the direct oblim solution. Factor loadings for the Teacher Form, Elementary Level were: Overall Academic performance .93; Intellectual functioning .91; Reading .90; Skills in mathematics .90; Skills in Reading .90; Mathematics .90; Overall motivation .84; Parent encouragement .73; and Overall classroom behaviour .62.

To further establish construct validity known groups were compared. Academic Competence Mean for non-handicapped (N = 769) was 34.0, SD 8.9; Academic Competence Mean for Learning Disabled (N = 116) was 26.7, SD 8.6; and Academic Competence Mean for Other Handicapped (N = 68) was 25.6, SD

9.8. Scheffe post hoc tests revealed significant differences between the groups ($p < .05$) (Gresham & Elliott, 1990).

Scoring of Academic Competence Checklist

A Likert-type scoring system was used to score the Academic Competence Checklist. Teachers were asked to rate students in their class according to the following instructions: (1) Lowest 10%, (2) Next lowest 20%, (3) Middle 40%, (4) Next Highest 20%, (5) Highest 10% (Gresham & Elliott, 1990). Raw scores were converted to standard scores.

3. Phase 3 - Teacher Checklist

a. Development of Teacher Checklist:

The Teacher Checklist was developed by the researcher using input from experienced teachers. Forty first-grade teachers from the Abbotsford School District generated items for a checklist of skills that they felt a child should possess at the beginning of grade one in order to meet with success in grade one (second year primary). Thirty-two items were rated on this list.

Two years later, ten first-grade teachers were given this list of 32 items and they were asked to select the 10 items that they felt were the most important skills for a second year primary (grade one) student to possess in order to meet with success in school. These teachers were also encouraged to include any other skills that they deemed essential. The ten teachers'

education ranged from B.Ed to M.Ed (4-6 years post secondary training) and teaching experience ranged from 9 to 22 years (Mean = 16.44 years).

The checklists submitted by the 10 teachers were tabulated. Skills that were selected as being essential by at least three teachers were included. New items suggested by any one of these ten teachers were also included. Items that were similar were combined. These items on the Checklist were compared with the Learning Descriptors in the Draft Document of the Primary Program, Ministry of Education, B.C. Items that had not been generated by the teachers but were listed as Learning Descriptors in the Primary Program were added to the Checklist. Refer to Appendix D for a further description of item selection and grouping of items. The items were categorized under the appropriate subheadings: Emotional Development, Social Development, Physical Development, Social Responsibility and Intellectual Development. These headings were taken from the Primary Program Draft Document (Primary Program, 1989). The approach of using the 10 grade one teachers, in order to prioritize the items on the Checklist, contributed to content validation of the Checklist. The procedures were used in an attempt to establish content validation of the Checklist.

The computer program SPSS-X RELIABILITY was used to perform an analysis of internal consistency in order to establish reliability of the Teacher Checklist which involved 17 teachers rating 336 students. Cronbach's Alpha reliability coefficient for the 28 item Checklist was .98. The reliability

coefficients for the Subscales were as follows: EM DEV (7 items) was .93, SOC DEV (3 items) was .96, PHYS DEV (4 items) was .86, and INT DEV (13 items) was .97.

In order to establish criterion-related validity, the SPSS-X CORRELATIONS computer program was used to compute Pearson Correlation Coefficients between the Checklist totals and the Checklist Subscales as well as with the Academic Competence Standard Scores. Correlation between ACSS and CH Total was .86. Correlations between AC standard scores and Subscales were: EM DEV .80, SOC DEV .57, PHYS DEV .67, and INT DEV .88. Correlations between the Checklist total and Subscales were: EM DEV .93, SOC DEV .78, PHYS DEV .87, and INT DEV .96. More detailed results are given in Chapter four.

b. Scoring the Experimental Checklist:

A Likert-type scoring system was used. A score of one indicated low skill development and five indicated high skill development. The scoring of the Checklist employed total raw scores for each of the subscales as well as a raw score total for the entire checklist.

C. DESIGN OF THE STUDY

1. Dependent Variables

The dependent variable in this study was group membership; (1) students at-risk, (2) students not at-risk. Group membership was

established by performing a Cluster Analysis of the Teacher Checklist using the five subgroupings (Emotional Development, Social Development, Social Responsibility, Physical Development, and Intellectual Development). A more detailed description will be provided later.

2. Independent Variables

The independent or predictor variables in this study were: (1) Teacher Classification, Early Identification (2) the Academic Competence Subtest from the Social Skills Rating System, and (3) An Experimental Teacher Checklist. Scoring of Teacher classification (EI Score) was: high (low risk) 1, average 2, and low (at-risk) 3.

The Academic Competence (AC) Raw Score was converted to Standard Scores. Cluster Analysis was used to group the like subjects.

The five subscale raw sores, Emotional Development (EM DEV), Social Development, (SOC DEV), Physical Development, (PHYS DEV), Social Responsibility (SOC RESP), and Intellectual Development (INT DEV), on the Teacher Checklist were used and like groups were established using a Cluster Analysis technique.

3. Research Design

This present study used a Cluster Analysis technique in order to make prediction based on high risk and low risk. Cluster Analysis is a method used

to discover homogeneous groups in data sets (Anderberg, 1973). Agglomerative methods were used to combine or cluster groups of individuals that were rated as most similar on the Academic Competence Questionnaire as well as on the Experimental Teacher Checklist. Agglomerative method is a hierarchical procedure which begins with individuals in separate clusters and combines the individuals that are most alike in subsequent steps to build new aggregate clusters (Hair, Anderson & Tatham, 1987). In each case, clustering with Euclidean distance using complete linkage was relied on.

A one-way analysis of variance was computed in order to find out if there were significant differences between the four clusters. A post hoc analysis using the Tukey procedure was used to asses the extent of the difference between groups.

Four basic clusters were derived. A 4 X 3 matrix was constructed to present the comparison of the four cluster solution on each phase (2 and 3) with teacher ratings on Phase 1. A Chi-Square Test was used in each of the two phases to test whether the two rating methods were related.

Several different approaches were used to provide validity information for the Teacher Checklist. In order to assess whether teacher differences were real differences or whether there were other contributing factors to the differences, a Pearson Correlation Coefficient was computed. The relationship between teachers' ratings and school population was assessed by correlating teacher mean scores on the Teacher Checklist and the SES score of the

corresponding school. Although it can be speculated that teachers' ratings may vary because of differences among teachers, differences in school populations are undoubtedly a contributing factor to varying scores.

In order to further establish validity of the experimental Checklist (Phase 3), Pearson Correlation Coefficients were computed. The Checklist was correlated with the Academic Competence subtest from the Social Skills Rating System. The Social Skills Rating System is a standardized measure with established validity and reliability.

A third analysis designed to provide further evidence of validity of the Teacher Checklist was conducted. Crosstabulations of the cluster membership in the AC Questionnaire with the cluster membership in the Teacher Checklist were presented. A Chi-square Test was used to test if there was a significant difference between the two sets of clusters for the two measures.

D. PROCEDURE

1. Subjects

The subjects in the study were 17 volunteer practising kindergarten teachers. A total of 336 kindergarten students were rated by these 17 teachers. The majority of the teachers rated all of the students in one of their kindergarten classes. One teacher in the study rated the students from both her morning and afternoon kindergarten classes. The number of students rated by each teacher varied, from 11 to 30. The distribution of subjects is displayed

on Table 1. The average number of students rated by each teacher was 19.76.

TABLE 1

DISTRIBUTION OF STUDENTS

Teacher ID	Number of Students	Percentage
1	24	7
2	19	6
3	20	6
4	11	3
5	19	6
6	15	4
7	17	5
8	18	5
9	30	9
10	21	6
11	24	7
12	20	6
13	21	6
14	22	7
15	22	7
16	20	6
17	13	4

2. Demographic Characteristics of Sample

Sixteen kindergarten teachers who participated in this study worked in the Abbotsford Public School District. One teacher worked in a separate parochial school in Abbotsford. Abbotsford is a small city located 70 kilometres

east of Vancouver. The city has a diverse professional demography; many residents commute to jobs in Vancouver while others are dairy or soft fruit farmers.

The teachers came from 13 elementary schools in the Abbotsford School District. Demographic data pertaining to the teachers' experience and training is presented in Table 2. The teachers' post secondary training ranged from three to six years (Mean = 4.52, SD = .7), and total teaching experience ranged from 4 to 25 years (Mean = 12.16, SD 6.02). The teachers' experience in teaching kindergarten ranged from .6 to 22 years (Mean = 7.39 years, SD = 5.96). All the teachers in the study were female. Eight teachers in the sample were distributed across four schools.

TABLE 2

CHARACTERISTICS OF SAMPLE - TEACHERS

	Training (years)	Experience (years)	Experience at this grade (years)
Mean	4.52	12.16	7.39
S.D.	0.70	6.02	5.96
Range	3 - 6	4 - 25	.6 - 22

Characteristics of the student sample are presented in Table 3. There were 336 kindergarten students who were rated in the three phases of the study. Their ages ranged from 5.43 years to 6.98 years (Mean = 5.94, SD = .436). The sample of students was comprised of 180 Males and 156 Females. The data was collected in May and June, thus the students had been attending

school for almost a complete school year.

TABLE 3
CHARACTERISTICS OF SAMPLE - STUDENTS

Sex	Age	
Male = 180 Female = 156	Mean	5.94
	S.D.	.436
	Range	5.43 - 6.98

Characteristics of the schools are presented in Table 4. The size of the schools ranged from 7 to 21 divisions (Mean = 14.18, S.D. = 4.59). School populations ranged from 181 to 556 with a Mean of 353.13 (S.D. 113.63).

TABLE 4
CHARACTERISTICS OF SCHOOLS

	Number of Divisions	School Population
Mean	14.38	352.13
S.D.	4.59	113.63
Range	7 - 21	181 - 556

Schools used in the sample included: city centre schools, schools from the suburbs, as well as schools from the surrounding farming community.

Schools represented in the sample ranged from 28% below the district average in Socio-economic Status (SES) to 21% above the district average. The school district calculated the SES Scores of the individual schools based on the mother's highest level of formal education.

In order to collect this SES information, surveys were sent to families throughout the school district. This survey is conducted annually by the school

district. There were 35 items on this particular survey, however the items are revised to some extent annually. The items that were included involve questions relating to parent's attitudes about their child's school. A section of the survey asked for General Information. One of the questions under the General Information section on the survey asked for the highest level of formal education completed by the mother and father. Every fourth family from each school roster was sent a survey. The rate of survey return was 49%. The SES scores for the schools were computed based on the information pertaining to the mother's highest level of education completed.

Some of the schools had a high percentage of immigrant students for whom English is their second language. This high ethnic representation in a few of the schools could have been a contributing factor to some of the SES scores. Moreover, this diverse ethnicity of the school populations could have had an effect on the teachers' ratings.

For participation in the research project, subjects were initially contacted by telephone. The researcher met individually with each teacher and explained the nature of the study as well as the requirements from each teacher. Each teacher was given a written description of the study and instructions for proceeding (See Appendix E). The school administrator was also informed about the study (See Appendix F). All 17 teachers indicated an interest in participating and there was 100% completion of all three phases by the 17 teachers.

E. DATA PREPARATION AND ANALYSIS

1. Data Preparation

Scoring procedure for the data followed the description provided earlier. All data were numerically coded and fed into a computer data file. All data entered were 100% verified and all corrections were made prior to analysis.

2. Data Analysis

Subject responses on the three different phases were analyzed using the SPSS-X (Statistical Package for Social Sciences Extended Version Release 2.0) computer program.

a. Reliability Analysis

A series of reliability analysis were conducted using the SPSS-X RELIABILITY program in order to establish internal consistency of the Teacher Checklist. In order to provide validity evidence for the Teacher Checklist, a correlation matrix was generated using AC Standard scores, Teacher Checklist Total Scores, and Subscale scores.

First, a reliability analysis was performed using the individual items on the Checklist. Coefficient Alpha reliability indices were provided.

Second, a correlational analysis was performed using the individual subscales in the Teacher Checklist: Em Dev, Soc Dev, Phys Dev, and Int Dev, as well as the AC standard score and the Teacher Checklist total score. The one-item domain, Soc Resp, was not included in this analysis.

b. Hierarchical Cluster Analysis - Academic Competence Questionnaire

In order to discover to what extent the Academic Competence Questionnaire identifies children as at-risk for learning problems, a procedure involving Hierarchical Clustering techniques were used. The computer program SPSS-X CLUSTER was employed to compute this analysis.

c. Analysis of Variance - AC Cluster Membership

A one-way analysis of variance was computed in order to investigate whether the clusters were significantly different. A post hoc analysis using Tukey procedure was used to assess the significance of the difference between groups and to investigate whether there were significant differences between all groups.

d. Crosstabulation of EI Scores with AC Cluster Membership

A 4 x 3 matrix was used to compare the students' placement in the four clusters with the teachers' initial (EI) classifications in Phase 1. A Chi-square Test was used to test whether the two rating methods were related.

e. Hierarchical Cluster Analysis - Teacher Checklist Subscales

In order to discern to what extent the Teacher Checklist Subscales identify children who are at-risk for learning problems, a Hierarchical Clustering procedure was used to group children with similar characteristics. The computer program SPSS-X CLUSTER was again employed to cluster the students who were similar in each of the five subscales. The five subscales that

made up the checklist were: Emotional Development (EM DEV), Social Development (SOC DEV), Physical Development (PHYS DEV), Social Responsibility (SOC RESP), and Intellectual Development (INT DEV).

f. Analysis of Variance - Teacher Checklist Subscales

A one-way analysis of variance was computed in order to investigate whether the clusters were significantly different in each of the subscales as well as the total Teacher Checklist. A post hoc analysis using Tukey procedure was used for each analysis to assess the significance of the difference between groups and to investigate whether there were significant differences between all groups.

g. Crosstabulation of EI Scores with Teacher Checklist Subscale Membership

A 4 x 3 matrix was used to compare the students' placement in the four clusters based on the Teacher Checklist Subscales with the teachers' initial (EI) classifications in Phase 1. A Chi-square Test was used to test whether the two rating methods were related.

h. Teacher Differences

To establish evidence for the validity of the Teacher Checklist a series of analyses were undertaken. Firstly, to assess whether teacher differences were real differences or whether there were other contributing factors to the differences, a correlational analysis between teacher mean ratings and corresponding school SES scores was computed. Secondly, a Pearson

Correlation between AC Questionnaire, Teacher Checklist Totals and Teacher Checklist Subscales was computed to provide further evidence of validity. Crosstabulation of cluster membership on the AC Questionnaire with cluster membership on the Teacher Checklist was provided as yet more evidence of validity.

F. SUMMARY

The purpose of this chapter was to describe the methods that were used to generate an answer to the research questions. Initially a general description of the nature of the study was given. Descriptions of the three phases that made up the study were then given in greater detail. A brief account of instrument development was included. Psychometric statistics of the standardized instrument were provided. As well, evidence of validity for the experimental Checklist was given. The research design was described, and following the research design, a description of the sample used in the study was provided. In the final section, brief procedures for data preparation and data analysis were outlined.

CHAPTER 4

RESULTS

This chapter presents the results of the psychometric analysis in the same order as previously described in Chapter 3. Results of the initial classification of students in Phase 1 is given and further results of Phases 2 and 3 are provided. Hierarchical Cluster Analysis of the AC Questionnaire and the Checklist subscales, which is the heart of this psychometric analysis, are presented in detail. Other analyses used to assist in establishing reliability and validity of the Teacher Checklist are also fully described. The complete results of all the analyses used to answer the research question are presented in this chapter.

ANALYSIS OF DATA

A. PHASE 1 - Early Identification

1. Summary of Classification

The summary of the classification of the students in Phase 1 - Early Identification is presented in Table 5. The teachers in the study rated 336 students. There were 97 students rated a 1 which is high or above average; 183 students were rated a 2 or average; and 56 students were rated a 3 or considered to be 'at-risk'.

TABLE 5
PHASE 1 - EARLY IDENTIFICATION CLASSIFICATION OF STUDENTS

	Group	Number
High	1	97
Average	2	183
At-risk(low)	3	56

2. Reliability

a. Internal Consistency

In order to establish internal consistency of the Teacher Checklist, a series of reliability analyses were conducted. Cronbach's Alpha reliability coefficient for the Total Checklist (28 items) was .98. The reliability coefficient for: the subscale EMDEV (7 items) was .93, the subscale SOCDEV (3 items) was .96, the subscale PHYSDEV (4 items) was .86, and INTDEV (13 items) was .97. The INTDEV coefficient was the highest and closest to the Total Checklist Score. According to Crocker and Algina (1986), adding items to a test should increase total variance. The proportion of increase in test variance is more marked when items are added to a short test than to a longer one (Crocker & Algina, 1986). Considering the subscales, INTDEV has 13 items, whereas EMDEV has 7 items, SOCDEV has 3 items and PHYSDEV has 4 items. Proportionately, INTDEV has the most items and therefore the highest variance and the highest coefficient.

The reliability analysis of the AC Questionnaire (9 items) yields a

Coefficient Alpha of .96. This is comparable with the Social Skills Rating System Manual where the Coefficient Alpha for Academic Competence is .95 (Gresham & Elliott, 1990).

b. Correlational Analysis

In order to provide concurrent validity evidence for the Checklist, a correlation matrix was generated using: AC Standard Scores, CH Total Scores, CH subscales - EM DEV, SOC DEV, PHYS DEV and INT DEV. Table 6 displays the correlation matrix for ACSS, CH TOTAL, EM DEV, SOC DEV, PHYS DEV, and INT DEV.

TABLE 6
PEARSON CORRELATION COEFFICIENTS

	1.	2.	3.	4.	5.	6.
1. AC STANDARD SCORES		.86	.80	.57	.67	.88
2. CH TOTAL			.93	.78	.87	.96
3. EM DEV				.79	.74	.83
4. SOC DEV					.65	.63
5. PHYS DEV						.80
6. INT DEV						

* All correlations were P= 0.001.

Correlation between ACSS and CH TOTAL was .86. The Checklist subscale correlations range from .57 to .88 with ACSS (Mean validity coefficient = .73). Slightly higher correlations were observed between INT DEV, the other subscales, as well as the CH TOTAL, and ACSS. ACSS correlates .88 with INT DEV. A reasonable explanation for the higher correlation of ACSS with INT

DEV than with the other subscales is that the AC Questionnaire tends to focus mainly on academics and consequently is more closely related to the INT DEV Subscale.

B. PHASE 2 - Academic Competence Questionnaire

1. Hierarchical Cluster Analysis - Academic Competence Questionnaire

In order to discover to what extent the Academic Competence Questionnaire identifies children as at-risk for learning problems, Hierarchical Clustering techniques were used. Nine items made up the AC Questionnaire. Both correlational and distance measures were used to test the stability of the analysis. Clustering with Euclidean distance, which utilized a complete linkage procedure was relied on. For the purpose of this study a four cluster solution appeared to be the most reasonable. The agglomeration coefficients demonstrated that four clusters explained the data the most adequately. The cluster analysis as shown on the horizontal icicle plots also displayed four distinct clusters. Means and standard deviations were provided for each of the nine items in the Questionnaire. As well, means and standard deviations were provided for each of the four clusters on each of the nine items in the Questionnaire. Using the SPSS-X CLUSTER computer program, an agglomeration schedule using complete linkage provided correlation coefficients. A plot of the means and standard deviations for each of the subscales for each cluster aided in the interpretation of the clusters.

The following results help to answer the question: "to what extent does the Academic Competence Questionnaire identify children as being at-risk for learning problems?" The cell means and standard deviations for each of the Early Identification classifications based on the Academic Competence Questionnaire are presented in Table 7. Means for the three groups are distinctly different. The mean difference between Groups 1 and 2 is 14.92 and 15.6 between Groups 2 and 3.

TABLE 7
CELL MEANS FOR EARLY IDENTIFICATION GROUPS FOR ACADEMIC COMPETENCE QUESTIONNAIRE

EI SCORE	MEAN	SD	NUMBER
1	106.95	7.54	97
2	92.03	6.50	183
3	76.43	7.10	56

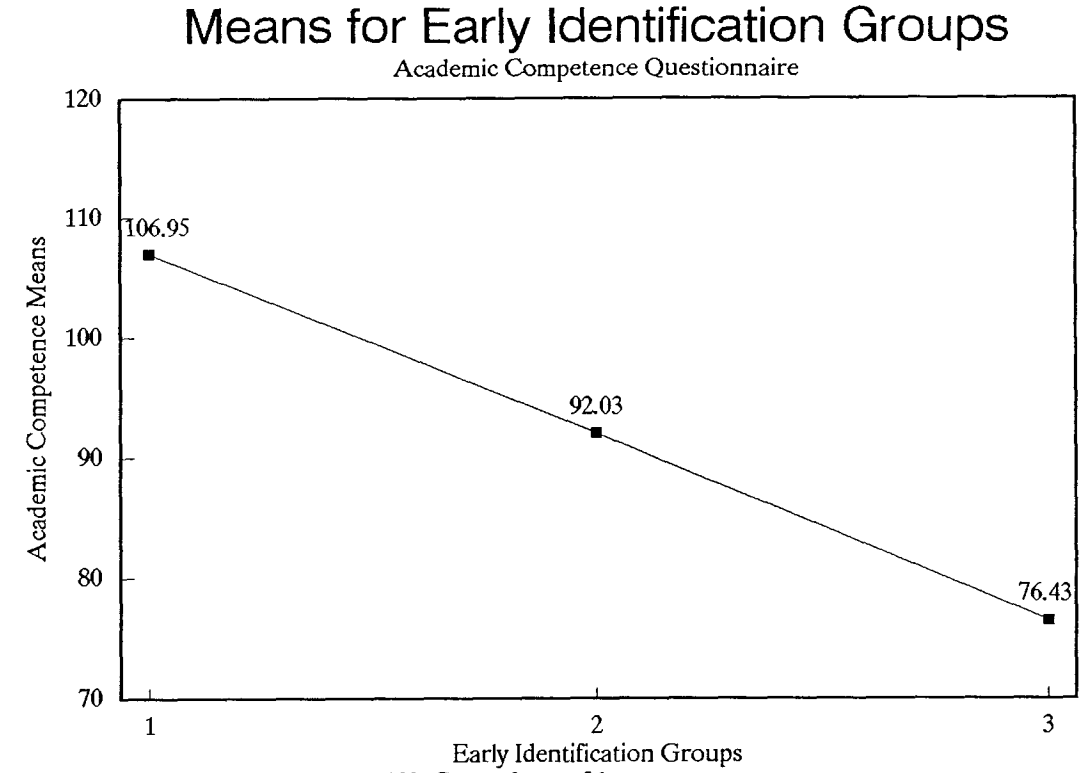
Table 8 presents the means for cluster membership on the Academic Competence Questionnaire.

TABLE 8
MEANS FOR CLUSTER MEMBERSHIP ON ACADEMIC COMPETENCE QUESTIONNAIRE

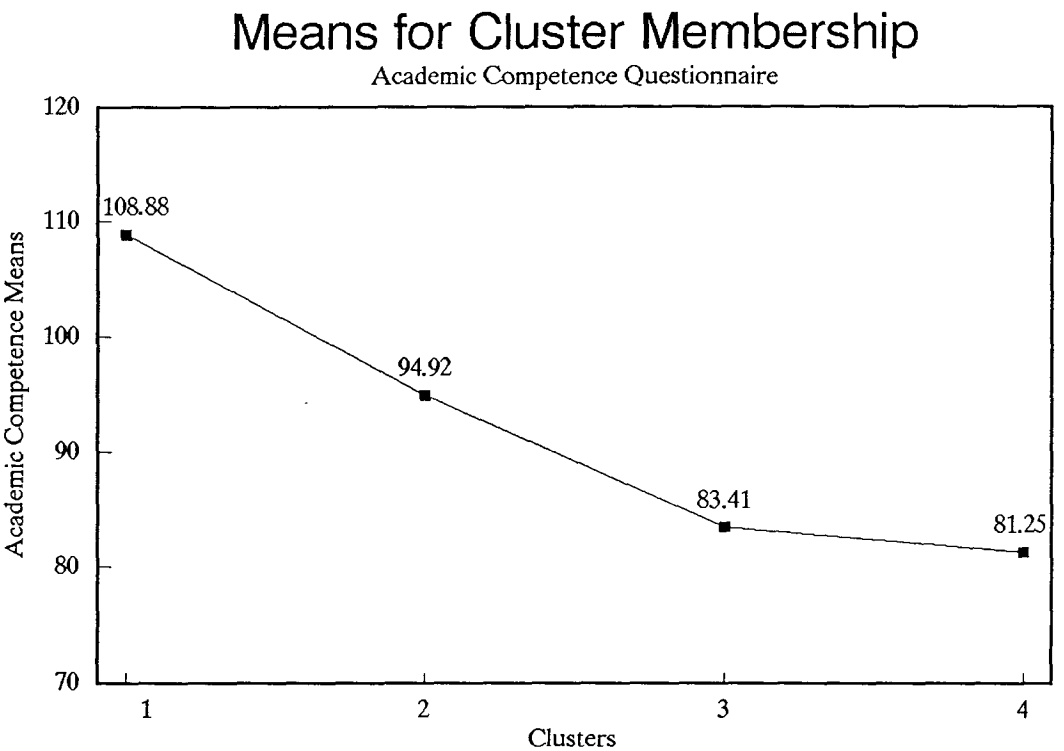
Cluster Membership	Mean	S.D.	Number in Cluster
1	108.88	5.46	90
2	94.92	4.28	118
3	83.41	4.16	44
4	81.25	8.63	84

Figure one shows graphic representation of EI Group Mean AC Scores and the Cluster Group Mean AC Scores.

FIGURE 1



Group 1: n = 97, Group 2: n = 183, Group 3: n = 56



Cluster 1: n = 90, Cluster 2: n = 118
Cluster 3: n =44, Cluster 4: n = 84

In order to facilitate interpretation of the four clusters, the cluster profile points for each group of students were plotted relative to the AC mean and standard deviation for each of the nine items for the total sample of children. Figure two illustrates the cluster profile points for the different groups of students.

Cluster 1, with 27% of the sample, was consistently one z-score higher than Cluster 2 in the first eight items and the same as Cluster 2 in overall classroom behaviour.

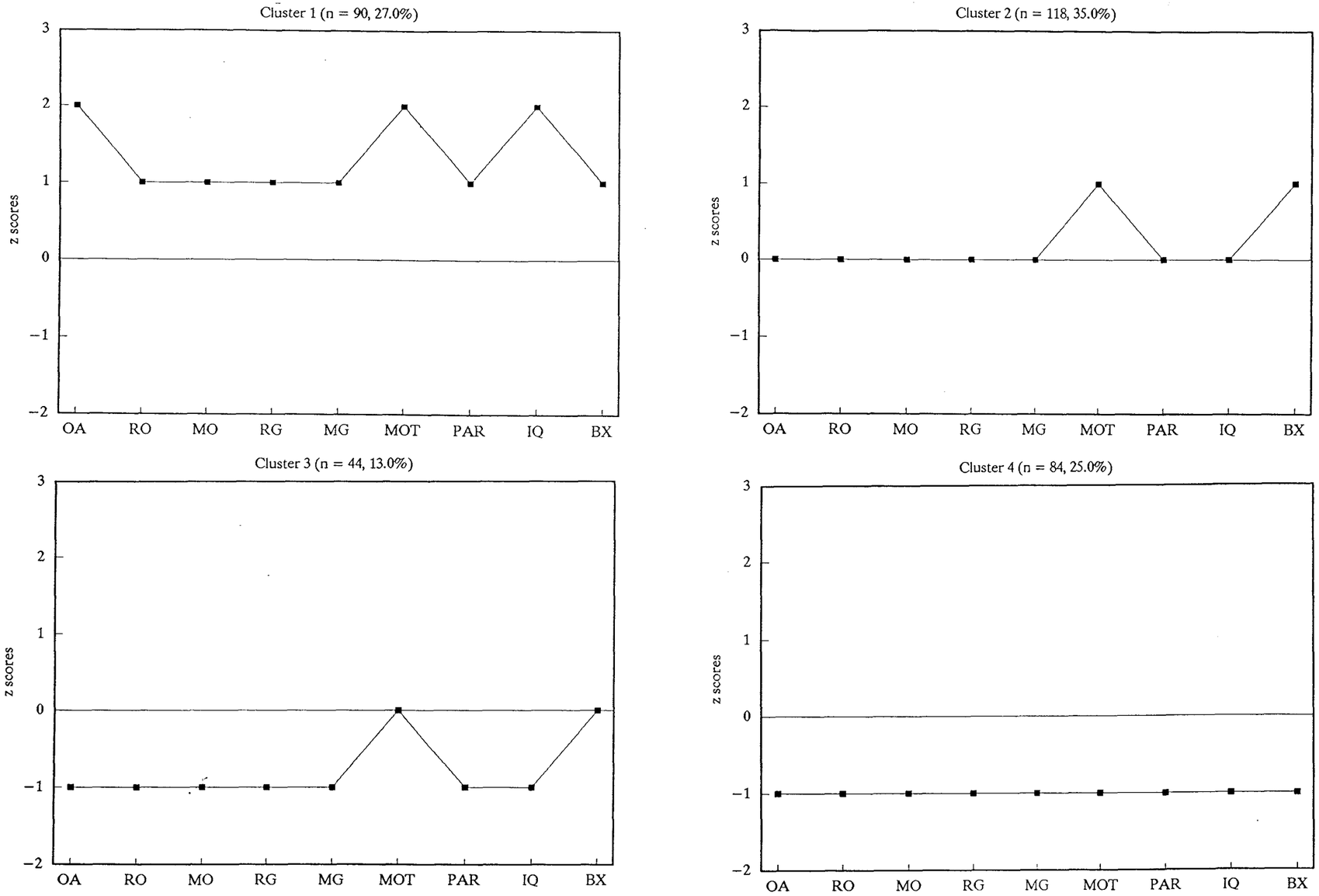
Cluster 2, containing 35% of the sample was approximately one z-score higher than Cluster 3 in all areas. With 62% of the sample in Clusters 1 and 2, it is unlikely that members from these two clusters would be at-risk because their ratings are all average or above average.

Cluster 3 is smaller, with only 13% of the sample, it is similar to Cluster 4 in seven items but is one z-score higher in motivation and overall classroom behaviour. Possibly some students from this cluster may be later identified as at-risk.

One fourth of the sample is in Cluster 4. The members of this cluster are low in all areas on the AC Questionnaire. Therefore, individuals in this cluster would likely be at-risk.

FIGURE 2

Cluster Membership Academic Competence



OA = Overall Academic Performance, RO = Reading compared to others, MO = Math compared to others, RG = Reading compared to grade-level expectations, MG = Math compared to grade-level expectations, MOT = Motivation, PAR = Parental encouragement, IQ = Intellectual functioning, BX = Overall classroom behavior.

2. Analysis of Variance - AC Cluster Membership

In order to investigate whether the clusters were significantly different, a one-way analysis of variance was computed. The independent variable was AC Total Standard Score and the dependent variable was AC Cluster membership. The results of the analysis of variance of the Academic Competence Standard Scores ratings are presented in Table 9. As displayed on Table 9, there is a significant difference between the individual groups. The significant F ratio between groups factor [$F(3,332) = 364.43$, $p = .001$] indicated that there were significant differences between the four clusters on the Phase 2 Academic Competence Questionnaire.

TABLE 9
ONE-WAY ANOVA TABLE AC STANDARD SCORE BY AC CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	38589.07	12863.02	364.43	.001
Within Groups	332	11718.36	35.30		
Total	335	50307.43			

A post hoc analysis using Tukey-HSD procedure found that some of the clusters were significantly different. Critical table value was 4.20 ($p = .05$). As indicated on Table 10, Clusters 3 and 4 were not significantly different. Every teacher rated students who were grouped in each of the four clusters.

TABLE 10
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
108.88	1		*	*	*
94.92	2			*	*
83.41	3				
81.25	4				

(*) denotes pairs of groups significantly different at the 0.05 level

3. Crosstabulation of Early Identification Scores with AC Cluster Membership

The following section displaying crosstabulation between AC clusters and EI groups will assist in clarifying who may possibly be at-risk. The crosstabulation of Academic Competence membership by Early Identification scores is displayed in Table 11. Based on EI scores, 97 students were classified as high (1), however only 76 of those students were members of Cluster 1, while 18 of those high students were members of Cluster 2 and three were in Cluster 4. These three students initially classified as high who were grouped in Cluster 4 on the AC Questionnaire were likely students who were initially missed by the teacher. Using the EI classifications, 183 students were classified as average (2), however 99 were in AC Cluster 2, 14 were in Cluster 1, 21 were in Cluster 3, and 49 were in Cluster 4. There were 56 students classified as low or at-risk on the EI classifications, yet 23 were in Cluster 3 on the AC groupings, 32 were in Cluster 4 and one student was in Cluster 2. It can be speculated that the 23 students from AC 3 - EI 3, and the 32 students from AC 4 - EI 3 are clearly at-risk. Approximately 16% of the sample would fall into

this group. Further consideration should be given to the 49 students from AC 4 - EI 2. These students, another 14% of the sample, may also be at-risk. Approximately 31% of the sample falls into the category of being clearly or likely at-risk and would likely benefit from prereferral intervention or further assessment. It is interesting to note that 70 students who were rated as average on the EI classifications were in Clusters 3 and 4 on the AC. Upon perusal of the AC Standard Scores for EI Group 2, 67 students who were classified as average (2) in Phase 1 were rated below the 25th percentile on the AC ratings. This may be an indication that when teachers use a more specific measure, they are able to make finer discriminations between students abilities.

TABLE 11
CROSSTABULATION OF ACADEMIC COMPETENCE MEMBERSHIP BY EARLY IDENTIFICATION SCORES

		EI SCORES			
AC MEM		1	2	3	Row Total
	1	76	14		90
	2	18	99	1	118
	3		21	23	44
	4	3	49	32	84
	Column Total	97	183	56	336

A Chi-square Test was used to test if there was a significant difference between the Academic Competence Group Membership when children with like characteristics were clustered using the Hierarchical Cluster Analysis technique

and the initial EI teacher classifications. Chi-square results (268.90, D.F.(6), p = 0.001) indicated a significant relationship between EI groups and AC cluster membership. Cramer's V statistic at 0.63 indicates a moderately strong relationship between Phase 1 EI classification and Phase 2 AC Questionnaire cluster membership.

C. PHASE 3 - Teacher Checklist

1. Hierarchical Cluster Analysis - Teacher Checklist Subscales

In order to discern to what extent the Teacher-made Checklist Subscales identifies children who are at-risk for learning problems, Hierarchical Clustering techniques were used. Five subgroups that made up the Checklist were: Emotional Development (EM DEV), Social Development (SOC DEV), Physical Development (PHYS DEV), Social Responsibility (SOC RESP), and Intellectual Development (INT DEV). Identical Cluster Analysis procedures that were described in the Academic Competence Questionnaire (section b) were also employed in the Teacher Checklist Subscales analysis. A four cluster solution also appeared to be the most reasonable. Means and standard deviations were provided for each of the five subscales in the Checklist. As well, means and standard deviations were provided for each of the four clusters on each of the five subscales in the Checklist. Using the SPSS-X CLUSTER computer program, an agglomeration schedule using complete linkage provided correlation coefficients. A plot of the means and standard deviations of the

subscales for each cluster aided in the interpretation of the clusters.

The cell means and standard deviations for each of the Early Identification classifications based on the Teacher Checklist Subscales are presented in Table 12. Means for the three groups are different. EM DEV, SOC DEV, PHYS DEV and SOC RESP show some differences among the three groups. Using a band of more than +/- one standard deviation, there is distinct separation between groups in the INT DEV Subscale. As indicated in Table 12, the mean difference between Groups 1 and 2 for the INT DEV Subscale was 14.98 and 17.42 between Groups 2 and 3. There is more than one standard deviation separating the three groups on the Total Checklist means.

TABLE 12
SUBSCALE CHECKLIST MEANS AND STANDARD DEVIATIONS
FOR EARLY IDENTIFICATION GROUPS

Subscales	EI Group 1 n=97	EI Group 2 n=183	EI Group 3 n=56
Em Dev Mean	30.30	23.78	17.09
SD	4.01	4.75	4.18
Soc Dev Mean	11.99	10.44	7.16
SD	2.53	2.60	2.61
Phys Dev Mean	16.58	13.78	9.70
SD	2.30	2.65	2.59
Soc Resp Mean	4.14	3.39	2.61
SD	.87	.81	1.01
Int Dev Mean	58.81	43.83	26.41
SD	5.52	7.82	7.25
Ch Total Mean	121.82	94.77	62.96
SD	11.89	15.72	14.07

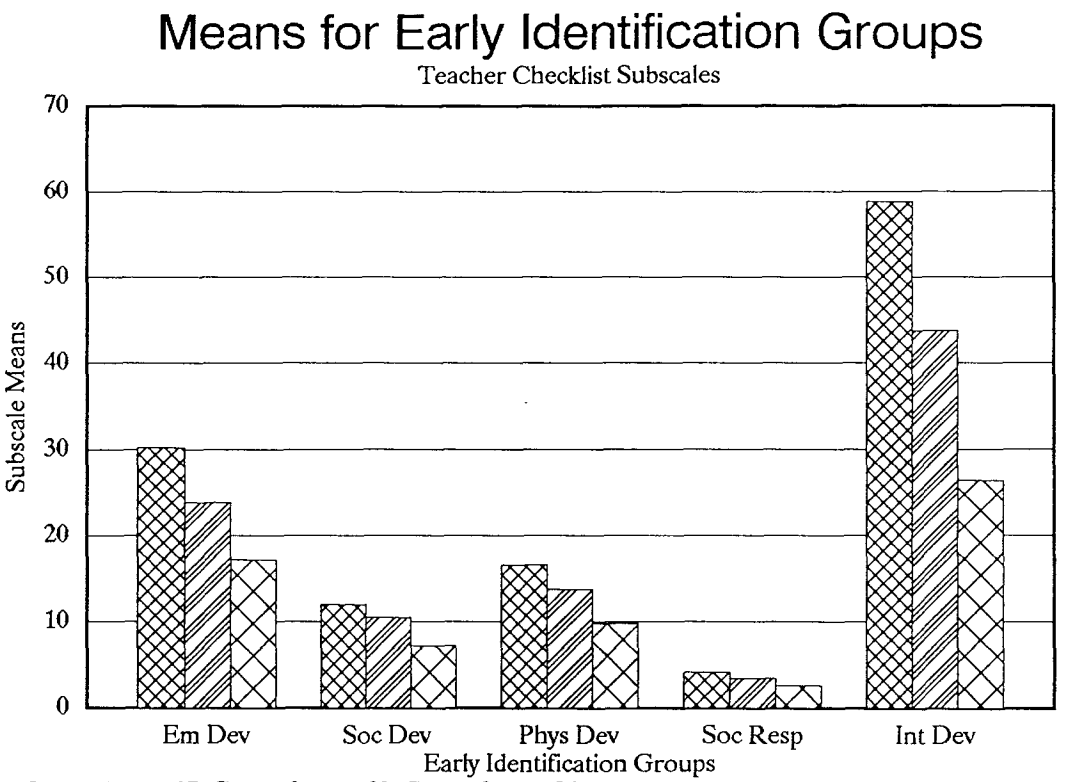
Table 13 presents the means of the clustering variables for the Checklist subscales (EM DEV, SOC DEV, PHYS DEV, SOC RESP, and INT DEV) as well as the Checklist Total means. Cluster 1 is clearly the high group and Cluster 4 is clearly the low group.

TABLE 13
SUBSCALE CHECKLIST MEANS AND STANDARD DEVIATIONS FOR CLUSTERS

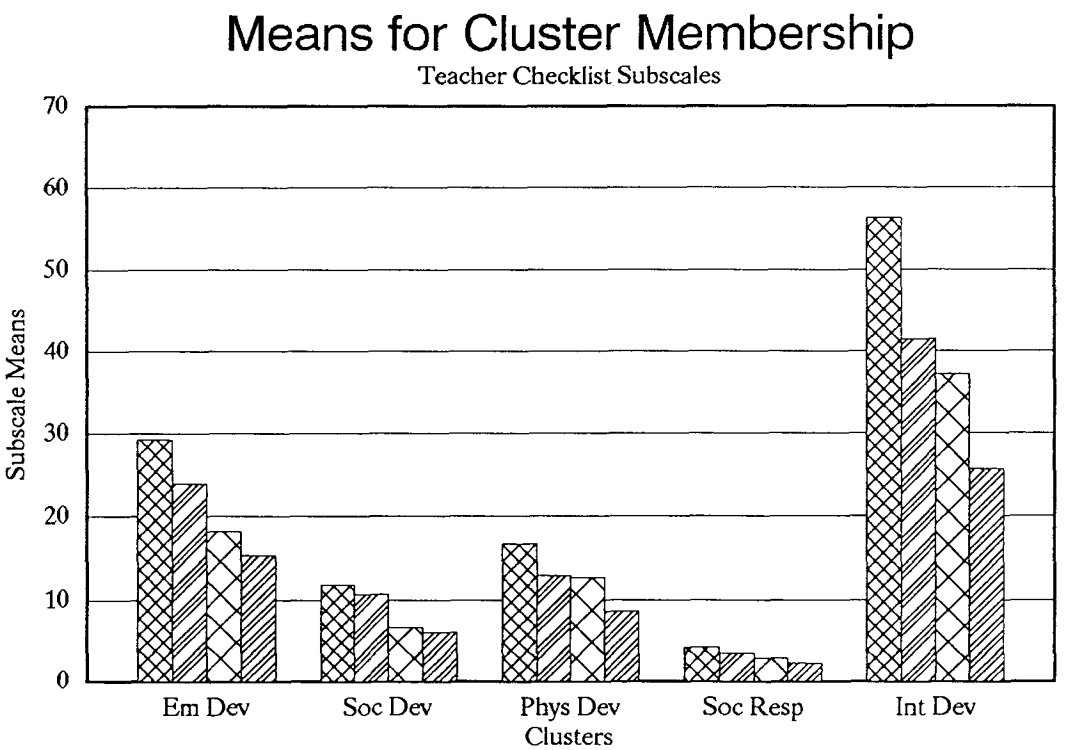
Subscales	Cluster 1 n=144	Cluster 2 n=114	Cluster 3 n=31	Cluster 4 n=47
Em Dev Mean	29.29	24.06	18.19	15.36
SD	4.31	3.19	3.60	3.05
Soc Dev Mean	11.79	10.65	6.58	6.06
SD	2.35	1.92	1.31	2.13
Phys Dev Mean	16.69	12.87	12.65	8.57
SD	1.94	2.02	1.60	1.90
Soc Resp Mean	4.10	3.41	2.83	2.15
SD	.81	.66	.64	.78
Int Dev Mean	56.31	41.54	37.16	25.70
SD	6.72	6.23	9.47	7.38
Ch Total Mean	118.18	92.54	77.42	57.85
SD	12.68	10.20	12.33	10.92

Clusters 2 and 3 are fairly similar in Physical Development and Social Responsibility. Clusters 3 and 4 are fairly similar in Social Development. The four clusters are distinctly separate in Emotional Development as well as Intellectual Development. Total mean scores are quite different for each of the clusters. A graphic presentation of the Checklist Subscale mean scores for the Early Identification Groups and the Checklist Subscale mean scores for the four

FIGURE 3

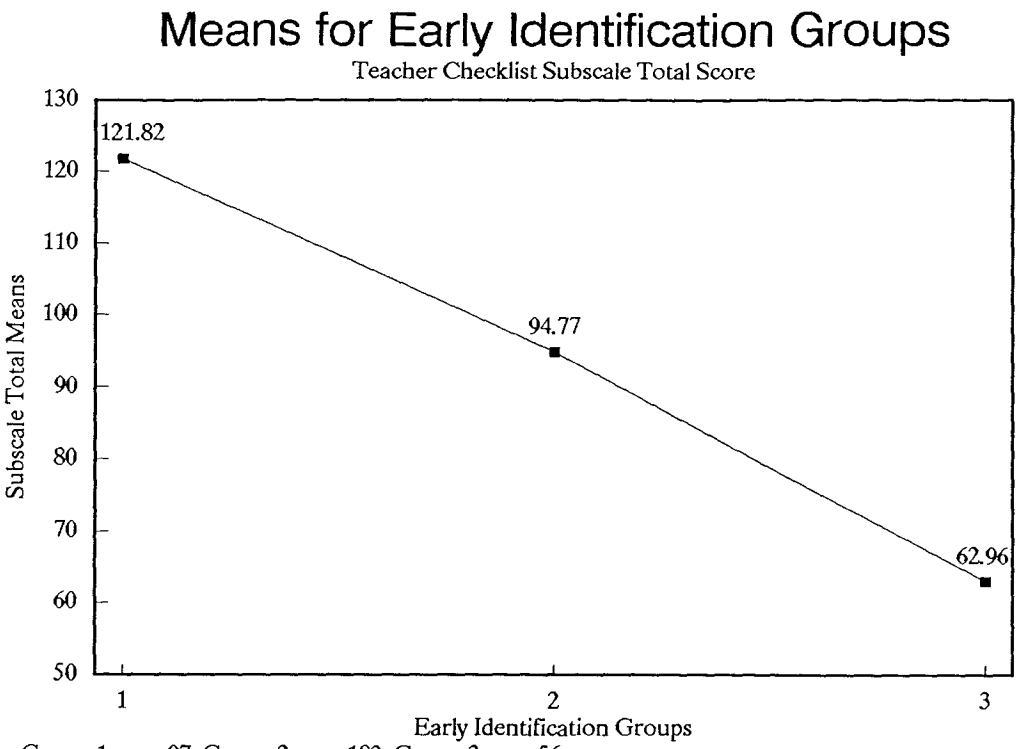


Group 1: n = 97, Group 2: n =183, Group 3: n = 56

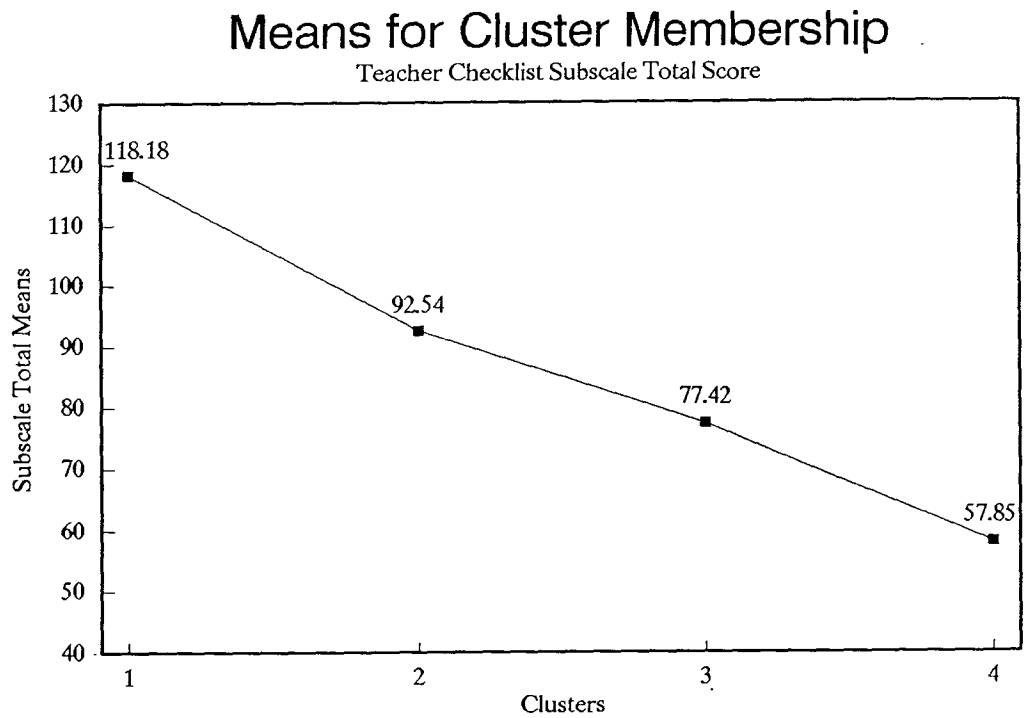


Cluster 1: n = 144, Cluster 2: n = 114
Cluster 3: n = 31, Cluster 4: n =47

FIGURE 4



Group 1: n = 97, Group 2: n = 183, Group 3: n = 56



Cluster 1: n = 144, Cluster 2: n =114
Cluster 3: n = 31, Cluster 4: n = 47

clusters is displayed on Figure 3. As well Figure 4 displays graphic comparison of the Checklist Subscale Total means for both EI classifications and cluster membership.

In order to facilitate interpretation of the four clusters, the cluster profile points for each group of students were plotted relative to the Teacher Checklist mean and standard deviation for each of the five subscales for the total sample of children. Figure 5 illustrates the cluster profile points for the different groups of students.

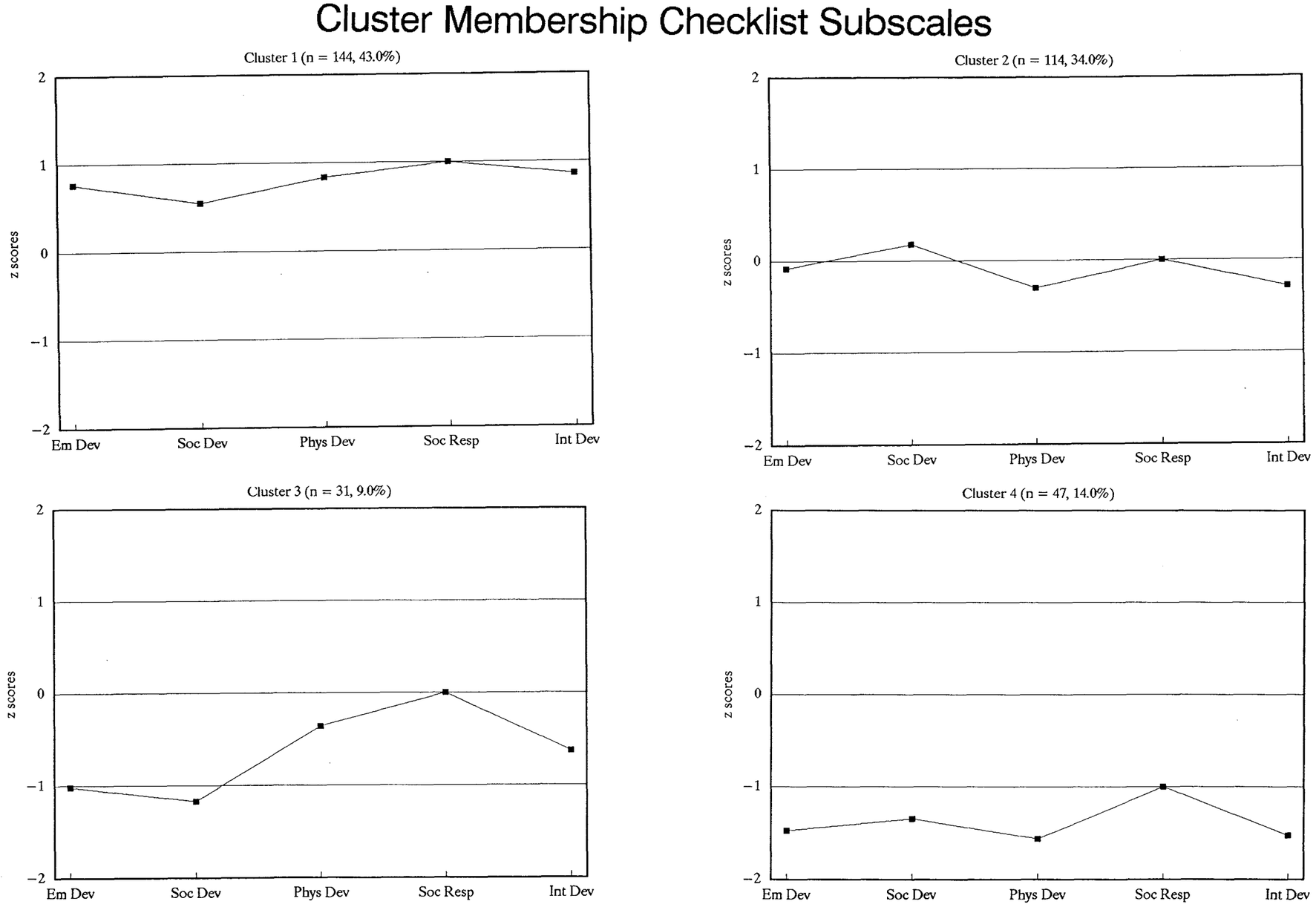
Cluster 1 with 43% of the sample, was consistently higher than the other clusters in all subscale sections. This group is clearly not viewed by teachers as at-risk.

Cluster 2 containing 34% of the sample was higher than Cluster 3 in all areas except for Soc Resp where the z-score was identical.

The pattern indicated in Cluster 3 was that this group was lower than Cluster 2 but consistently higher than Cluster 4 in all areas. This smaller cluster had 9% of the sample.

Fourteen per cent of the sample was in Cluster 4. The members of this cluster were low in all areas. Individuals in this cluster would likely be at-risk. The section displaying crosstabulation between Checklist subscales and EI groups will assist in clarifying who may possibly be at-risk.

FIGURE 5


Cluster Membership Checklist Subscales
Cluster 1 (n = 144, 43.0%)
Cluster 2 (n = 114, 34.0%)
Cluster 3 (n = 31, 9.0%)
Cluster 4 (n = 47, 14.0%)
z scores
2
1
0
−1
−2
Em Dev
Soc Dev
Phys Dev
Soc Resp
Int Dev

2. Analysis of Variance - Checklist Subscales

a. Emotional Development (EM DEV)

In order to find out to what extent Teacher Checklist Subscale clusters are different, a one-way ANOVA was computed. The independent variable in this analysis was the Checklist Subscale, EM DEV. The dependent variable was the Checklist Subscale EM DEV cluster membership. The mean for the total population (336 students) was 24.54. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale EM DEV ratings are presented in Table 14. The significant F ratio for the differences between groups factor [$F(3, 332) = 203.34$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist EM DEV subscale.

TABLE 14
ONE-WAY ANOVA TABLE EM DEV BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	8485.32	2828.44	203.34	.001
Within Groups	332	4618.01	13.91		
Total	335	13103.33			

Post hoc analysis using Tukey-HSD procedure found that all of the clusters were significantly different. Critical value was 2.64 ($p = .05$). Table 15 displays the results indicating the groups that are significantly different. Every teacher rated students who were identified in each of the four clusters.

TABLE 15
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
29.29	1		*	*	*
24.06	2			*	*
18.19	3				*
15.36	4				

(*) denotes pairs of groups significantly different at the 0.05 level

b. Social Development (SOC DEV)

In order to find out to what extent clusters in the Teacher Checklist Subscales are different, a one-way ANOVA was computed. The independent variable in this analysis was Checklist subscale SOC DEV. The dependent variable was the Checklist subscale SOC DEV cluster membership. The mean for the total population (336 students) was 10.12. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale SOC DEV ratings are presented in Table 16. The significant F ratio for the differences between groups factor [$F(3, 332) = 120.30$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist SOC DEV subscale.

TABLE 16
ONE-WAY ANOVA TABLE SOC DEV BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	1595.93	531.98	120.30	.001
Within Groups	332	1468.07	4.42		
Total	335	3064.00			

Post hoc analysis using Tukey-HSD procedure found that most of the clusters were significantly different. Critical value was 1.49 (p = .05). Table 17 displays the results indicating the groups that are significantly different. Groups 3 and 4 were not significantly different. Every teacher rated students who were identified in each of the four clusters.

TABLE 17
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
11.7917	1		*	*	*
10.6491	2			*	*
6.5806	3				
6.0638	4				

(*) denotes pairs of groups significantly different at the 0.05 level

c. Physical Development (PHYS DEV)

In order to find out to what extent the clusters on the Teachers Checklist PHYS DEV subscales were different, a one-way ANOVA was computed. The independent variable in this analysis was Checklist Subscale PHYS DEV. The dependent variable was the Checklist Subscale PHYS DEV cluster membership. The mean for the total population (336 students) was 13.88. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale PHYS DEV ratings are presented in Table 18. The significant F ratio for the differences between groups factor [$F(3, 332) = 233.52$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist PHYS DEV Subscale.

TABLE 18
ONE-WAY ANOVA TABLE PHYS DEV BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	2621.92	873.97	233.52	.001
Within Groups	332	1242.55	3.74		
Total	335	3864.47			

Post hoc analysis using Tukey-HSD procedure found that most of the clusters were significantly different. Critical value was 1.37. (p = .05). Table 19 displays the results indicating the groups that were significantly different. Clusters 2 and 3 were not significantly different from each other. Every teacher rated students who were identified in each of the four clusters.

TABLE 19
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
16.69	1		*	*	*
12.87	2				*
12.65	3				*
8.64	4				

(*) denotes pairs of groups significantly different at the 0.05 level

d. Social Responsibility (SOC RESP)

In order to find out to what extent clusters on the Teachers Checklist Subscale SOC RESP reflected real differences, a one-way ANOVA was computed. The independent variable in this analysis was Checklist Subscale SOC RESP. The dependent variable was the Checklist Subscale SOC RESP cluster membership. The mean for the total population (336 students) was

3.48. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale SOC RESP ratings are presented in Table 20. The significant F ratio for the differences between groups factor [$F(3, 332) = 92.20$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist SOC RESP subscales.

TABLE 20
ONE-WAY ANOVA TABLE SOC RESP BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	152.64	50.88	92.20	.001
Within Groups	332	183.21	.55		
Total	335	335.85			

Post hoc analysis using Tukey-HSD procedure found that all of the clusters were significantly different. Critical value was .53 ($p = .05$). Table 21 displays the results indicating the groups that were significantly different. Every teacher rated students who were identified in each of the four clusters.

TABLE 21

MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
4.10	1		*	*	*
3.41	2			*	*
2.84	3				*
2.15	4				

(*) denotes pairs of groups significantly different at the 0.05 level

e. Intellectual Development (INT DEV)

In order to find out to what extent clusters on the Teacher Checklist Subscale INT DEV reflected real differences, a one-way ANOVA was computed. The independent variable in this analysis was Checklist Subscale INT DEV. The dependent variable was the Checklist Subscale INT DEV cluster membership. The mean for the total population (336 students) was 45.25. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale INT DEV ratings are presented in Table 22. The significant F ratio for the differences between groups factor [$F(3, 332) = 270.46$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist INT DEV subscale.

TABLE 22

ONE-WAY ANOVA TABLE INT DEV BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	39154.14	13051.38	270.46	.001
Within Groups	332	16020.86	48.26		
Total	335	55175.00			

Post hoc analysis using Tukey-HSD procedure found that all of the clusters were significantly different. Critical value was 4.91 ($p = .05$). Table 23 displays the results indicating the groups that were significantly different. Every teacher rated students who were identified in each of the four clusters.

TABLE 23
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
56.31	1		*	*	*
41.54	2			*	*
37.16	3				*
25.70	4				

(*) denotes pairs of groups significantly different at the 0.05 level

f. Checklist Subscales Total Score

In order to find out to what extent clusters on Teacher Checklist Subscale total were different, a one-way ANOVA was computed. The independent variable in this analysis was Checklist Subscale Total. The dependent variable was the Checklist Subscale Total cluster membership. The mean for the total population (336 students) was 97.28. Means and standard deviations were presented earlier in Table 13.

The results of the analysis of variance of the Teacher Checklist Subscale total ratings are presented in Table 24. The significant F ratio for the differences between groups factor [$F(3, 332) = 372.37$, $p = .001$] indicated there were significant differences between clusters on the Phase 3 Teacher Checklist Total.

TABLE 24
ONE-WAY ANOVA TABLE CHECKLIST SUBSCALE TOTAL BY CLUSTER MEMBERSHIP

Source of Variation	df	Sum of Squares	Mean Squares	F Ratio	F Prob.
Between Groups	3	150766.53	50255.51	372.37	.001
Within Groups	332	44807.17	134.96		
Total	335	195573.70			

Post hoc analysis using Tukey-HSD procedure found that all of the clusters were significantly different. Critical value was 8.21 (p = .05). Table 25 displays the results indicating the groups that were significantly different. Every teacher rated students who were identified in each of the four clusters.

TABLE 25
MULTIPLE RANGE TEST - TUKEY PROCEDURE

Mean	Clusters	1	2	3	4
118.18	1		*	*	*
92.54	2			*	*
77.42	3				*
57.85	4				

(*) denotes pairs of groups significantly different at the 0.05 level

3. Crosstabulation of Early Identification Scores with Teacher Checklist Subscale Membership

The crosstabulation of Subscale membership by Early Identification score is presented in Table 26. Although 97 students were classified as high or 1 on the teachers' initial classification, 144 students were clustered in Cluster 1 using the Subscale scores. Of the 144 students in the first cluster, 88 were EI

1 , 55 were from EI 2 and 1 was from EI 3. Meanwhile, 183 students were classified as average (2) on EI, but there were 114 in Cluster 2. Of the 114 students in Cluster 2, six came from EI 1, 94 came from EI 2 and 14 came from EI 3. Using the EI classifications, 56 were classified as low (3), however 31 were in Cluster 3 and 47 were in Cluster 4. The distribution of Cluster 3 indicates three came from EI 1, 19 came from EI 2 and nine came from EI 3. Cluster 4 had 15 students from EI 2 and 32 students from EI 3.

TABLE 26
CROSSTABULATION OF SUBSCALE MEMBERSHIP BY EARLY IDENTIFICATION SCORE

	EI SCORE				
SC MEM		1	2	3	Row Total
	1	88	55	1	144
	2	6	94	14	114
	3	3	19	9	31
	4		15	32	47
	Column Total	97	183	56	336

It would appear that the 32 students who were in SC 4 - EI 3 and the 9 students in SC 3 - EI 3 are at-risk. This group makes up 12% of the sample. It is also likely that the 15 students in SC 4 - EI 2 and the 19 students in SC 3 - EI 2 are likely to be at-risk. This group of likely students made up another 10% of the sample. Upon perusal of the scores of the EI 2 group, which was initially classified as average, it was noted that an average of 15.8 students were more

than one standard deviation below the mean in each of the subscales. It was interesting to note that 17 students were more than one standard deviation below the mean in the INT DEV Subscale in the average group (EI 2).

A Chi-square Test was used to test if there was a significant difference between the Subscale Group Memberships when children with like characteristics were clustered using the Hierarchical Cluster Analysis technique and the initial EI teacher classifications. Chi-square results (220.19, D.F.(6), $p = 0.001$) showed a significant relationship in group membership. Cramer's V statistic at 0.57 indicates a moderate relationship between Phase 1 (EI) teacher classification and Teacher Checklist cluster membership.

Cluster analysis using individual items on the Teacher Checklist was also computed, however it did not indicate group separation as clearly. The results of the Total Checklist analysis are included in Appendix I. The Subscale cluster patterns seem to identify the at-risk students the most clearly and are also most consistent with the AC Questionnaire. This is visually demonstrated in Table 28.

TABLE 27
SUMMARY OF GROUP MEMBERSHIPS FOR THE 3 PHASES

	Phase 1	Phase 2	Phase 3	
Group/Cluster	EI Score	AC SS	CHECKLIST	SUBSCALES
1	97	90	222	144
2	183	118	62	114
3	56	44	31	31
4		84	21	47

A summary of the distribution of group membership for the three phases is displayed in Table 27. It appears that both the Academic Competence Questionnaire and the Teacher Checklist Subscales were able to identify a number of students who would likely be at-risk for learning problems in school.

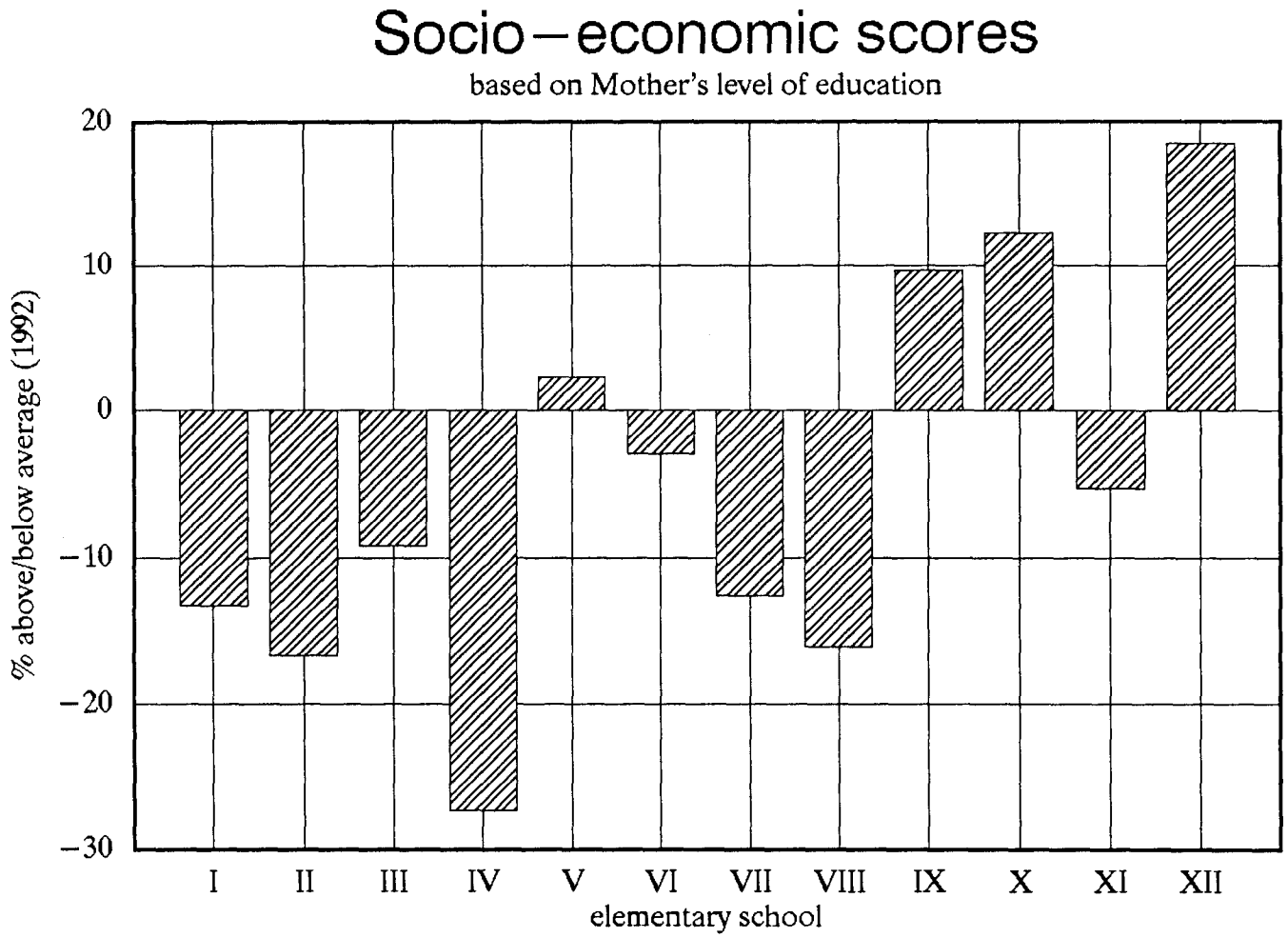
D. TEACHER DIFFERENCES

In order to focus on evidence for validity of the Teacher Checklist, a series of analyses were undertaken. Initially teacher differences were investigated.

1. Correlation of Teacher Ratings with SES Scores

In order to investigate to what extent teachers' Checklist mean ratings reflected real differences, or whether these differences were related to other important factors, correlations of teacher Checklist mean ratings with SES scores were computed. In this case, one important factor to explain differences could be student achievement. An extensive review of research conducted by the local school board indicated that there is evidence to show that student achievement can be predicted from SES scores. A description of how the SES scores were computed was given in Chapter 3. The SES scores were calculated solely on the mother's level of formal education. The mother's level of formal education is known to be a key factor in predicting a child's success in school. See Figure 6 for a graphic presentation of the SES scores of the schools used in the study.

FIGURE 6



% Above/below average refers to combined elementary and secondary parent average.

Correlational analysis was computed using 16 teachers because SES scores were available for these schools. One teacher in the study worked in a private school that did not have SES scores available.

Correlation between teacher mean ratings on the Checklist and school SES scores was .52. This indicates a moderate relationship between teacher ratings and SES scores. That is, although teachers differ from school to school generally in rating students, their ratings are a reflection of the general level of education in the schools.

2. Correlation of Teacher Checklist with AC Questionnaire

Further evidence supporting concurrent validity involved the inclusion of the results of the correlation between the AC Questionnaire and the Teacher Checklist. Correlation between the Teacher Checklist and the Academic Competence Questionnaire was .86. Correlations of the AC Questionnaire with the Teacher Checklist Subscales were: EM DEV .80, SOC DEV .57, PHYS DEV .67, and INT DEV .88. The higher correlations of AC with EM DEV and INT DEV could be because the AC is more academically oriented and thus is more like the two subscales.

3. Crosstabulation of AC Cluster Membership with Teacher Checklist Subscale Cluster Membership

To further establish validity of the Teacher Checklist, crosstabulation of AC Cluster Membership with the Teacher Checklist Cluster Membership was

provided. The crosstabulation of the Academic Competence Membership by Subscale Membership (Table 28) makes an interesting comparison. Although on AC, 84 students (25% of the sample) were in Cluster 4 and appear to be likely at-risk, the Subscales identify a slightly different at-risk group. It is likely that the 12 students from SC 3 - AC 4, the 35 from SC 4 - AC 4, the 6 from SC 3 - AC 3 and the 12 from SC 4 - AC 3 will be in need of some type of intervention or further assessment. This would include 65 students (19% of the sample). The 37 other students who were in Cluster 4 on AC Membership likely may include some at-risk students.

TABLE 28
CROSSTABULATION OF TEACHER CHECKLIST SUBSCALE MEMBERSHIP AND ACADEMIC COMPETENCE MEMBERSHIP

		AC MEMBERSHIP				
SC MEM		1	2	3	4	Row Total
	1	81	50	1	12	144
	2	7	57	25	25	114
	3	2	11	6	12	31
	4			12	35	47
	Column Total	90	118	44	84	336

A Chi-square Test was used to test if there was a significant difference between the Academic Competence Group Membership when children with like characteristics were clustered using the Hierarchical Cluster Analysis technique and the Subscale Cluster Group Membership. Chi-square results

(200.84, D.F.(9), p = 0.001) showed a significant relationship in group membership on the two measures. Cramer's V statistic at 0.45 indicates that the strength of the relationship is moderate. One reason for the differences in the groups may be that the AC Questionnaire is more specifically academic, whereas the Teacher Checklist includes all aspects of the child.

The results of the analyses presented in this chapter and possible implications are discussed in Chapter 5.

CHAPTER 5

SUMMARY AND CONCLUSIONS

The purpose of this chapter is to present a summary of the methods used and the results obtained in the present study as well as to provide an interpretation of the results. Implications for education as well as for further research are also presented.

A. SUMMARY OF METHODS AND RESULTS

The primary purpose of this study was to address the problem, if teachers apply their own standards in making global assessments of children's performance, to what extent are these ratings consistent with subsequent specific ratings using several global items or several specific ratings of skills and abilities?

Research points to the value of early identification and intervention, yet children seem to be missed in the identification procedure and frequently fail in school before they are picked out for intervention (Siegel, 1988). This study attempted to determine the extent to which teachers were able to identify children who are at-risk for failure in school if they use their own methods for classification rather than referring to specific guidelines or benchmarks. This relates to the question "To what extent do each of these measures: Teacher

Rating, Academic Competence Questionnaire, and Teacher Checklist identify children as being at-risk for learning problems."

A second purpose of this study was to validate a Teacher-made Checklist of skills intended for kindergarten or early primary children. With the development of the New Primary Program in British Columbia and the trend away from standardized measures in classrooms, the need for developing a number of benchmarks as guidelines for teachers to follow in order to identify those students who are in need of specific intervention or further assessment is increasing. The Checklist focused on skills that experienced early primary (grade 1) teachers felt that children should possess in order to meet with success in school. The documents from the New Primary Program were also perused for additional benchmarks that should be included. The Checklist was correlated with The Academic Competence Questionnaire, a subtest from the Social Skills Rating System which is a current, reliable, and valid measure.

Scores from the Checklist were computed based on the total score. As well, raw scores from the subscales: Emotional Development, Social Development, Physical Development, Social Responsibility and Intellectual Development were also computed to determine if there would be a difference in the manner in which the children were grouped.

Another question this study attempted to answer was, "What is the evidence of the validity for teachers' ratings of at-risk students?"

This research was conducted in three phases. In the first phase, teachers

were asked to classify their students into one of three categories: high (3), average (2), or low (3). Second, in Phase 2, the teachers were asked to use an individual Likert-type rating scale on each of their kindergarten students. The rating scale used was the Academic Competence Questionnaire taken from the Social Skills Rating System (Gresham & Elliott, 1990). The third phase involved the teachers in rating their students using a 28 item Checklist which had been developed and was being validated during this study. The Checklist also used a Likert-type scoring system. In the initial teacher rating, students were grouped using general or global characteristics. In the second phase students were rated on several global characteristics or abilities. In the final phase students were rated on a number of specific skills thought to affect success in the Primary School Program.

The study involved 17 volunteer kindergarten teachers who rated 336 kindergarten children. The study was conducted in Abbotsford, British Columbia. The children's age ranged from 5.43 to 6.98 years.

Discussion of Research Question 1: How does the Teacher Checklist compare to acceptable reliability standards? In order to establish internal consistency, a series of reliability analyses were conducted. Cronbach's Alpha reliability coefficients for the Checklist total as well as the subscales were all above .90 except for Physical Development which was .86. The majority of these results meet the standards for minimum criteria of internal consistency of .90 for

placement and diagnostic test scores as well as internal consistency of .80 for screening measures (Bracken, 1987). Therefore, we have established that the Teacher Checklist could be used as a part of a screening or diagnostic measure.

Discussion of Research Question 2: To what extent do each of these measures: Teacher Rating, Academic Competence Questionnaire, and Teacher-made Checklist identify children as being at-risk for learning problems?

In Phase 1 of the study, teachers classified 97 students as high (Group 1), 183 students as average (Group 2), and 56 students as low or at-risk (Group 3). This first phase involved the most general classification where teachers used their own methods to rate children.

A Hierarchical Cluster Analysis was used in Phase 2 to cluster the students who were most similar on the Academic Competence Questionnaire. The Academic Competence Questionnaire was more specific than the broad classifications in Phase 1 in that it investigated 9 specific areas. These areas included: intellectual ability, behaviour, motivation, parental support, academic functioning, as well as achievement in reading, and mathematics. Using teacher ratings of children's ability or performance, the Cluster Analysis of the ratings indicated that four groups were a more reasonable solution. Clusters 3 and 4 were closest in mean scores, 83.41 and 81.25 respectively. These two groups were rated similarly across the majority of areas. The most outstanding difference between students on Cluster 3 and Cluster 4 was the considerable

difference between Cluster 3 and Cluster 4 students in items of "the child's overall motivation to succeed academically" and "the child's overall classroom behaviour when compared with other children in the classroom". Cluster 4 students were generally rated as being in the lowest 10-20% of their respective class in terms of motivation and behaviour although both Clusters 3 and 4 students were rated as lower in most aspects of this scale. Based on the Phase 1 classifications, 78% of the students who were classified in Group 1 were also in Cluster 1 on the Academic Competence Questionnaire. There was some variability between the EI Group 1 and Cluster 1, however 97% of the students who were initially classified as High or Group 1 were in either AC Cluster 1 or 2. It is unlikely that any of these students would be at-risk for learning problems. The group that was initially classified as average or EI 2 is interesting in that 38% of this group (70 students) was in Cluster 3 or 4 on the AC ratings. This number of students is comparable with the 67 students from EI 2 who were rated below the 25th percentile on the AC ratings. It is possible that these students who were originally classified as average in Phase 1 but later classified in Cluster 3 or 4 in Phase 2 may be at-risk for learning problems in school. This is the group that may include the children who are typically missed when global measures are used. Of the 56 students who were originally classified as low on Phase 1, 55 were in Cluster 3 or 4 on Phase 2. These children are likely those who stand out and who could be identified as at-risk using global methods. The results of the Cluster Analysis indicated that when

the guidelines became more defined, the discrimination among the groups was finer. It can be speculated in this analysis that the students who were in Clusters 3 (Mean = 83.41), and 4 (Mean = 81.25) are likely to be at-risk for learning difficulties in school. This group includes 37% of the sample. This figure, while high, is consistent with the estimate that approximately one third of America's school children are educationally at-risk (Roth, McCaul, Barnes, 1993). Although there may be some students in this group who may not have learning problems, per se, it is reasonable to assume all would benefit from receiving stimulating early intervention (Cornell & Gottfried, 1976). However if learning disabled students were missed, the problem of failure would have to be dealt with. It is likely that some of the children identified in Clusters 3 and 4 from Phase 2 may be at-risk for future learning problems and would benefit from further assessment or monitored for intervention.

Initial analysis of Phase 3, using individual items on the Checklist yielded a four cluster solution. The Cluster Analysis using the 28 individual items to identify similar clusters did not give a clear picture of group separation. (Results of this analysis are included in Appendix I for information.) However, the Hierarchical Cluster Analysis that was used to cluster similar individuals on the Checklist Subscales, which made up Phase 3 of the study, appeared to be most interpretable. This Checklist is different from the Questionnaire used in Phase 2 as it is skill specific and looks at all areas of the child's development. The AC Questionnaire in Phase 2 investigated 9 specific areas that are more

academically inclined while the Checklist looked at 28 specific skills. These skills were grouped into five specific groups which are referred to as subscales. The clusters derived using these five scales: Emotional Development, Social Development, Physical Development, Social Responsibility, and Intellectual Development were more easily interpretable. When the teachers used this specific skills Checklist to rate their students, the Cluster Analysis of the ratings indicated that four groups were a more reasonable solution. The four clusters were significantly different. Clusters 2 and 3 were similar in Physical Development whereas Clusters 3 and 4 were similar in Social Development. The most outstanding differences between the students in the four clusters can be seen in the Intellectual, Emotional and Social Responsibility Subscales. It appeared that of the 97 students who were originally classified as high (EI 1), 97% were in either Clusters 1 or 2. It is unlikely that any of these children would be at-risk. The average group is interesting in that 81% of the children who were classified as average were in Clusters 1 or 2 in Phase 3, however 19% of this average group were in Clusters 3 or 4. This 19% is possibly the group of students who would be missed. The children who were initially classified as low (EI 3) tend to be spread over the four clusters with 27% of these low students fitting into Clusters 1 and 2. These children are a puzzle in that although the teacher initially classified them as low, her opinion may have been in some way biased but changed when she focused on specific skills. Of the students initially classified as low, 73% remained in Clusters 3 and 4. These are

the children who may be to be at-risk for learning problems. It is predicted that the students who are in Cluster 3 and 4 are at-risk. This would make up 23% of the total sample. Although the Chi-square Test of the Crosstabulations between the cluster membership in Phases 2 and 3 shows a moderate relationship between the groups, approximately 15% of the sample does not match in Clusters 3 and 4. A reasonable explanation for the part of the sample that doesn't match may be that the Phase 3 Checklist is skill specific consequently the discrimination between students is finer. This is an indication that for these students who are in Clusters 3 or 4 in one measure but not the other, more assessment or intervention follow-up is required.

Based on the results of the Hierarchical Cluster Analysis, there are indications that there was a significant difference between the clusters in each of Phases 2 and 3. The ANOVA's that were computed for the AC Questionnaire, the Checklist and the individual subscales all indicated significant differences in group separation. Further Post hoc analysis indicated that Clusters 3 and 4 on the AC were not significantly different. This is likely an indication that the nine items are too broad to make fine discriminations. It can be safely concluded that some students in Cluster 3 are likely at-risk for learning in school. Checklist subscale Clusters 3 and 4 were not significantly different in Social Development. As well Clusters 2 and 3 were not significantly different in Physical Development. This may be an indication that Social and Physical Development are not effective indicators of children who

may be at-risk. Another factor is that both SOCDEV and PHYSDEV had fewer items than EMDEV and INTDEV so they were not as effective in indicating student differences. Based on Cramer's V statistic there is a moderate relationship between the groups in Phases 2 and 3. Depending on the method used, there will be a small difference as to who is identified as being at-risk. However it is clear that when the guidelines for rating children changed from general to more specific, it appeared that teachers made finer discriminations, therefore, the use of 3 groupings was insufficient for classification purposes. Using more specific guidelines, four groups emerged under Cluster Analysis. It would be reasonable to assume that by using both the Academic Competence Questionnaire and the Checklist Subscale Classifications most at-risk students would be identified.

The 56 students who were identified as at-risk in the initial broad classification undoubtedly stood out. With more skill specific ratings, a larger number of students stood out as being at-risk. It would be reasonable to consider Clusters 3 and 4 in both the Checklist Subscale Analysis and AC Questionnaire to most likely be students who would require further assessment or intervention. It may also be an indication that the two measures used are suitable for initial screening for the complete group but that more assessment may be required for those students who are in Clusters 3 and 4 in either measure.

Discussion of Research Question 3: What is the relationship between the Academic Competence Questionnaire and the Teacher Checklist? In order to provide concurrent validity evidence for the Teacher Checklist, Pearson Correlation Coefficients were computed using AC Standard Scores (ACSS, Checklist (CH) Total, and Checklist (CH) Subscales: EMDEV, SOCDEV, PHYSDEV, and INTDEV. The correlation between ACSS and CH Total was high. Similarly, the relationships between ACSS, INTDEV and EMDEV were also high. All correlations were significant at the $p = .001$ level. Strong concurrent validity is evidenced by the high correlations.

Discussion of Research Question 4: "What is the evidence of the validity of teachers' ratings of at-risk students?" Several different approaches were used to provide validity information.

To investigate to what extent teachers' mean ratings were reflecting real differences, a correlation analysis was computed between teacher mean ratings and SES scores of each school. The moderate correlation between these scores indicates that some of the differences in rating were a reflection of the general level of education in the schools. Although there are indications that teachers vary to some extent in rating students, it must be noted that the sample is varied. Some of the schools in the study had a higher proportion of immigrant children where English was their second language. Teachers rated their students on the basis of their classes, therefore, it is expected that because the classes are different teachers will vary in ratings. As noted earlier, the SES of

the schools was very different and may have been a contributing factor to these differences. From this study, it appears that teachers' ratings vary according to the nature of their classes.

Another evidence of validity is that all the teachers used the scales to identify children in each of the groups. That is, every teacher identified children in each group or cluster.

The high correlation of the Checklist with the Academic Competence Questionnaire is an indication of concurrent validity.

The initial contribution of the teachers initially in generating the items and subsequently, the ten experienced teachers who prioritized the items both contributed to content validation.

A test is considered to have face validity if it appears to measure a construct that is meaningful to laypersons or typical examinees (Crocker & Algina, 1986). The Checklist appears to be a reasonable screening measure and thus has face validity.

The Cross-tabular analyses of the AC Questionnaire cluster membership with the Checklist Subscales cluster membership indicated that there was a moderate relationship between the groups. This is yet another indication of validity.

Of course, the strongest predictor of validity is related to follow-up. The validity question cannot be fully answered until a follow-up of these children is carried out.

B. CONCLUSIONS

The difficulties involved with identifying children who may be at-risk for learning disabilities were evidenced in this study. From the study, it appeared that the method used for identification of at-risk children makes a difference as to who is identified. Using the present data, the students did not consistently fall into 3 groups. Because of method variance, instead of 3 groups a fourth group emerged. One possibility is teachers had four groups in their minds when they were initially rating students, however they were limited to three categories. But, the skill specific checklist sensitized the teacher to focus in on specific areas of the child's development and thus the four groups emerged.

It is interesting to note that in-class attention span, distractibility, or memory span as well as verbal fluency are considered to be "the best overall in-class indicators of future academic achievement when considered in relation to other potential in-class warning signs" (Simner, 1983, p.24). Upon perusal of the individual items on the Checklist, it was interesting to note that Cluster 4 was two Z-scores below the mean and Clusters 2 and 3 were one z-score below the mean in attention span. Also, Clusters 3 and 4 were both one and two z-scores below the mean, respectively in the oral language item. There is some relationship with these scores and Simner's study.

The data and analyses appear to indicate that teachers are uniform in classifying students. It can be expected that the individual teacher's rating will vary in relationship to the variance in the sample. As evidenced in the

Checklist total scores, using total scores can be misleading. Subscale cluster memberships appeared to be more reasonable. It can be argued that students who were members of Cluster 4 stood out as being at-risk; however, a number of students who clustered in Cluster 3 are also likely to be at-risk and should be monitored on intervention.

It can be concluded from this study that although young children vary in their rate of development, it is important to identify those who would benefit from specific intervention in order to stimulate some areas of development that may be otherwise delayed. There is evidence that the "most meaningful model of intervention would be one that would increase successful school experiences and "intervene" (i.e., appropriately program) before serious school problems develop" (Roth, McCaul, Barnes, 1993, p. 349).

Based on Cramer's V statistic, the Early Identification Classification indicates a slightly stronger relationship with AC Questionnaire cluster membership (0.63) than Early Identification Classification with Checklist Subscale cluster membership (0.57). The same statistic indicates a moderate but weaker relationship between AC Questionnaire Cluster membership and Checklist Subscale membership (0.45). An explanation of the strength of these relationships is that the classification focus in each phase was slightly different. Results of the study support the contention that the Checklist is useful in identifying at-risk students. By using the Checklist, there is some indication that total scores are consistent with teacher nominations but identify more at-

risk students than initial nominations. Cluster Analysis within Subscales differentiated students into four groups when teachers were asked to make finer discriminations. Therefore, it can be concluded that teachers need guidelines to assist them in identifying those children who are at-risk. A screening battery that focuses on essential skills that children need in order to make progress in the classroom is essential.

C. LIMITATIONS OF THE STUDY

The study was conducted over a short period of time. A longitudinal study would perhaps be useful for this topic. Also, it may have been more informative to use a measure that used a different scoring system in one of the phases. In this type of study the teacher is the rater so the teacher really becomes the measuring instrument. The instrument then is only as good as the teacher's ability to evaluate and organize information (Cadwell & Jenkins, 1986). A multi-method approach would generally have a higher accuracy rate. That is, the use of a measure where the students would have to respond would involve a different type of scoring system.

The Checklist items need to be further analyzed to see if they are all effective predictors of children at-risk for learning problems. Items in each of the Subscales should be analyzed to see if they fit more than one domain. The Subscales with very few items should have more items added. It would be interesting to use the Checklist as part of a screening battery with other

measures. Measures of receptive and expressive language as well as a measure of pencil skills would give added information and provide for a variety of responses.

Boys and girls could have been analyzed separately to see if there were differences between sexes in assessing learning problems.

Analysis of age differences may have added to the information about different factors that should be considered when identifying at-risk students.

The question of prediction cannot be fully answered in this type of study. A follow-up study when the students are in third or fourth year Primary would help to validate the accuracy of prediction.

D. IMPLICATIONS FOR EDUCATION

Early identification of children who are in need of intervention has been emphasized in theoretical research studies. Recently researchers and practitioners have become concerned with the increasing number of today's children who are considered to be "at-risk for school failure" (Davis & McCaul, 1990, p.348). The need for early identification continues to exist. Presently education is undergoing change. This change is invoking a trend that is focusing on a more authentic approach to assessment. Assessment that relates to curriculum is being promoted (Duran, 1989). Assessment tools that are curriculum focused, provide useful information and that are quick and easy to administer will be required. Because teachers vary in the type of training they

receive, as well as years of experience they have, it is important to have some basic standard guidelines or benchmarks for them to help to identify the children in need of intervention.

E. IMPLICATIONS FOR RESEARCH

Further research is necessary to validate the Checklist. This can be accomplished by doing a follow-up study of the students who were members of Clusters 3 and 4 in approximately two years to find out the number from these groups who developed learning problems. It is speculated that the students who are usually missed would likely be in Cluster 3.

REFERENCES

Algozzine, B. & Curran, T.J. (1978). Teachers' predictions of children's school success as a function of their behavioral tolerances. Journal of Educational Research, 72, 344-347.

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1985). Standards for educational and psychological testing. Washington DC: Author.

Anastasiow, N.J. (1990). Implications of the neurobiological model for early interventions. In S.J. Meisels & J.P. Shonkoff (eds.), Handbook of early childhood intervention (pp. 196-216). Cambridge: Cambridge University Press.

Anderberg, M.R. (1973). Cluster analysis for applications. Academic Press.

Badian, N.A. & Serwer, B.L. (1975). The identification of high-risk children: A retrospective look at selection criteria. Journal of Learning Disabilities, 8 May, 5.

Barkley, R.A. (1981). Hyperactive children: A handbook for diagnosis and treatment. New York: Guilford Press.

Becker, W.C. & Gersten, R.A. (1982). A follow-up of follow through: The later effects of the direct instruction model on children in fifth and sixth grades. American Educational Research Journal, 19, 75-92.

Beery, K.E. (1968). Comprehensive research, evaluation and assistance for exceptional children. Exceptional Children, 35, 223-28.

Black, F.W. (1974). Self-concept as related to achievement and age in learning disabled children. Child Development, 45, 1137-1140.

Book, R.M. (1980). Identification of educationally at-risk children during the kindergarten year: A four-year follow-up study of group test performance. Psychology in the Schools, 17, 153-158.

Bowlby, J. (1969). Attachment and loss. Vol.1, Attachment. New York: Basic Books.

Bracken, B.A. (1987). Limitations of preschool instruments and standards for minimal levels of technical adequacy. Journal of Psychoeducational Assessment, 4, 313-326.

Bredekamp, S. & Shepard, L. (1989). How best to protect children from inappropriate school expectations, practices, and policies. Young Children, March, 14-24.

Brophy, J.E. (1983). Research on the self-fulfilling prophecy of teacher expectations. Journal of Educational Psychology, 75, 631-661.

Cadwell, J. & Jenkins, J. (1986). Teachers' judgments about their students: The effect of cognitive simplification strategies on the rating process. American Educational Research Journal, 23(3), 460-475.

Casto, G. & Mastropieri, M. (1986). The efficacy of early intervention programs. Exceptional Children, 52(2), 417-424.

Cooper, D.H. & Speece, D.L. (1988). A novel methodology for the study of children at-risk for school failure. The Journal of Special Education, 22, 186-197.

Cooper, D.H. & Speece, D.L. (1990). Maintaining at-risk children in regular settings: Initial effects of individual differences and classroom environments. Exceptional Children, 57(2), 117-126.

Cornell, E.H. & Gotfried, A.W. (1976). Intervention with human premature infants. Child Development, 47, 32-39.

Crocker, L. & Algina, J. (1986). Introduction to classical and modern test theory. Holt, Rinehart & Winston Inc.

Cunnigham, C.E. & Barkley, R.A. (1978). The role of academic failure in hyperactive behaviour. Journal of Learning Disabilities, 11, 15-21.

Darlington, R.B. (1981). The consortium for longitudinal studies. Education Evaluation and Policy Analysis, 3, 37-45.

Davis, W.E., & McCaul, E. (1990). At-risk children and youth: A crisis in our schools and society. Orono, ME: College of Education, University of Maine. In Roth, M., McCaul, E., & Barnes, K. (1993) Who becomes an "at-risk" student? the value of a kindergarten screening battery. Exceptional Children, 59(4). 348-358.

DeHirsch, K., Jansky, J., & Langford, W.S. (1966). Predicting reading failure. Harper & Row.

Diamond, K.E. (1990). Effectiveness of the revised Denver Developmental Screening Test in identifying children at risk for learning problems. Journal of Educational Research, 83(3), 115-122.

Douglas, V.I. & Peters, K.G. (1979). Toward a clearer definition of the attentional deficit of hyperactive children. In G.A. Hale & M. Lewis (eds.) Attention and the development of cognitive skills, (pp. 173-247). New York: Plenum Press.

Dunst, C.J., Snyder, S.W., & Mankinen, M. (1987). Efficacy of early intervention. In M. Wang, H. Walberg, & M. Reynolds (eds.) Handbook of special education: Research and practice, Vol. 3 (pp. 259-290). Oxford, England: Pergamon Press.

Duran, R.P. (1989). Assessment and instruction of at-risk hispanic students. Exceptional Children 56(2), 154-158.

Egan, O. & Archer, P. (1985). The accuracy of teachers' ratings of ability: A regression model. American Educational Research Journal, 22(1), 25-34.

Everitt, B. (1977). Cluster analysis. Social Sciences Research Council. Heineman Educational Books.

Fedoruk, G.M. & Norman, C.A. (1991). Kindergarten screening predictive inaccuracy: First-grade teacher variability. Exceptional Children, 57(3), 258-263.

Feshbach, S., Adelman, H. & Fuller, W.W. (1974). Early identification of children with high risk of reading failure. Journal of Learning Disabilities, 7(10), 639-644.

Fincham, J.G. (1977). A comparison of moral judgment in learning disabled and normal achieving boys. Journal of Psychology 96, 153-160.

Funder, D.C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. Psychological Bulletin, 101, 75-90.

Gerber, M.M. & Semmel, M.I. (1984). Teacher as imperfect test: Reconceptualizing the referral process. Educational Psychologist, 19, 137-148.

Glass, G.V. & Hopkins, K.D. (1984). Statistical methods in education and psychology. New Jersey: Prentice-Hall.

Greenwood, L.K. & Morton, L.L. (1989). Using teacher ratings to determine if the learning disabled are ready for the regular classroom. B.C. Journal of Special Education, 13(1), 67-73.

Gresham, F.M., Reschly, D.J. & Carey, M.P. (1987). Teachers as tests: Classification accuracy and concurrent validation in the identification of learning disabled children. School Psychology Review, 16(4), 543-553.

Gresham, F.M. & Elliott, S.N. (1990). Social skills rating system: Manual. American Guidance Service, Inc.

Guralnick, M.J. (1991). The next decade of research on the effectiveness of early intervention. Exceptional Children, 58(2), 174-183.

Hair, J., Anderson, R.E. & Tathan, R.L. (1987). Multivariate Data Analysis. New York: McMillan.

Hallahan, D.P. & Cruickshank, W.M. (1973). Psychoeducational foundations of learning disabilities. New Jersey: Prentice-Hall Inc.

Hammill, D.D., Bartel, N.R. & Bunch, G.O. (1984). Teaching children with learning and behaviour problems. (Canadian ed.). Toronto: Allyn and Bacon.

Harbin, G.L. (1988). Implementation of P.L. 99-457: State technical assistance needs. Topics in Early Childhood Special Education, 8, 24-36.

Harrington, R. & Ford, L. (1988). A national study of needs in preschool assessment. Paper presented at the meeting of the National Association of School Psychologists, Chicago, Il. in Lee, S. et al. Diagnostique, 16(2-3).

Hayden, A.H. (1974). Perspectives of early childhood education in special education. In N.G. Haring (ed.), Behaviour of exceptional children: An introduction to special education. Columbus, Ohio: Charles E. Merrill.

Heward, W. & Orlansky, M. (1988). Exceptional Children. Columbus, Ohio: Charles E. Merrill.

Horn, W.F. & O'Donnell, J.P. (1984). Early identification of learning disabilities: A comparison of two methods. Journal of Educational Psychology, 76(6), 106-1118.

Horn, W.F. & Packard, T. (1985). Early identification of learning problems. Journal of Educational Psychology, 77, 597-607.

Howell, K.W. (1986). Direct assessment of academic performance. School Psychology Review, 15, 324-335.

Johnson-Fedoruk, G.M. (1990). Teacher recommendations for at-risk students: The relative impact of academic and nonacademic factors. B.C. Journal of Special Education, 14(3), 287-293.

Johnson, G.M. & Johnson, J.A. (1992). Teacher-student congruence and first grade reading achievement: Predictive validation of an interactional educational construct. B.C. Journal of Special Education, 16(2), 183-197.

Judy, J. (1986). Early screening is essential for educational accountability: Response to Salzer and to Shepard and Smith. Educational Leadership, 44(3), 87-88.

Kerlinger, F.N. (1973). Foundations of behavioral research. New York: Holt, Rinehart & Winston Inc.

Kerlinger, F.N. (1979). Behavioral research: A conceptual approach. New York: Holt, Rinehart & Winston Inc.

Kishor, N. (1992). Information utilization in teachers' judgments of students' performance: Evidence for a systematic error hypothesis. (Unpublished. Paper presented at CAEP, UPEI)

Klein, A.E. (1978). The validity of the Beery Test of Visual-motor Integration in predicting achievement in kindergarten, first and second grades. Educational and Psychological Measurement 38, 457-461.

Kochanek, T.T., Kabacoff, R.I. & Lipsitt, L.P. (1990). Early identification of developmentally disabled and at-risk preschool children. Exceptional Children, 56(6), 528-538.

Koppitz, E.M. (1971). Children with learning disabilities. Grune & Stratton.

Kornblau, B. (1982). The teachable pupil survey: A technique for assessing teachers' perceptions of pupil attributes. Psychology in the schools, 19, 170-174.

Lacayo, N., Sherwood, G. & Morris, J. (1981). Daily activities of school psychologists: A national survey. Psychology in the Schools, 18, 184-190.

Lazar, I., Darlington, R., Murray, H., Royce, J. & Snipper, A. (1982). Lasting effects of early education. Monographs of the Society for Research in Child Development (Serial No. 195).

Lee, S., Unruh, L., Ford, L. & Mace, W. (1991). The reliability and validity of the Cognitive Levels Test (CLT) for preschool children. Diagnostique, 16(2-3), 94-100.

Lerner, J.W. (1976). Children with learning disabilities. (2nd ed.). Boston, MA: Houghton Mifflin Company.

Levine, M.D., Oberklaid, F. & Meltzer, L. (1981). Developmental output failure: A study of low productivity in school-aged children. Pediatrics, 67(1), 18-25.

Lichtenstein, R. (1981). Comparative validity of two preschool screening tests: Correlational and classificational approaches. Journal of Learning Disabilities, 14(2), 68-73.

Mardell-Czudnowski, C. (1990). Infant and toddler intervention: Project links. B.C. Journal of Special Education, 14(2), 110-115.

Marsh, D.T., Stoughton, N.L. & Williams, T.A. (1985). Effects of role, gender, age and parental status on perception of childhood problems. Exceptional Children, 52, 170-177.

Meisels, S.J. (1985a). Developmental screening in early childhood: A guide. Washington, D.C.: NAEYC.

Meisels, S.J. (1985b). A functional analysis of the evolution of public policy for handicapped young children. Educational Evaluation and Policy Analysis.

Meisels, S.J. (1987). Uses and abuses of developmental and school readiness testing. Young Children, 42(2), 4-6, 68-73.

Meisels, S.J., Harbin, G., Modigliani, K. & Olson, K. (1988). Formulating optimal state early intervention policies. Exceptional Children, 55(2), 159-165.

Mercer, Cecil D., Algozzine, B. & Trifiletti, J. (1988). Early identification - An analysis of the research. Learning Disabilities Quarterly, 11, 176-188.

Messick, S. (1984). Assessment in context: Appraising student performance in relation to instructional quality. Educational Researcher, 13(3), 3-8.

NAEYC (National Association for the Education of Young Children). (1988). Position statement on standardized testing of young children three through eight years of age. Young Children, March, 42-47.

Oliver, Carolyn E. (1990). A sensorimotor program for improving writing readiness skills in elementary-age children. The American Journal of Occupational Therapy, 44, February, 111-116.

Primary Program. (1989). Program Development. Ministry of Education, Province of British Columbia.

Reynolds, C., Gutkin, T., Elliott, S. & Witt, J. (1984). School psychology: Essentials of theory and practice. New York: Wiley.

Romesburg, H.C. (1990). Cluster analysis for researchers. Robert E. Krieger Publishing Company.

Roth, M., McCaul, E., & Barnes, K. (1993) Who becomes an "at-risk" student? the predictive value of a kindergarten screening battery. Exceptional Children. 59(4). 348-358.

Salvia, J. & Ysseldyke, J. (1981). Assessment in special and remedial education. (2nd ed). Boston: Houghton-Mifflin.

Sattler, J.M. (1988). Assessment of children. San Diego, CA: Jerome M. Sattler.

Self, H., Benning, A., Martson & Magnusson, D. (1991). Cooperative teaching project: A model for students at risk. Exceptional Children, 58(1), 26-33.

Shepard, L.A. & Smith, M.L. (1986). Synthesis of research on school readiness and kindergarten retention. Educational Leadership, 44(3), 78-86.

Shepard, L.A. & Smith, M.L. (1988). Escalating academic demand in kindergarten: Counterproductive policies. The Elementary School Journal, 89, 133-145.

Siegel, L.S. (1988). A system for the early detection of learning disabilities. Canadian Journal of Special Education, 4(2), 115-122.

Simner, M.L. (1983). The warning signs of school failure: An updated profile of the at-risk kindergarten child. Topics in Early Childhood Special Education, October, 17-26.

Solomon, D. & Kendall, A.J. (1977). Dimensions of children's classroom behaviour as perceived by teachers. American Educational Research Journal, 14, 411-421.

Speece, D.L. & Cooper, D.H. (1990). Ontogeny of school failure: Classification of first grade children. American Research Journal, 27(1), 119-140.

Stevens, J. (1986). Applied multivariate statistics for the social sciences. London: Lawrence Erlbaum Assoc., Publishers.

Stone, B., Cundick, B.P. & Swanson, D. (1988). Special education screening system: Group achievement test. Exceptional Children, 55(1), 71-75.

Strag, G.A. (1972). Comparative behavioral rating of parents of severely mentally handicapped, special learning disability and normal children. Journal of Learning Disabilities, 5, 631-635.

Swanson, L. & Watson, B.L. (1989). Educational and psychological assessment of exceptional children. Merrill.

Tabachnick, B.G. & Fidell, L.S. (1983). Using multivariate statistics. Harper & Row.

Tucker, J.A. (1985). Curriculum-based assessment: An introduction. Exceptional Children, 52, 199-204.

Ysseldyke, J.E., Thurlow, M., Graden, J., Wesson, C., Algozzine, B. & Deno, S. (1983). Generalizations from five years of research on assessment and decision making: The University of Minnesota Institute. Exceptional Education Quarterly, 4(1), 75-91.

APPENDIX CONTENTS

Appendix A

Information Sheet

Dear _____,

Thank-you for participating in this study. Please fill in the requested information on the following form.

1. Number of years of university training_____
2. Number of years of experience _____
3. Number of years of teaching at this level _____
4. Number of divisions in the school _____
5. Total school population _____
6. General SES of the school _____
7. Is this a typical Year I class or was there a specific selection process? ...for instance, were the students generally older, younger. higher (lower) ability, or ESL? Mention anything that would make your class different from the typical year 1 class.

_____

_____

_____

_____

**Please Note. When filling in the Appendix A Form, for those students who are low, use just a word or 2 to indicate the area of concern. For instance is it behavior, ability, language, immaturity, etc...

If you have any questions, please call me at home at 853-8989 or at Wm. A. Fraser School 859-6794.

***** When you have completed this first phase, please call me. I will pick it up and give you Phase 11.**

Thank-you for your interest!
Roz Francis

Appendix B

Phase 1: Early Identification

Please rank your students as Low (at risk), Average, or High. **Use each child's given name and first initial of his/her surname.**

Teacher's ID. number:_____ School:_____

Name	High	Average	Low (At Risk)	Specific area of Concern
1.				
2.				
3.				
4.				
5.				
6.				
7.				
8.				
9.				
10.				
11.				
12.				
13.				
14.				
15.				
16.				
17.				
18.				
19.				
20.				
21.				
22.				
23.				
24.				
25.				

Appendix C

Phase 2: Academic Competence Questionnaire

Case # _____

Name _____ Date _____

The nine items require your judgments of this student's academic or learning behaviors as observed in your classroom. Compare the student with other children who are in the same classroom.

Rate all items using a scale of 1 to 5. Circle the number that best represents your judgement. The number 1 indicates the lowest or least favorable performance, placing the student in the lowest 10% of the class. Number 5 indicates the highest or most favorable performance, placing the student in the highest 10% compared with other students in the classroom.

FOR OFFICE USE ONLY		Lowest 10%	Next Lowest 20%	Middle 40%	Next Highest 20%	Highest 10%
	49. Compared with other children in my classroom, the **overall academic performance** of this child is:	1	2	3	4	5
	50. In **reading**, how does this child compare with other students?	1	2	3	4	5
	51. In **mathematics**, how does this child compare with other students?	1	2	3	4	5
	52. In terms of grade-level expectations, this child's skills in **reading** are:	1	2	3	4	5
	53. In terms of grade-level expectations, this child's skills in **mathematics** are:	1	2	3	4	5
	54. This child's **overall motivation** to succeed academically is:	1	2	3	4	5
	55. This child's **parental encouragement** to succeed academically is:	1	2	3	4	5
	56. Compared with other children in my classroom this child's **intellectual functioning** is:	1	2	3	4	5
	57. Compared with other children in my classroom this child's **overall classroom behavior** is:	1	2	3	4	5
AC	SUM OF COLUMN					

Stop. Please check to be sure all items have been marked.

For Office Use:
Academic Competence Total Rating _____
Competence Level ___Below ___Average ___Above

Standard Score _____ Percentile Rank_____
SEM +/- _____ Confidence Level ____68% ____95%
Confidence Band _____to _____

Gresham & Elliott, 1990

Appendix D

Phase 3: Experimental Teacher Checklist

Case # _____
Student's Name _____ **Date** _____
Birth Date _____ **Age** _____
School _____ **Teacher** _____

Purpose and Instructions:
The purpose of this checklist is to identify those students who may be at risk and may require intervention in order to help them to meet with success in school.

1. Base rating on the student's current behavioral characteristics.
2. Compare the student with normal children his/her age.
3. Rate each item quickly and base the rating on your own experience with the student.
4. **5 indicates high performance and 1 indicates low performance.**

Circle 5 if skill development is high.
Circle 4 if skill development is between average and high.
Circle 3 if skill development is average.
Circle 2 if skill development is between average and low.
Circle 1 if skill development is low.

Emotional Development

1. Is able to share and take turns.

Low				**High**
1	2	3	4	5

2. Has a good positive attitude toward school.

1	2	3	4	5

3. Is willing to take risks. (i.e. answering questions, writing, predicting)

1	2	3	4	5

4. Has a willingness to learn.

1	2	3	4	5

5. Is able to follow routines.

1	2	3	4	5

6. Is able to work independently.

1	2	3	4	5

7. Is persistent, even when task appears challenging.

1	2	3	4	5

Appendix D

Checklist...

Social Development

8. Is able to co-operate with others.
   1 2 3 4 5

9. Is able to co-operate and function in a large group.
   1 2 3 4 5

10. Has skills in making friends, sharing, doing for others.
    1 2 3 4 5

Physical Development

11. Large muscle development - walks, runs, climbs.
    1 2 3 4 5

12. Fine muscle development - knows how to use tools, scissors, pencils, glue.
    1 2 3 4 5

13. Has early printing skills, can hold a pencil print name properly, knows left to right progression, can draw circles and lines.
    1 2 3 4 5

14. Practices personal safety.
    1 2 3 4 5

Social Responsibility

15. Looks after personal belongings and school property.
    1 2 3 4 5

Intellectual Development
Using Language For Communication

16. Has acquired some listening skills and is able to attend in a large group situation.
    1 2 3 4 5

17. Has had extensive exposure to literature and has an enthusiasm for books and stories.
    1 2 3 4 5

18. Has some book knowledge; print has meaning.
    1 2 3 4 5

19. Oral Language (Speaking Vocabulary) is well developed for this level.
    1 2 3 4 5

20. Shows an increasing ability to retell stories.
    1 2 3 4 5

Appendix D

Checklist...

21. Can read familiar pattern books.
    1 2 3 4 5

22. Recognizes own name in print.
    1 2 3 4 5

23. Uses some letter-sound relationships.
    1 2 3 4 5

24. In writing, uses letter names as sound cues, may use one letter to represent a whole word. (Emergent Writing)
    1 2 3 4 5

25. Can recognize # 1 to 10, count to 30 and use one to one correspondence.
    1 2 3 4 5

26. Can classify objects by single attributes.
    1 2 3 4 5

27. Thinking skills and problem solving skills appropriate for this level.
    1 2 3 4 5

28. Attention span is developed appropriately for this level - the ability to stay on task.
    1 2 3 4 5

If you have any questions, call Roz Francis at 853-8989 or 859-6794.

For Office Use:

Total Raw Score _____

Appendix D.1

Development of Teacher Checklist

Step 1: Approximately 40 Grade one teachers brainstormed to generate a list of skills or qualities they felt a child should possess in order to meet with success in first grade (second year primary). Refer to Appendix D.2 for this original list.

Step 2: Two years later the researcher asked ten experienced first grade teachers to peruse the original list and select the ten items that they felt were most important for a child to possess. They were also encouraged to add any items that they thought were important. A copy of this list is in Appendix D.3.

Step 3: The results were tabulated. Similar items were combined and skills that were chosen by at least three teachers were included. New items that were suggested by any of these ten teachers were also included. Similar items were combined. Refer to Appendix D.4 for a copy of these results.

Step 4: Several established checklists were perused for wording and style. These checklists were used as models for wording the items. The Learning Descriptors in the Draft Document in the Primary Program were perused. Items from the Teacher Checklist were compared with the Learning Descriptors and were placed under the headings which were already established. The headings used in the checklist were Emotional Development,

Social Development, Physical Development, Social Responsibility and Intellectual Development. Items 11, 14 and 20 were taken from the Learning Descriptors and included in the Checklist. Refer to Appendix D for the Teacher Checklist that was used in the study.

Appendix D.2

Step 1: Teacher Input for Checklist

Bridging the Gap

Your point of view

Teachers' Group

Grade One

In order to be successful in a grade one program children entering grade one should possess the following qualities:

-perhaps a 2nd year of kindergarten for those who are not emotionally mature
-early printing skills - hold a pencil, left to right progression, print name properly, circles and lines
-knowledge of birthdate, address, telephone number, full name
-recognizing letters (upper and lower)
-recite alphabet
-book knowledge - what is a story?, what is a word?
-listening skills - eye contact, hands still, no talking when someone else is
-work independently - quietly
-recognize #1-10, count to 30, one-to-one correspondence
-able to function in a large group - cooperate
-attention span
-ability to stay on task
-knows how to use tools, scissors, pencils, glue
-heavy exposure to literature
-some training in listening skills - following directions
-looking after personal belongings and school property
-letter and numerical recognition
-lots of oral reading (pocket chart etc.) chants, singing
-language development (brainstorming, etc., farm animals....
-training in school bathroom routine
-training in changing for gym, clothes for outside
-lots of exposure to literature and drama and an enthusiasm for books and reading
-thinking skills - problem solving - not necessarily math
-a good positive attitude towards school and self worth
-an expanded speaking vocabulary
-feeling good about themselves so they are not afraid to take risks in answering questions, writing, predicting, science activities
-free exploration with a variety of materials - science and math
-some awareness of names of alphabet and order and number
-cut and glue
-listening skills
-ability to stay on task independently (less direct help from parent aides)

Appendix D.2

-book knowledge and handling skills
-left to right, top to bottom, directions
-print has meaning
-social skills - cooperate in group
-some experience in attending to large group direction/work
-decision making skills
-experience with writing
-willingness to take risk

Appendix D.3

Step 2: More Teacher Input

Dear _____,

For my thesis to complete my M.A. degree in School Psychology, I am planning to do research using early primary children. I am hoping to use two recently published tests and compare them with how the classroom teacher evaluates the child. I am specifically interested in motor-output and how it effects learning and progress in school. My plan is to develop a teacher checklist. Enclosed you will find a checklist that was generated by local early primary teachers a few years ago. Would you please help me by choosing the **ten** items that you think are the most important skills that the child needs in order for him to be successful in second year primary in school.

In order to successful in a grade one (Second year primary) program, children entering grade one should possess the following skills:

Please check the ten items which you feel are the most important qualities to possess:

_____ **early printing skills - hold a pencil, left to right progression, print name properly, circles and lines**

_____ **knowledge of birth date, address, telephone number, full name**

_____ **recognize letters (upper and lower)**

_____ **recite alphabet**

_____ **book knowledge - what is a story? , what is a word?**

_____ **listening skills - eye contact, hands still, no talking when someone else is**

_____ **work independently - quietly**

_____ **recognize #1-10, count to 30, one-to-one correspondence**

_____ **able to function in a large group - cooperate**

_____ **attention span, ability to stay on task**

_____ **knows how to use tools - scissors, pencils, glue**

_____ **heavy exposure to literature**

_____ **some training in listening skills - following directions**

Appendix D.3

____ looking after personal belongings and school property

____ lots of oral reading (pocket chart etc.) chants, singing

____ language development (Brainstorming, etc., farm animals)

____ training in school bathroom routine

____ training in changing - for gym, clothes for outside

____ lots of exposure to literature and drama and an enthusiasm for books and reading

____ thinking skills, problem solving - not necessarily math

____ a good positive attitude towards school and self worth

____ an expanded speaking vocabulary

____ feeling good about themselves so they are not afraid to take risks in answering questions, writing, predicting, science activities

____ free exploration with a variety of materials - science and math

____ book knowledge and handling skills

____ left to right, top to bottom, directions

____ print has meaning

____ social skills, co-operate in a group

____ some experience in attending to large group direction/work

____ decision making skills

____ experience with writing

____ willing to take risks

Please add any other skills which you think are important for a child entering second year primary (Grade one) to have. _____

_____

_____

Appendix D.4

Step 3: Tabulated Checklist Results

* Priority - chosen by 10 teachers January 25, 1992

This is a revised version of the Teacher Checklist (like items are combined).

Int. * ✓ early printing skills - hold a pencil, left to right progression, print name properly, circles and lines

____ knowledge of birth date, address, telephone number, full name

✓ recognize letters (upper and lower)

✓ recite alphabet

✓ work independently - quietly

* ✓ recognize #1-10, count to 30, one-to-one correspondence

* ✓ attention span, ability to stay on task

* ✓ knows how to use tools - scissors, pencils, glue

Lang * ✓

____ some training in listening skills - following directions

____ listening skills - eye contact, hands still, no talking when someone else is speaking

____ some experience in attending to large group direction/work

* ✓ looking after personal belongings and school property

Com. ____ lots of oral reading (pocket chart etc.) chants, singing

Com. * ✓ language development (Brainstorming, etc., farm animals)

____ training in school bathroom routine

____ training in changing - for gym, clothes for outside

Com. * ✓

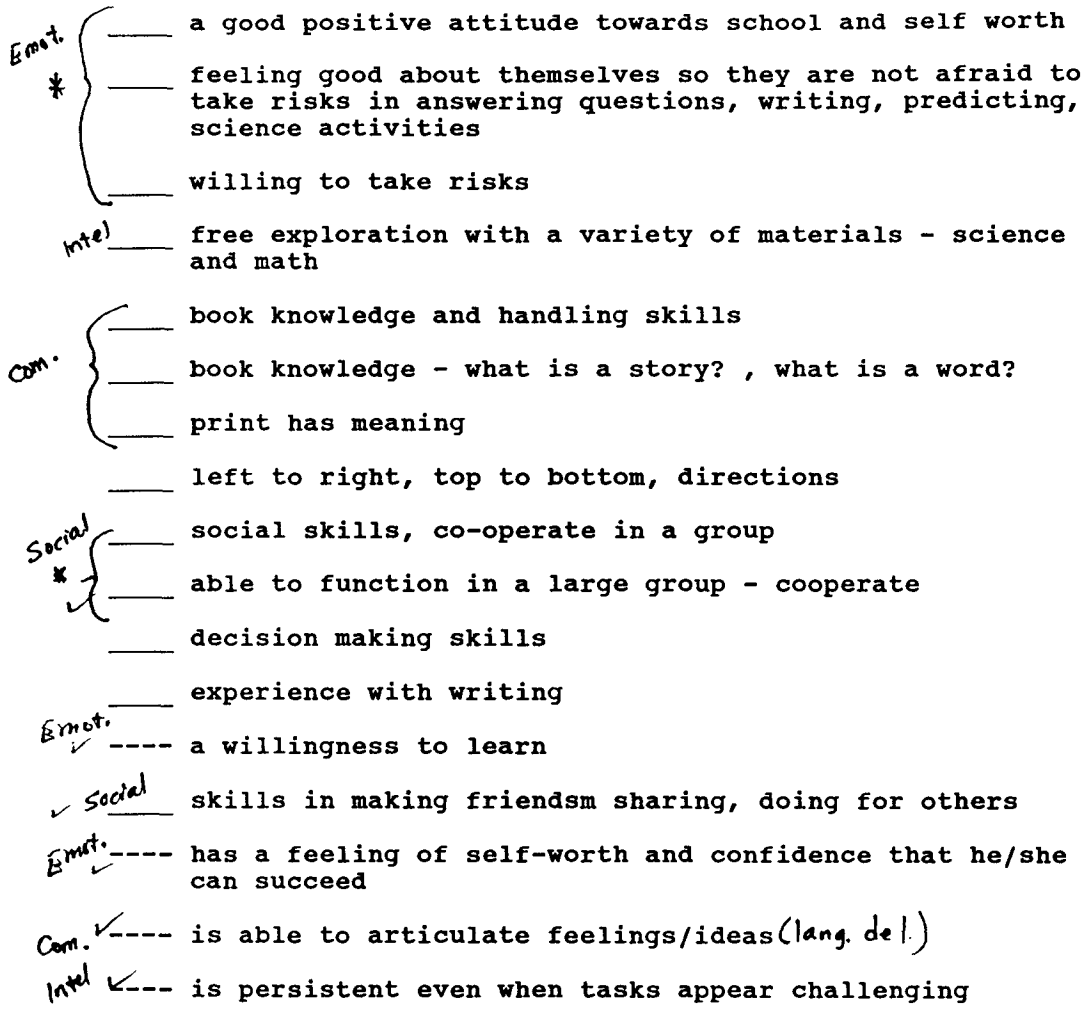
____ lots of exposure to literature and drama and an enthusiasm for books and reading

____ heavy exposure to literature

Arith * ✓ thinking skills, problem solving - not necessarily math

Com. ✓ an expanded speaking vocabulary

Appendix D.4


Emot.
a good positive attitude towards school and self worth
feeling good about themselves so they are not afraid to take risks in answering questions, writing, predicting, science activities
willing to take risks
Intel
free exploration with a variety of materials - science and math
Com.
book knowledge and handling skills
book knowledge - what is a story? , what is a word?
print has meaning
left to right, top to bottom, directions
Social
social skills, co-operate in a group
able to function in a large group - cooperate
decision making skills
experience with writing
Emot.
a willingness to learn
Social
skills in making friendsm sharing, doing for others
Emot.
has a feeling of self-worth and confidence that he/she can succeed
Com.
is able to articulate feelings/ideas (lang. del.)
Intel
is persistent even when tasks appear challenging

Appendix E

Letter to Kindergarten Teacher

Dear _____,

Thank-you for participating in this study. The purpose of my research is to validate an experimental checklist of skills for end of Year 1 Primary students and to look at the relationships between that measure and other measures of student competence and performance. Research emphasizes the importance of early identification and intervention. The Title of my thesis is, "Identifying At-Risk Primary Students: Global, Academic and Specific Skills Assessments". I am particularly interested in children between the ages of 5 years 3 months and 6 years 9 months.

I plan to do my research in three stages. Following is a description of the stages:

- *In stage one I will ask you to rank all your students as average, low (at risk), or high.*
- *In stage two I will ask you to fill out a 9 item Academic Competence Checklist, (which has been standardized) on each of your students.*
- *In stage three I will ask you to fill out a Teacher-made checklist on each of your students. This checklist was generated by Grade 1 teachers in this district.*
- ** *Your total time commitment for the above 3 stages would be approximately 3 to 5 hours.*

Mr. Jim Dyck has approved this research.
This study depends on the support of many people.
If you have further questions, please call me at home at 853-8989 or at Wm. A. Fraser Elementary School at 859-6794.
My goal is to complete the three stages before the end of June.

Thank-you for your interest and co-operation.

Yours truly,

Roz (Rosalyn) Francis

Appendix F

Letter to Principal

_____, *Principal*

_____ *Elementary School*

Dear _____,

This is a request for you to grant me permission to approach your kindergarten teacher to solicit her interest in participating in this research study. The study would involve the teacher only.

As a student completing research for a Master of Arts degree in the Department of Educational Psychology at the University of British Columbia, I am interested in the early identification of At Risk Primary Students. The purpose of my research is to validate a checklist that was generated by Grade 1 (Year 2) teachers in this district and to look at the relationships between that measure and other measures of student competence and performance. The title of my thesis is, "Identifying At-Risk Primary Students: Global, Academic and Specific Skills Assessments". I am specifically interested in children aged between 5 years, 3 months and 6 years, 9 months. I am hoping that the three phases will be completed before the end of June. This research will be done under the supervision of Dr. William T. McKee from U.B.C.

My research study has been approved by Mr. Jim Dyck, Superintendent of Schools and has been endorsed by Paul Brinton, Director of Special Programs. A brief description of the study is presented below.

There will be three phases in the study. In Phase I, Teachers will be asked to rank their students as low (at risk), average or high. In Phase II, the teacher will be asked to fill out a 9 item general Academic Competence Checklist that is standardized. In Phase III, the Teacher will be asked to complete the Experimental Teacher-made Checklist on each of her students. Teacher participation will require approximately 3 to 5 hours.
Research supports early identification and intervention.

This study depends on the support of many people. Please do not hesitate to contact me with any questions you may have. I would be happy to discuss the project with you.

Thank-you.

Rosalyn (Roz) F. Francis
853-8989 Home 859-6794 Wm. A Fraser Elementary

Appendix G

Teacher Consent Form

To: Roz Francis

I have read the letter to Kindergarten teachers containing information about the research study, Teacher Judgment in Identifying At Risk Primary Children.

I consent/do not consent to participate in in this research study.

Signature of Teacher _____

Date _____

(Please return this form to Roz Francis.)

Appendix H

Definition of Terms

At-risk - a child is "at-risk" when he has a greater than average chance of developing a disability. Risk is not a condition but rather a circumstance to indicate an increased probability that a disorder will occur.

Screening test - a test used to identify children who may be in need of special services which is used as a first step in identifying children needing further diagnosis. This test focuses on the child's ability to acquire skills.

Readiness test - assessment of child's level of preparedness for a specific academic or pre-academic program.

Early identification - refers to the practice of screening young students in an attempt to discover those likely to be "at-risk" for experiencing school problems at a later time.

Consistency - refers to the stability of teacher's scoring when rating students using different measures.

Appendix I

Teacher Checklist - Hierarchical Cluster Analysis

In order to discover to what extent the Teacher Checklist identifies children as at-risk for learning problems, Hierarchical Clustering techniques were used that were identical to the Academic Competence Questionnaire as well as the Checklist Subscale Clustering Analysis. Means and standard deviations were provided for each of the 28 items on the Checklist. In addition, means and standard deviations were provided for each of the four clusters on each of the 28 items on the Checklist. A plot of the means and standard deviations of the individual items for each cluster aided in the interpretation of the clusters.

A 4 x 3 matrix was used to compare the students' placement in the four clusters with the teachers' initial (EI) classifications in Phase 1. A Chi-square Test was used to test whether the two rating methods were related.

The cell means and standard deviations for each of the EI classifications based on the Teacher Checklist are presented in Table 29. Means for the three groups are different. Group 1 and 2 means are closer and actually overlap when applying a band of +/- one standard deviation. There is more than one standard deviation separation between the Group 2 and 3 means. As indicated in Table 29, the mean difference between Groups 1 and 2 is 27.05 and 31.81 between Groups 2 and 3.

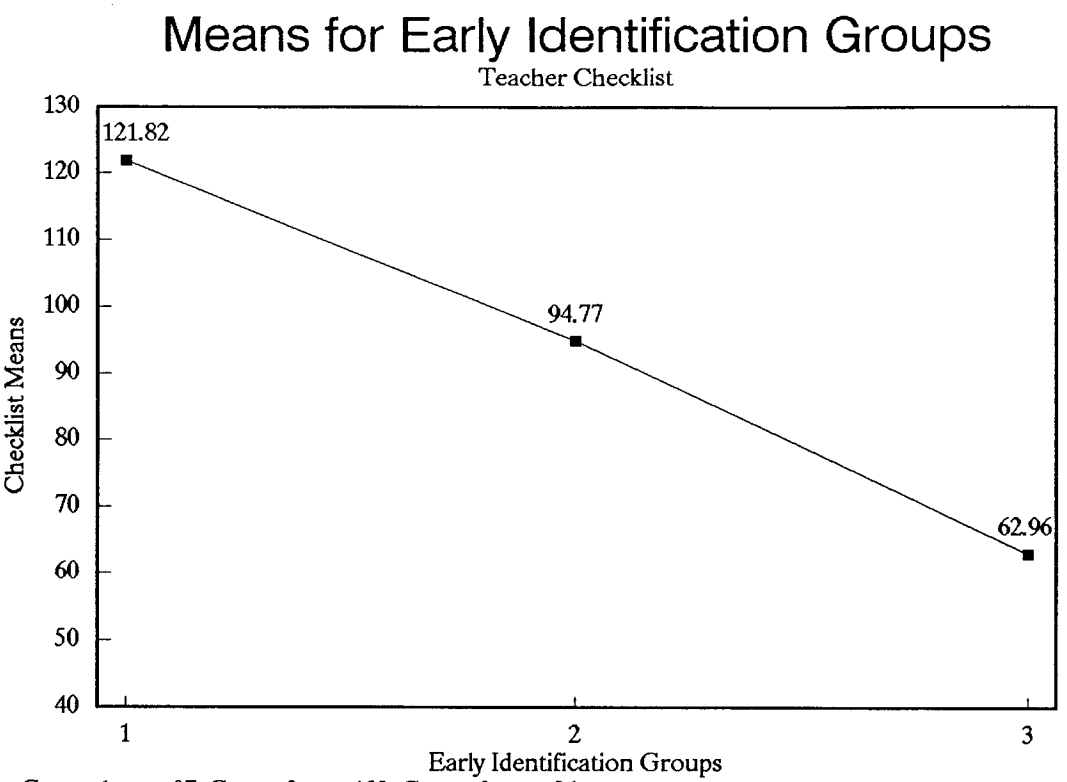
TABLE 29

CHECKLIST MEANS AND STANDARD DEVIATIONS FOR EI GROUPS

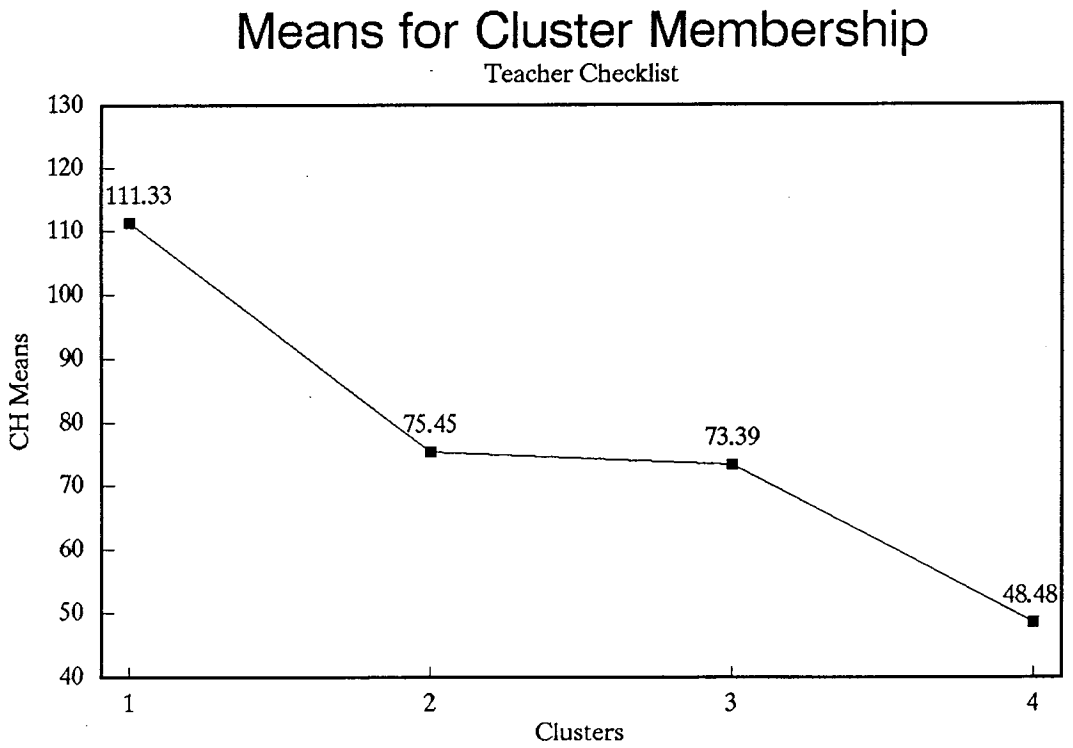
EI SCORE	MEAN	SD	NUMBER
1	121.82	11.89	97
2	94.77	15.72	183
3	62.96	14.07	56

Table 30 presents the total score means of the clustering variables for each of the four clusters. The mean of Cluster 1, or the high group, was 111.33 (SD 14.30). Two hundred and twenty-two students were in Cluster 1. The mean of Cluster 2, containing 62 students, was 75.45 (SD 9.11), and the mean of Cluster 3, with a total of 31 students, was 73.39 (SD 10.57). The mean of Cluster 4, with 21 students, was 48.48 (SD 6.72). Clusters 2 and 3 are similar in most areas. Cluster 2 was lower than Cluster 3 in working independently, being persistent and co-operating in large groups. Conversely, Cluster 3 was lower than Cluster 2 in the reading readiness skills such as: enthusiasm about books, realising that print has meaning, oral language skills, telling stories and reading familiar pattern books. Both Clusters 2 and 3 were similarly low in reading familiar pattern books, letter-sound relationships and emergent writing skills. Figure seven displays graphic representation of the Checklist means for the three EI Groups and the Checklist means for the four Clusters.

FIGURE 7



Group 1: n = 97, Group 2: n =183, Group 3: n = 56



Cluster 1: n = 222, Cluster 2: n = 62
Cluster 3: n = 31, Cluster 4: n =21

TABLE 30

CHECKLIST MEANS AND STANDARD DEVIATIONS FOR THE CLUSTERS

Cluster Membership	Mean	SD	Number in Cluster
1	111.33	14.30	222
2	75.45	9.11	62
3	73.39	10.57	31
4	48.48	6.72	21

In order to facilitate interpretation of the four clusters, the cluster profile points for each group of students were plotted relative to the Teacher Checklist mean and standard deviation for each of the 28 items for the total sample of children. Figure eight illustrates the cluster profile points for the different groups of students.

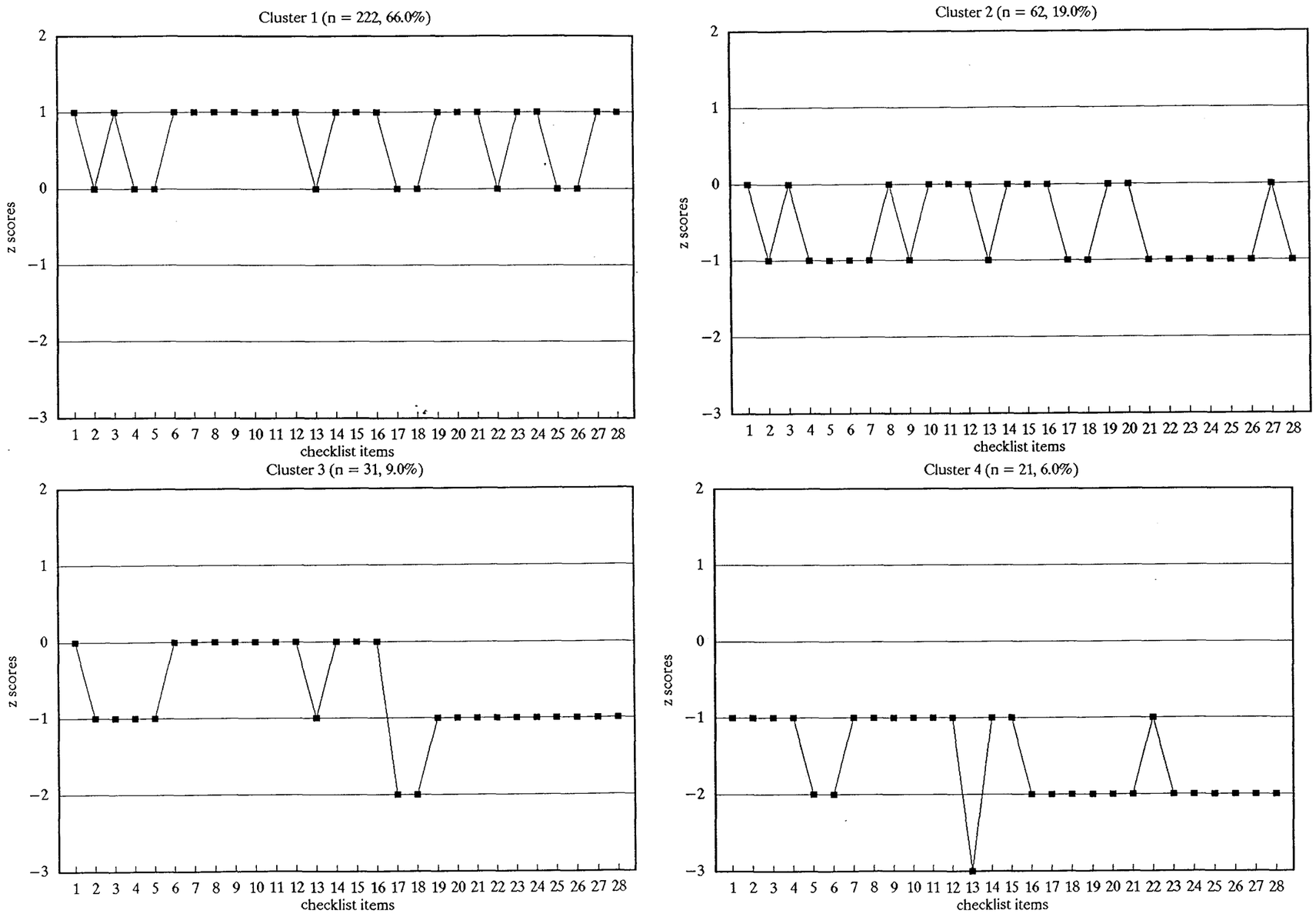
Cluster 1 with 66% of the sample, was consistently one z-score higher than Cluster 2 in all items.

Cluster 2 containing 19% of the sample was similar to Cluster 3 in some areas. But Cluster 3 was higher than Cluster 2 in working independently, persistence, and co-operating in large groups. Meanwhile Cluster 2 was higher than Cluster 3 in enthusiasm about books, knowledge about print, oral language and problem solving skills. Other items on the Checklist indicated similar patterns between Clusters 2 and 3.

Cluster 3 is smaller than Cluster 2 with only 9% of the sample. It is similar to Cluster 4 in positive attitude, taking risks, willingness to learn, enthusiasm about books, knowledge about print and recognizing name in

FIGURE 8

Cluster Membership Checklist



print. Possibly some students from Cluster 3 may be later identified as at-risk.

Six percent of the sample is in Cluster 4. The members of this cluster are low in all areas and stand out as very low in pencil skills. Individuals in this cluster would likely be at-risk. The following section displaying crosstabulation between CH and EI groups will assist in clarifying who may possibly be at-risk.

A comparison of the 3 group classifications on Phase 1 (EI) with the 4 cluster groupings on Phase 3 in the Checklist is given in Table 31. As indicated in Table 31, 96 students were classified as Group 1 or high on both the Phase 1 teacher classification and Phase 3 cluster grouping of the Checklist. Forty-seven students were placed in Group 2 on both the Teacher Classification and the cluster grouping of the Checklist, however 126 students who were classified as average or Group 2 on the initial classification were in Cluster 1 based on the cluster ratings of the Checklist. Eight students who were originally classified in EI 2 were in Cluster 3 and 2 students from EI 2 were in Cluster 4. Twenty-two students were in Group 3 (Cluster 3) on both classifications but 19 of the Group 3 students from Phase 1 classifications were in Cluster 4, and 15 of the Group 3 students were placed in Cluster 2 on the cluster analysis of the checklist.

It can be speculated that the 22 students from CH 3 - EI 3, the 19 students from CH 4 - EI 3 are clearly at-risk. Approximately 12% of the sample would fall into this group. Further consideration should be given to the 2

students from CH 4 -EI 2, the 8 students from CH 3 - EI 2 and 15 students from CH 2 - EI 3. These students, another 7% of the sample may be at-risk. Approximately 19% of the entire sample falls into the category of being clearly or likely at-risk and would likely benefit from further assessment.

Upon perusal of the scores, it was interesting to note that 27 students who were classified as average on Phase 1 were rated more than one standard deviation below the mean of the average group on the Teacher Checklist. Another point of interest was that 10 students who were initially classified as low in Phase 1 were rated more than one standard deviation above the mean of the Cluster 3 scores, as well, they fit within minus one standard deviation below the mean of Group 2.

TABLE 31

CROSSTABULATION OF CHECKLIST MEMBERSHIP BY EARLY IDENTIFICATION SCORES

		EI SCORES			
CH MEM		1	2	3	Row Total
	1	96	126		222
	2		47	15	62
	3	1	8	22	31
	4		2	19	21
	Column Total	97	183	56	336

A Chi-square Test was used to determine if there was a significant difference between the teachers' initial classification of students into three

groups and the four group classifications that were derived from the Cluster Groupings obtained when students with like characteristics were clustered using the Hierarchical Cluster Analysis technique. Chi-square test results (227.37, D.F.(6), $\underline{p}=0.001$) demonstrated a significant difference in group separation. Cramer's V statistic of 0.58 indicates a moderate relationship between EI Classification and Checklist Cluster membership.