

Intelligence Tests:
A Perspective from Psychologists

by

Donna L. Johnson

A thesis
Submitted to the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of

Master of Education

Department of Educational Psychology
University of Manitoba
Winnipeg, Manitoba

(c) May, 1996



National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-612-13219-6

THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION

INTELLIGENCE TESTS: A PERSPECTIVE FROM PSYCHOLOGISTS

BY

DONNA L. JOHNSON

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University
of Manitoba in partial fulfillment of the requirements for the degree
of**

MASTER OF EDUCATION

(c) 1996

**Permission has been granted to the Library of The University of Manitoba to lend or sell
copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis
and to lend or sell copies of the film, and to University Microfilms Inc. to publish an
abstract of this thesis/practicum.**

**This reproduction or copy of this thesis has been made available by authority of the
copyright owner solely for the purpose of private study and research, and may only be
reproduced and copied as permitted by copyright laws or with express written
authorization from the copyright owner.**

Abstract

The purposes of this study were to describe the perceptions of Manitoba school psychologists as to (a) what intelligence tests were used in Manitoba with children (b) for what purposes do psychologists in Manitoba use intelligence testing (c) what they perceived to be the strengths and (d) the limitations of intelligence testing. All members of the Manitoba Association of School Psychologists (MASP) and the Child Guidance Clinic (CGC) were surveyed by mailed questionnaire as well as psychologists who were not members of these two organizations. Demographic information included years as a school psychologist, type of practice, degrees held and whether they used intelligence tests with children. Results identified (a) the intelligence tests used in Manitoba with children (b) the purposes for which psychologists in Manitoba use intelligence testing (c) the perceived strengths or positive aspects of intelligence testing and (d) the perceived limitations or negative aspects of intelligence testing. As well there is evidence of:

1. continued use of intelligence tests; 2. alternative assessment procedures; and 3. caution in using the tests. Purposes for using intelligence tests related to the identification of learning potential, placement in programs and assessment of student strengths and weaknesses. Psychologists focused on using tests as a tool to understand and help students, parents and educators.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Gerry Bravi for his time, patience and advise in assisting me in the completion of this project. As well, I would like to acknowledge my committee members, Dr. Paul Madak and Dr. Ken McCluskey for their work and assistance.

I would also like to thank the school psychologists from the Manitoba Association of School Psychologists and the Child Guidance Clinic. They have made it possible to understand a complex task and I am grateful for their time, trust and support.

TABLE OF CONTENTS

Abstract.....	-i-
Acknowledgments.....	-ii-
Table of Contents.....	-iii-
Tables.....	-viii-
CHAPTERS	
1. Introduction.....	1
Statement of the Problem.....	1
Purpose of the Study.....	4
Educational Significance.....	6
2. Review of the Literature.....	13
Purposes of Testing.....	15
Group Tests.....	16
Purposes of Group Tests.....	16
Individual Tests.....	17
Purposes of Individual Tests.....	17
Strengths of Intelligence Testing.....	19
Concept of Intelligence.....	19
Culture Fairness Tests.....	20
Group Tests.....	22
Advantages.....	22
Technical Adequacy.....	23
Individual Tests.....	24
Advantages.....	24
Technical Adequacy.....	26
Summary.....	27
Limitations of Intelligence Testing.....	28
Concept of Intelligence.....	28
Culture Fairness.....	31
Group Tests.....	34
Disadvantages.....	34
Technical Adequacy.....	34

Individual Tests.....	35	
Disadvantages.....	35	
Technical Adequacy.....	38	
Reliability.....	38	
Validity.....	41	
Summary of Review of the Literature.....	41	
Purposes of Testing.....	41	
Strengths of Intelligence Testing.....	42	
Limitations of Intelligence Testing.....	42	
3. Method.....	45	
Population.....	45	
Instrument.....	46	
Pretesting.....	46	
Procedure.....	47	
Analysis.....	48	
4. Results.....	49	
Part 1 Results.....	50	
Demographic Information.....	50	
Part 2 Results.....	55	
Question #1	Do you use <u>group</u> intelligence tests with children?.....	55
Question #1a	Which <u>group</u> intelligence tests do you use?.....	55
Question #1b	For what purposes do you use <u>group</u> intelligence tests with children?.....	56
Question #2	Do you use <u>individual</u> intelligence tests with children?.....	56
Question #2a	Which <u>individual</u> intelligence tests do you use with children?.....	57

Question #2b	For what purposes do you use <u>individual</u> intelligence tests with children?.....	59
Question #3	If you do not use intelligence tests with children, please explain why.....	63
Question #4	If you do not use intelligence tests with children, what other assessment procedures do you use?.....	64
Question #4a	Why do you use other assessment procedures?.....	69
Question #5	What do you perceive to be the strengths or positive aspects of the intelligence tests you use with children?....	71
Question #6	What do you perceive to be the limitations or negative aspects of the intelligence tests you use with children?...	76
Question #7	Do you have any other comments you would like to make about intelligence testing with children?.....	86
5. Discussion.....		90
Conclusion.....		104
Limitations.....		106
Recommendations.....		107
References.....		108

Appendices.....	118
Appendix A - Permission Letter Mailed to the Manitoba Association of School Psychologists.....	118
Appendix B - Permission Letter Mailed to the Child Guidance Clinic.....	119
Appendix C - Letter Mailed to the Manitoba Association of School Psychologists Survey Respondents.....	120
Appendix D - Letter Mailed to the Child Guidance Clinic Survey Respondents.....	121
Appendix E - Survey with Manitoba Psychologists.....	122
Appendix F - Reminder Letter to All Psychologists.....	130
Appendix G - Results Request Form.....	131
Appendix H - Comments from Psychologists for Question # 2b.....	132
Appendix I - Comments from Psychologists for Question #3.....	141
Appendix J - Comments from Psychologists for Question #4.....	143
Appendix K - Comments from Psychologists for Question #4a.....	148
Appendix L - Comments from Psychologists for Question #5.....	152

Appendix M - Comments from Psychologists for Question #6.....	170
Appendix N - Comments from Psychologists for Question #7.....	186

Tables

Table 1.	Number of Years as a Practicing Psychologist.....	50
Table 2.	Type of Psychological Practice You Have.....	51
Table 3.	Percentage of Work Day Spent on Testing.....	52
Table 4.	What Degrees Do You Hold?.....	53
Table 5.	Do You Hold a Teacher's Certificate?.....	54
Table 6.	Do You Use Intelligence Tests with Children?....	54
Table 7.	Individual Tests Used.....	57
Table 8.	Purposes for Using Individual Intelligence Tests.....	60
Table 9.	Why Psychologists Don't Use Intelligence Tests....	63
Table 10.	Tests Used.....	66
Table 11.	Other Assessment Procedures Used.....	68
Table 12.	Reasons for Other Assessment Procedures.....	70
Table 13.	Strengths of Intelligence Tests (Group Tests).....	72
Table 14.	Strengths of Intelligence Tests (Individual Tests)...	73
Table 15.	Limitations of Intelligence Tests (Group Tests).....	76
Table 16.	Limitations of Intelligence Tests (Individual Tests)..	78
Table 17.	Other Comments About Intelligence Tests.....	86

CHAPTER I

Introduction

Statement of the Problem

The problem of intelligence testing with children today centers around three issues: (a) the purposes (b) the strengths and (c) the limitations of intelligence testing. First, purposes of intelligence testing rest on the expectations regarding the information that the testing will provide. Here lies the rub. The information that intelligence testing provides is dependent upon the strengths and limitations of the tests themselves. Second, if the expectations are that intelligence testing will measure intelligence, one must remember that "no IQ test is known to measure *the* best construct of intelligence, [thus] different IQ tests in effect measure different "intelligences" "(Carroll & Horn, 1981). Third, some intelligences such as musical and bodily-kinesthetic intelligences are not part of these formalized tests (Gardner, 1993). These are three major considerations for educators to keep in mind when psychologists use intelligence testing with students aged 5 - 18 years.

Information as to the strengths of intelligence testing encompass technical and social features. From a technical standpoint, patterns of high and low scores have been associated with specific disabilities (Bannatyne, 1974; Hale & Saxe, 1983; Kaufman, 1979) neuropsychological deficits (Rourke, Bakker, Fisk & Strang,

1983) and emotional disorders (Dean, 1977). Further, intelligence tests have been standardized so they can be uniformly applied.

From a social standpoint, three aspects could be described as adding strength to the use of intelligence tests. First, Reschly (1979) stated that test use protects students from inappropriate classification due to referrals. Second, variables such as ethnicity and socioeconomic status are important criteria of representativeness in standardization samples. Kaufman (1992) stated that tests such as the WPPSI-R and WISC-III encompass factors such as an appropriate standardization sample whereby it is "very representative on major variables, especially those factors like ethnicity and socioeconomic status and are known to correlate with one's obtained IQ's " (p. 155). This means that the norming sample within these tests are appropriate according to the above mentioned variables. Third, Robinson and Chamrad (1986) found that (a) group and individual intelligence tests can identify gifted children (b) the attitude of a child during testing is confirmation of performance in the classroom and (c) the testing situation may reveal other characteristics regarding the child's fluency, fine motor skills and mood. This means that in their research they found evidence that intelligence tests have the ability to (a) identify specific abilities of children, (b) that their abilities during testing is similar within the classroom, and that (c) characteristics about the child, such as verbal expression can be understood.

Criticisms of intelligence testing with children range from problems with the labeling of students for selection into programs (Myers, 1981; Ysseldyke, Algozzine & Thurlow, 1992), and bias of standardized tests related to specific types of students being identified (Ysseldyke, Algozzine and Thurlow, 1992) to tests developed for one purpose and their use being diverted for other purposes (Ysseldyke, Algozzine and Thurlow, 1992). With these revelations in mind, one must understand that intelligence testing will not result in 100% accuracy (Robinson & Chamrad, 1986).

A review of the literature concerning psychologists' use and views of intelligence testing has revealed studies by psychologists revealed studies by Lubin, Wallis and Paine (1971), Brown and McGuire (1976), Wade and Baker (1977), Ryan and Sackett (1987), Goh, Teslow and Fuller (1981) and Hutton, Dubes and Muir (1992). Lubin et al. (1971) obtained a frequency list of intelligence tests used by psychologists in state hospitals, counseling centers and Veterans Administration stations in the United States. The Brown and McGuire (1976) study concerned the frequency of usage of different intelligence tests by professionals in community mental health agencies and hospitals throughout the U.S. Wade and Baker (1977) surveyed clinical psychologists regarding the use and opinions of psychological tests. It was found that (a) personal clinical experience with a test was more important than psychometric considerations and (b) they emphasized the subjective, experiential nature of the

testing process. As well, a list of the tests used was obtained. Ryan and Sackett (1987) studied personnel-related individual assessment processes and content in a management arena. Again, a compilation of assessment procedures was gained. Hutton, Dubes and Muir (1992) conducted a study of the frequency of usage of instruments used by members of the National Association of School Psychologists. This study was a comparison of a study done 10 years earlier by Goh, Teslow and Fuller (1981). This study listed the percentage of psychologists who worked with preschool, elementary, junior high and senior high students, the number of different assessment devices administered, and the psychological instruments most frequently used in different assessment areas (e.g. intelligence and achievement).

I was unable to find any studies that examined psychologists perceptions of the strengths and limitations of intelligence testing or the purposes for which psychologists use intelligence testing with children in Manitoba. Studies of this type conducted in the United States included a wide range of age levels and psychologists practicing in a variety of locales. Since practices in Manitoba could be different, it is necessary to survey psychologists in this province.

Purpose of the Study

It is necessary therefore to view intelligence testing of children through the eyes of a psychologist to understand (a) what intelligence tests are used in Manitoba with children (b) for what purposes do

psychologists in Manitoba use intelligence testing (c) what they perceive to be the strengths or positive aspects and (d) what they perceive to be the limitations or negative aspects of intelligence testing.

It is necessary to understand their views for two reasons. First, for the past number of years, researchers have written about the controversies surrounding the continued use of intelligence testing. Second, while much has been written in the literature regarding the use of intelligence tests, little is known about how practicing psychologists view the use of intelligence testing. The researcher was unable to find studies in the literature which pursued the idea of what psychologists thought regarding the purposes, strengths and limitations of intelligence testing. After reviewing this literature, the researcher finds it necessary to pursue the following questions:

1. What individual and group intelligence tests do Manitoba school psychologists use?
 - 1a. If they do not use intelligence tests, why have they chosen not to?
 - 1b. What other assessment procedures do they use instead of intelligence tests?
2. For what purpose(s) do school psychologists in Manitoba use intelligence testing?
3. What do Manitoba school psychologists perceive as the strengths or positive aspects of intelligence testing?

4. What do Manitoba school psychologists perceive as the limitations or negative aspects of intelligence testing?

Educational Significance

Research concerning the views of school psychologists regarding the strengths, limitations and purposes for the use of intelligence testing with children could not be found in the literature. The relevance and timeliness of this investigation can be evaluated according to the need for educators to understand the purposes for which school psychologists in Manitoba utilize intelligence testing plus their views about the strengths and limitations of such testing.

This is important as testing is used within an educational context for a variety of reasons, for example, for evaluation of individual progress and eligibility for special programming. As well, it may be used as a diagnostic tool to detect learning difficulties as well as to achievement discrepancies such as reasons for poor school performance. Testing may also provide data regarding program development. Most importantly, educators need to understand students' strengths and weaknesses. The limitations of intelligence testing include technical adequacy of the instruments, culture fairness of tests, varying concepts of intelligence (definitions of what intelligence is), and influences, both external and internal, that affect performance on tests.

Further, the decisions made regarding students fall into five areas; "evaluation of the individual, program evaluation, screening, placement and intervention planning" (Ysseldyke, Algozzine & Thurlow, 1992, p. 170). The information gained has the potential to assist in appropriate decision making. Decisions made lead one to make plans regarding an individual. The tests used from which to gain information need to be chosen with care as each has its strengths and weaknesses. One purpose of individually administered intelligence tests is "for making educational placement decisions" (Salvia & Ysseldyke, 1988, p. 161), whereas group intelligence tests are given "as screening devices to identify those who are different enough from average to warrant further assessment . . . [and] to provide descriptive information about the level of capability of students in a classroom" (Salvia & Ysseldyke, 1988, p. 200). It is therefore imperative that the reason for testing match the type of test given since different kinds of tests, be they individual or group, provide different information. One must also evaluate the precision of each test given as each individual or group test has its own strengths and limitations. A case in point refers to the reliability scores reported for the tests to be used. "If test scores are to be used for administrative purposes and are reported for groups, a reliability of .60 should probably be the minimum" (Salvia & Ysseldyke, 1988, p. 128). Individual data have higher standards so that for tracking

and placement in a special class "the minimum standard should be .90" (Salvia & Ysseldyke, 1988, p. 129).

In addition, one must examine the use of personnel and time for testing of students. Group tests enable many students to be tested at one time, whereas individual tests limit the testing situation to one student at a time. However, one must remember the results desired from the testing situation. Individual tests will provide a one - to - one situation which is conducive to observing how a child approaches a problem. The gains achieved by testing must be weighed carefully so that testing is not overused or misused.

Ysseldyke, Algozzine, Richey and Graden (1982) found that there was little correspondence between decisions made by placement teams and the data to support the decisions. It was found that placement teams who made decisions concerning a student's eligibility based their decisions on the amount of information presented regarding a student. For example, if a lot of information was given about a student, the likelihood was higher that the student would be identified as learning disabled (LD) and labeled. As well, "the data provide little evidence to suggest that teams use specific, formal criteria when making eligibility decisions" (Ysseldyke et al., 1982, p. 37). Thus one must be very conscious of the assessment procedures used with students and the reasons they are employed.

If we are to use intelligence testing within an educational setting we must keep in mind limitations concerning a lack of

personal impressions of the student during a group testing situation. If one needs to understand how a child tackles a question one needs to observe that child. Further, limitations concerning language and culture bias with individual tests can affect the results. Students who are tested in a language other than their mother tongue will be at a disadvantage. As well, those who are not familiar with the objects used in the test, or with questions which are culturally loaded will score lower. Knowing this will make us aware of discrepancies in scoring and aid the interpretation of the results. If these dilemmas play a role in the outcome of the testing situation, the end result may be an inaccurate measure of a child's capabilities. Following this line of thought, if major decisions are to be made concerning a child's education, one must be aware of the limitations and strengths of testing.

Much has been written regarding the ability of IQ tests to predict or assess academic performance. Wexler-Sherman, Gardner and Feldman (1988) wrote that "because these tests are intended to predict future performance in school, questions and tasks focus on those skills considered to guarantee success in standard school settings" (p. 77). If this is so, the tests are already structurally biased, and abilities which fall beyond mathematical or linguistic capabilities are seldom examined. Further, "evidence of the long-term predictive validity of individual intelligence tests scores for identifying the academically gifted comes largely from Terman's

longitudinal study of children with IQ scores of 140 or higher on the Stanford-Binet" (Fox, 1981, p. 1105). Evidence of the predictability of school success from IQ tests is described by Fox (1981) as well. Here, she mentions two important details. First, "factors not assessed by intelligence tests are also necessary to enable individuals to become highly productive [and second,] measures of global intelligence do not provide sufficient information about patterns of specific abilities and levels of achievement for program placement" (p. 1105-1106). The former item finds fault with the tests themselves by saying that IQ tests miss out on vital information such as the ability to solve real-world problems that enable a student to be successful in school. The latter item refers to the scores achieved on an IQ test. One student may show strengths in one area and weaknesses in another area while another student achieves differently, yet both students receive the same total test score. Sternberg (1984) also mentioned that IQ tests measured skills for largely academic kinds of environments. Observations by Gardner (1987) stated that "if you do well in language and logic, you will do well in IQ tests and SAT's . . . but whether you do well once you leave [school] is probably going to depend as much on the extent to which you possess and use the other intelligences" (p. 190). Gardner (1984) also stated that "intelligence tests are most effective at predicting how individuals will perform on other tests, especially tests created in the image of the original" (p. 700). Lastly, two critical articles by Wexler-Sherman,

Gardner and Feldman (1988) and Haney (1981) describe intelligence testing as sorting, ranking and classifying students. One may conclude then that the validity of predicting academic performance through the use of intelligence tests is suspect due to circumstances such as questions that are limited in scope, individual factors outside the testing realm such as motivation and educational opportunity which can determine success, and that IQ tests are set up to accomodate those who excell in linguistic and mathematical topics.

My research will attempt to show for what purposes educational psychologists in Manitoba utilize intelligence testing and the strengths and limitations they perceive in testing situations. It will also uncover which tests are used in Manitoba. This information will clarify to what extent intelligence testing is used and how educational psychologists in Manitoba perceive intelligence testing.

The importance of this study is three-fold. First, the results could be utilized to identify perceptions of these tests. Second, the results will help in developing compilations of the tests used by Manitoba educational psychologists. Third and adjacent to this point, is what do we do with this information? The outcome of this study should have its effects within the teaching community for it is fundamental to those who are educating future psychologists to enlighten them as to the opinions of practicing psychologists regarding the use, as well as the perceived strengths and limitations

of intelligence tests with children. Lubin, Wallis and Paine (1971) stated that "the attitudes of university faculty in clinical psychology are important influences on the testing practice of students" (p. 73).

Further, it is necessary for classroom teachers to understand the results because these might provide useful information pertaining to educational programming.

CHAPTER II

REVIEW OF THE LITERATURE

A review of the literature concerning psychologists' use and views of intelligence testing has revealed studies by Lubin, Wallis and Paine (1971), Brown and McGuire (1976), Wade and Baker (1977), Ryan and Sackett (1987), Goh, Teslow and Fuller (1981) and Hutton, Dubes and Muir (1992). Each study was conducted in the United States and described a compilation of psychological tests used. Lubin et al. (1971) collected data regarding a list of tests used in state hospitals, counseling centers and Veterans Administration stations. Brown and McGuire (1976) consolidated data on the overall frequency of usage of different psychological tests and cites the three most frequently used instruments in each of seven age groups for both personality and intellectual assessment, used in community mental health agencies and hospitals. Wade and Baker (1977) found that personal clinical experience with a test was more important than psychometric considerations. They noted that "The two most important criteria measures of test performance in research (ie., behavioral prediction and diagnostic accuracy) were not viewed by clinicians as highly important in test selection or use" (Wade & Baker, 1977, p. 879). Ryan and Sackett (1987) studied assessment practices in the management field while Hutton, Dubes and Muir (1992) compared the frequency of usage of instruments used by members of the National Association of School Psychologists ten

years after a similar study done by Goh, Teslow and Fuller (1981). In the Hutton study, a compilation of the percentage of psychologists who work with preschool, elementary, junior high and senior high students was listed as was a list of assessment instruments used. The Goh et al (1981) study identified the eight most frequently used instruments for eight assessment areas: Intelligence, Achievement - General, Achievement - Specific, Perceptual Functioning, Personality, Behavior Rating, Preschool Functioning and Vocational Interests. For the intelligence assessment area, the eight most used tests were 1. WISC-R, 2. Stanford-Binet, 3. Peabody Picture Vocabulary Test, 4. WAIS, 5. Goodenough Draw - a Man Test, 6. Adaptive Behavior Scales, 7. WPPSI and 8. The McCarthy Scales of Children's Abilities. The Hutton et al. (1992) study was an updated version of the Goh (1981) study. Again, there was a list of the tests most frequently used in the same eight categories, with Adaptive Behavior, a category added onto the original eight, for a total of nine. The top eight intelligence tests in this study were the same eight as in the previous study and in the same order. These studies omitted an important investigation regarding the views of psychologists concerning the strengths and limitations of intelligence testing and the purposes for which they used these instruments.

There are three important aspects to consider regarding intelligence testing with children: the purposes of using intelligence testing, the strengths, and limitations of intelligence testing. Each of

these aspects is important, yet considering the substance of each, one comes to the conclusion that the purposes of intelligence testing have, as their foundation, the information gained through testing. Thus the importance of the purposes of testing is tantamount to success with the instrument. The strengths and limitations of testing directly affect the information gained.

Purposes of Testing

Purposes of using intelligence tests for educational purposes with children revolve around three major objectives: (a) evaluation of individual progress (b) screening or (c) eligibility for special programming (Ysseldyke, Algozzine & Thurlow, 1992 p. 172).

Intelligence tests are also used to develop instructional programs (Ysseldyke, Algozzine & Thurlow, (1992). The important element within this area is the designation, educational purposes. The purposes of IQ testing for educational decision making is the most important aspect of this study as the school psychologists to be surveyed work with children who would be a part of an educational setting.

Ysseldyke (cf. Ysseldyke, Algozzine & Thurlow, 1992, p. 170) "identified five kinds of decisions regarding students--evaluation of the individual, program evaluation, screening, placement, and intervention planning". The choice of assessment tool, a group test or an individual test, is dependent on the purpose for which one is testing. As well, if one wishes to observe the individual, then an

individual tests would be more appropriate. Further, if one values the validity of one type of test over the other, then it is that kind of test that would be used. For this reason, a discussion of the purposes of individual and group tests will be carried out.

Intelligence tests are used most frequently "for making educational placement decisions" (Salvia & Ysseldyke, 1988, p. 161). "Intelligence tests are frequently employed as preliminary screening instruments, to be followed by tests of special aptitudes" (Anastasi, 1988, p. 238). These tests could be related to specific disciplines such as reading or math. As well, another common use "is to be found in clinical testing, especially in the identification and classification of the mentally retarded" (Anastasi, 1988, p. 238).

Group Tests

Group tests consist of those that employ the following characteristics: they are used with groups of students, and "consist of a single battery to be administered in one sitting, [while] others contain a number of subscales or subtests and are administered in two or more sittings" (Salvia & Ysseldyke, 1988, p. 200). A typical test employs the multiple-choice format for uniformity and objectivity of scoring (Anastasi, 1988).

Purposes of Group Tests. Group tests are generally used for two main purposes. "They are routinely administered as screening devices to identify those who are different enough from average to warrant further assessment Group intelligence tests are also used

to provide descriptive information about the level of capability of students in a classroom, district or even state" (Salvia & Ysseldyke, 1988, p. 200).

Individual Tests

Individual tests of intelligence differ from group tests in that they assess a student on an individual basis. Most importantly the examiner has the opportunity to observe the student and the way the student responds to the questions. This is important as influences such as anxiety and fatigue may influence the outcome of the testing.

Purposes of Individual Testing. According to the WISC-R manual, purposes of using the test encompass a description of assessment reasons ranging from (a) diagnosis of mental disease, (b) application in counseling and (c) application in education. In the WISC-III manual, the stated purposes are that it can "provide a considerable amount of information about an individual's cognitive strengths and weaknesses in a relatively short amount of time" (p. 2). It stated further that significantly low performance is a necessary part of a diagnosis of mental deficiency, that it is suitable for identifying children of high intellectual ability and can be an integral part of neuropsychological evaluations (Wechsler, 1992).

Some researchers have linked the patterns of high and low scores to the presence of specific developmental learning disabilities (Bannatyne, 1974; Hale & Saxe, 1983; Kaufman, 1979); neuropsychological deficits (Rourke, Bakker, Fisk & Strang, 1983);

and emotional disorders (Dean, 1977). The use of these tests, according to the researchers, can help to identify specific problems therefore, testing may be a good place to start. The use of patterning within the subtest scores has disadvantages though. First, "performance successes and failures can be attributed to a number of different possible underlying cognitive skills" (Kramer, 1993, p. 193). "The types of errors made and problem-solving strategies used often provide more clues about the status of an examinee's intellectual operations than the final achievement score" (Kramer, 1993, p. 193). This means that performance successes and failures may be attributed to a number of different possible underlying cognitive skills. Second, one must use caution when attempting to interpret the subtest scores as they don't have good reliability. For example, the reliability coefficients for the WISC-R range from 0.70 to 0.86 while the reliability coefficients for the WISC-III range from 0.69 to 0.87. "When tests are to be used to make decisions regarding individuals, they must have reliability coefficients greater than .90" (Ysseldyke, Algozzine & Thurlow, 1992, p. 190). This is also true for group tests. Therefore, some of the subtest scores must be treated with caution.

Thus when an individual test is employed, identification and classification are the primary outcomes. One may wish to observe the student during the testing situation as well as during classroom time to aid in understanding the manner in which the student

tackles the questions presented. Thus, it is important to collect both observations and scores regarding a student.

Strengths of Intelligence Testing

An awareness of the strengths of intelligence testing will help to explain the continued use and support of intelligence testing in educational settings. The strengths reflect many issues: the concept of intelligence, culture fairness tests, and the advantages of group and individual testing instruments. It is necessary to understand that proponents of IQ testing make the following claims. There are differences between what educators and psychologists claim are strengths of intelligence testing.

Concept of Intelligence

"A psychological test is essentially an objective and standardized measure of a sample of behavior" (Anastasi, 1988, p. 23). This idea is repeated by Salvia and Ysseldyke (1988) when they stated that "it is impossible to administer every item in the domain to a child whose intelligence we want to assess" (p. 148). Thus it would be impossible to test everything you wanted to know about a student. However, the value of a sample of behaviors gained through testing is beneficial as "the *diagnostic or predictive value* of a psychological test depends on the degree to which it serves as an indicator of a relatively broad and significant area of behavior" (Anastasi, 1988, p. 24). The behaviors and abilities analyzed during an intelligence test are those that predominate in school learning (Anastasi, 1988,

Wexler - Sherman, Gardner & Feldman, 1988). Proof of this may be found in the tests themselves where they ask for definitions of words, and investigate logical - mathematical inquiry. If that is so, "the score on a test of academic intelligence is also an effective predictor of performance in many occupations and other activities of daily life because the functions taught in the educational system are of basic importance in modern, technologically advanced cultures" (Anastasi, 1988, p. 363). One may have a general understanding of a student as a function of the performance on the test and daily performance within the classroom (Anastasi, 1988). Therefore, an intelligence test may be used as an assessment tool in the broadest sense with an emphasis on the correspondence between the student's knowledge and a picture of the total package of knowledge gained through observation.

Culture Fairness Tests

Tests such as the Culture Fair Intelligence Test by Cattell (1950) is a paper and pencil test developed because of "the need for a test which would fairly measure the intelligence of persons having different languages and cultures, or influenced by very different social status and education" (Cattell, 1962, p. 5). The tests developed by Cattell are used with children aged 4 to 8 years (Scale 1), 8 to 14 years (Scale 2) and 14 + (Scale 3). Thus the purpose for this test was to be more sensitive to the background cultures of students. "Group

tests can be administered simultaneously to as many persons as can be fitted comfortably into the available space" (Anastasi, 1988, p. 312).

Other culture-fair tests such as the Goodenough-Harris Drawing Test (1963) have been developed for this purpose.

Culture fair intelligence tests were developed due to criticism with traditional intelligence tests which "emphasize verbal and culture-related skills, thereby discriminating against children with language handicaps and from deprived or culturally different backgrounds" (McCluskey & Walker, 1986, p. 27). Cattell (1950) and Cattell and Cattell (1960, 1963) developed scales for use with children in three different age groups, 4-8 years old, 8-14 years, and 14 plus and believed that "culture-fair intelligence tests are more adequate measures of learning potential than are traditional intelligence tests" (Salvia & Ysseldyke, 1988, p. 201). Another culture fair test developed includes Raven's Progressive Matrices which has a set of designs arranged into rows or columns from which a part has been removed. From alternatives one must decide which portion is missing. There is also a Coloured Progressive Matrices for use with young children and special groups, and the Advanced Progressive Matrices for use with adolescents and adults. A third approach is the Goodenough Draw-a-Man Test (Harris, 1963). Here a child is asked to draw a man, and to do the best picture he/she can. The accuracy and amount of detail in each drawing are the criteria upon which the drawing is scored.

In addition, the Leiter International Performance Scale is another individually administered test developed within the culture fair arena. Here though, there is almost a complete elimination of instructions, verbal or otherwise. "The comprehension of the task is treated as part of the test" (Anastasi, 1988, p. 299).

Culture fair tests were developed to facilitate testing with people who were from cultures other than the culture from which traditional intelligence tests have been designed. Thus, they are an alternative for testing children who come from other cultural backgrounds and "assess intelligence with relatively little contamination by formal instruction" (Salvia & Ysseldyke, 1988, p. 204-205). However, there are criticisms of these tests as well.

Group Tests

This next section explains the advantages and technical adequacy of group administered instruments.

Advantages. A group test can be administered to many people at once. Using printed items and simple responses, the examiner's role in the testing situation is limited. An extensive amount of training is unnecessary for the examiner to give the test due to the uniform routine used. "Scoring is typically more objective in group testing . . . [or] can also be scored by computers" (Anastasi, 1988, p. 312). The norming samples are much larger for group tests as opposed to individual tests due to the ease of gathering data. "It is customary to test very large, representative samples in the

standardization process. . . . it is not unusual for the normative samples to number between 100,000 and 200,000, in contrast to the 1,000 to 8,000 cases laboriously accumulated in standardizing the most carefully developed individual intelligence tests" (Anastasi, 1988, p. 312-313). This is a strength because the norms are better established for group tests than for individual tests.

Technical Adequacy. Reliability refers to the consistency of measurement. For administrative purposes and when test data are reported for groups of individuals the reliability coefficients should be in excess of .60. When making decisions regarding individuals however, .90 is necessary (Ysseldyke, Algozzine & Thurlow, 1992). Validity refers to how much a test measures what it purports to measure and "is assessed depending on the purpose, population and environmental characteristics in which measurement takes place" (Schumacher & McMillan, 1993). Major group tests such as the Cognitive Abilities Test (Thorndike & Hagen, 1986) the Otis-Lennon School Ability Test (Otis & Lennon, 1982) and the Goodenough-Harris Drawing Test are those which will be discussed as to their technical adequacy.

The Cognitive Abilities Test was standardized using 332,000 students with 14,000 at each grade level. All reliability coefficients exceed .83. The Otis-Lennon School Ability Test (Otis & Lennon, 1982) has reported reliability coefficients exceeding .93. The Goodenough-Harris Drawing Test (Harris, 1963) reports "reliability

coefficients ranging from the low .80s to as high as .96" (Salvia & Ysseldyke, 1988, p. 211). These three tests seem technically adequate for groups.

The latter test is said by Short-DeGraff and Holan (1992) that it may be used more as a screening of perceptual-motor functions than as a general gauge of intelligence. Short-DeGraff, Slansky and Diamond (1989) stated that figure drawing ability as analyzed from the Goodenough-Harris Drawing Test, is more highly correlated with the Performance than with the Verbal components of IQ tests. Reliability of this test has been done by various researchers (McCarthy, 1944; Dunn, 1967; Harris, 1963). McCarthy (1944) found the retest correlation to be .68 and split-half was .89. Studies by Dunn (1967) and Harris (1963) yielded similar results. With .60 the adequacy point for use with groups of individuals, this test appears adequate. Strommer and Smith (1987) found that the 72-item full scale showed good internal consistency at all ages. In addition, as McCabe and Hilmo (1985) found, this test, when given at regular intervals during periods of remediation, may show clear evidence of improvement in behavior and attitude in learning disabled students.

Individual Tests

Advantages. There are two main advantages to administering an individual intelligence test. First, the examiner may be able to establish rapport and maintain the interest with the person being tested. In doing so, the examiner might discern any condition of the

examinee such as, fatigue or anxiety which could affect the outcome of the testing situation. Following this line of thought, original and extraordinary responses of a student become available to the examiner so that they may be taken into account as anecdotal information about the student (Anastasi, 1988; McCluskey & Walker, 1986). An individual test may indicate to an alert examiner those attributes of the testing situation which may reveal useful information about the student, for example, "fluency and maturity of expression, the fine motor skills, posture, restlessness, sense of humor, poise and mood" (Robinson & Chamrad, 1986, p. 160-161).

Second, some individual tests such as the Peabody Picture Vocabulary Test-Revised, the Quick Test and the Slosson Intelligence Test can be administered by resource teachers thus they don't require a great deal of training to administer them (McCluskey & Walker, 1986). The amount of training one has in regards to using tests of this nature is important. One must know what behaviors to look for and how to evaluate the responses obtained from the test taker. Through testing, the examiner develops: (a) several standard scores which give normative data; (b) incidental observations about the student such as fluency, maturity, restlessness etc.; and (c) a rich body of information, most of it reasonably objective (Robinson & Chamrad, 1988).

One criterion of representativeness involves the standardization samples. They must be normed on major variables

such as ethnicity and socioeconomic status. Kaufman (1992) explained that tests such as the WPPSI-R and WISC-III "pass this criterion with flying colors" (p. 55). A further strength is that standardized tests yield stable scores. A combination of these scores and achievement scores permits a more accurate prediction of future academic accomplishments (Robinson & Chamrad, 1988).

Technical Adequacy. The technical adequacy (which includes the reliability and validity) of tests such as the WISC-R and WISC-III have been discussed extensively in the literature. The Wechsler Scales have impressive Full Scale IQ reliabilities of .95-.96 for the WISC-R, .96-.97 for the WPPSI, .96-.98 for the WAIS-R and .96 for the WISC-III. Kramer (1993) stated that the Wechsler scales provided more than a separate score, by giving individual scores on separate tasks. In addition, it has been noted that Digit Span in the WISC-R (Kaufman, 1979) and WISC-III (Kamphaus & Platt, 1992) could be interpreted individually with confidence by the examiner. "Using individual subtest scores, however, requires consideration of several theoretical and psychometric issues . . . performance successes and failures can be attributed to a number of different possible underlying cognitive skills" (Kramer, 1993, p. 193). Another strength noted by Carroll and Horn (1981) concerned that ability tests have been standardized so they can be uniformly applied.

The Stanford-Binet: Fourth Edition, consisting of 15 subtests grouped into four areas, verbal reasoning, quantitative reasoning,

abstract-visual reasoning and short-term memory, has a standardization sample of over 5,000 cases. The composite score on the entire battery yielded high reliabilities, ranging from .95 to .99. The reliabilities of each cognitive area range from .80 to .97. With standards of reliability for making decisions about individuals being .90 (Salvia & Ysseldyke, 1988), there are some subtests that agree with this standard while others do not.

The Kaufman Assessment Battery for Children (K-ABC) (Kaufman & Kaufman, 1983) was standardized on 8,000 children aged 2 1/2 to 12 1/2 years. Test-retest reliabilities for global scores range from .77 to .97. Anastasi (1988) stated that "the K-ABC is an innovative, cognitive assessment battery whose development meets high standards of technical quality" (p. 269-270).

The Peabody Picture Vocabulary Test-Revised lists reliabilities for standard scores for single age groups from .54 to .90 (Salvia & Ysseldyke, 1988). Due to the large range of these reliabilities, Salvia & Ysseldyke (1988) stated that the PPVT-R "has satisfactory reliability for screening purposes, the intended use of the test" (p. 183). However, this range poses a limitation for the use of this test as the standard reliability is .90 for making decisions about individuals.

Summary. The technical adequacy for each of the above tests have reliabilities that if used properly, and for the purposes they were intended, are reasonably adequate.

Limitations of Intelligence Testing

Some people feel there are many limitations that underlie IQ testing in the literature. They underlie the concept of intelligence, culture fair tests, the disadvantages of using group and individual tests, technical adequacy of the instruments used, and the controversy over whether IQ is controlled by heredity, environment or both. There is also an assumption that each individual has had an equal opportunity to learn what is tested, and an equal quality of opportunity. Further, problems could ensue because intelligence tests are used for inappropriate purposes.

IQ testing evaluates certain areas of knowledge and those areas have influenced subject areas and curriculum (Gardner, 1987a). Thus, the educational system seeks to sort, rank, and classify those according to those areas valued on an IQ test (McKenna, 1982).

Concept of Intelligence

Definitions of intelligence are numerous. Major theorists such as Spearman, Thurstone, Wechsler, Guilford and more recently, Gardner, explained the concept of intelligence in various ways. Spearman (cf. Wagner and Sternberg, 1984, p. 181) stated that "intelligence comprises a general factor (g) common to performance on all tests that are used to measure intelligence plus a specific factor (s) involved in performance on each individual test". Later, Thurstone (cf. Wagner and Sternberg, 1984, p. 181) described intelligence as "comprising a set of approximately seven primary

mental abilities, namely, verbal comprehension, verbal fluency, number, spatial visualization, perceptual speed, memory, and reasoning". Wechsler (1974) stated that "intelligence is the overall capacity of an individual to understand and cope with the world around him" (p. 5). Guilford (cf. Wagner and Sternberg, 1984, p. 181) stretched his understanding of intelligence to encompass "150 factors, each involving an operation, a content and a product". More recently, Gardner (1983) chose to include cultural fairness within his definition which states that an "intelligence" as "the ability to solve problems, or to create products, that are valued within one or more cultural settings." (p. x.) Taking all these definitions into consideration, one realizes the complexity of intelligence, as well as the difficulty with which one might attempt to measure it. It is this complexity which alludes to the fact that intelligence may be too complicated to measure because it seems to exist in so many forms for so many people.

Intelligence is not "a single unitary ability, but a composite of several functions. . . . The term is commonly used to cover that combination of abilities required for survival and advancement with a particular culture" (Anastasi, 1988, p. 363). Moreover, these abilities and their relative value in a culture may change over time. "In different cultures and at different historical periods within the same culture, the qualifications for successful achievement differ" (Anastasi, 1988, p. 363). "In a society that harbors traditional religious

school, linguistic knowledge comes to be esteemed In modern secular educational settings, logical mathematical knowledge is at a premium, and certain forms of linguistic competence are also of value" (Gardner, 1983, p. 337).

Tests of this nature were never developed to measure mechanical, motor, musical and artistic aptitudes. Thus if one is using an intelligence test for that purpose or using it as the sole indicator of performance, one would clearly be misusing the instrument (Anastasi, 1988). Intelligence tests demand different things of different people even if identical items are used, due to the interaction between acculturation and the behavior sampled (Salvia & Ysseldyke, 1988). The kinds of behaviors, sampled by intelligence tests, such as discrimination, generalization and motor behavior etc., interact with the background experiences of the student being assessed so that the combination among different people are endless. Thus, we are measuring different things for different children. This seriously limits how one might use data from such tests.

Discussion regarding the controversy over whether IQ is controlled by heredity, environment or both is a moot point. IQ tests, by their very nature, must invite the notion of environment into the equation as each person who takes the test has experienced life differently. Added to this, is the realization that each individual has had varied opportunities to learn. Therefore, children from homes who have had few or no opportunities to visit cultural events such

as the symphony, museums or travel, will have limited knowledge of such experiences. The same can be said for questions on intelligence tests. One would score differently dependent on the experiences with questions of that sort.

One major difficulty with the concept of intelligence is that "no one has seen a thing called intelligence" (Salvia & Ysseldyke, 1988, p. 197). One may only infer intelligence from a sample of behavior gained through testing. If we are to use the test to predict future behavior, one must choose a test that relates to the behavior one is trying to predict (Salvia & Ysseldyke, 1988). Further, one must remember that intelligence is not a fixed entity, that it changes with time, instruction and experiences (Anastasi, 1988).

These are limitations for three reasons. First since the definitions are so diverse, one person's idea of who is intelligent and who isn't could be in conflict with what another person thinks. Second, definitions by earlier scholars are more specific and number oriented as compared to Gardner who invites more wholistic assessment in context with the classroom. Third, "decisions about which abilities to measure in a particular IQ test have been based upon judgements that have no rigorously defended rationale in theories about human intelligence (Carroll & Horn, 1981, p. 1017).

Culture Fairness

Much has been written regarding culture fairness within IQ tests and it has been well documented that culture bias affects the

scoring of mental tests (Feuerstein, 1979; Gardner, 1983, 1991; Jensen, 1980; Mercer, 1972, 1973, 1974, 1977; Mercer & Brown, 1973). Further, if a member of a different culture, one on which the test has not been normed, is administered the test, the bias is set in motion. This poses severe restrictions upon the use of the test itself. "Persons from backgrounds other than the culture in which the test was developed will always be penalized" (Jensen, 1980, p. 2). It is a disservice to students who live in a culture inside or outside the North American realm to test them using an instrument which has inherent culture bias imbedded in many test items. A specific example, illustrated by the street beggar - organized charity item in the WISC-R, is particularly faulted by Freides (1978), who criticizes it for its cultural bias and "general inappropriateness for youngsters" (p. 232). Children may have their own understanding of who they would rather give money to (a street beggar or an organized charity) dependent on the culture in which they have been raised. If they are unfamiliar with organized charities, they may choose the answer of "street beggar" and visa versa.

Some tests such as the Culture Fair Intelligence Test by Cattell, the Goodenough-Harris Drawing Test and the Kahn Intelligence Test: A Culture-Minimized Experience focus on nonverbal tasks deemed to be culturally neutral. However, "the nonverbal tasks that permeate the culture-fair tests may be as culturally loaded as the verbal items of the more traditional intelligence inventories"

(McCluskey & Walker, 1986, p. 27). Anastasi (1988) concurred with this criticism. This is due to tests using either the manipulation of objects or pantomimed directions. Here, one must be familiar with the objects used during the test and the hand gestures associated with the giving of instructions.

The problem of cultural loading on tests demands that one look at the ways cultural differences may operate to bring about differences in behavior. First, some examinees respond to some test items poorly due to the lack of familiarity with the objects (Anastasi, 1988). If one hasn't had the opportunity to see or use the objects, it is unfair to assess someone using them. For example, in the BITCH Test, (Black Intelligence Test of Cultural Homogeneity), it compares the understanding of slang used by black Americans with white and black populations. This would serve as an index of the degree of familiarity with this language. Other cultural differences include attitudinal or motivational differences which are likely to affect test performance. This is especially true for students who have had adverse experiences with tests and intellectual activities with authority figures (Anastasi, 1988). For this reason, "culture permeates nearly all environmental contact" (Anastasi, 1988, p. 357), so that cultural influences will affect test performance. "It is therefore futile to try to devise a test that is *free* from cultural influences" (Anastasi, 1988, p. 357).

Group Tests

Disadvantages. Concerns have been raised regarding the limitations of group tests. Anastasi (1988) maintained that "in group testing, the examiner has much less opportunity to establish rapport, obtain cooperation, and maintain the interest of the examinees" (p. 313). Due to this lack of rapport, the examiner also has less opportunity to make note of conditions such as fatigue, worry or anxiety on the part of the examinee. A second concern noted by Anastasi (1988) pertained to the fact that group tests have multiple choice items, all which must be answered. This set-up limits the responses of the examinee and inhibits the creative and original answers one might obtain with more open-ended responses.

Technical Adequacy. Since the definition of reliability refers to the consistency of measurement and to the minimization of the influence of chance or other variables (Schumacher & McMillan, 1993), that, which would interfere with the reliability, or affect the variables, would cause the reliability to be less. Anxiety, worry or fatigue may influence the outcome of the testing due to the examiner being unaware of these characteristics being present (Anastasi, 1988; Salvia & Ysseldyke, 1982). "Persons unaccustomed to testing may be somewhat more handicapped on group than on individual tests" (Anastasi, 1988, p. 313). Validity refers to "the degree to which the test actually measures what it purports to measure" (Anastasi, 1988). Since other factors, mentioned above, may very well affect the

outcome of the test, the validity of the outcome could be called into question. Too, the amount of anxiety, worry or fatigue is an unknown factor in the final analysis.

Individual Tests

Disadvantages. One major disadvantage of individual tests such as the Wechsler and Stanford-Binet Scales is the amount of training needed to administer them. One must have the necessary qualifications developed through a long period of intensive training and supervised experience (Anastasi, 1988). However, some tests such as the Test of Nonverbal Intelligence and the Slosson Intelligence Test "can be given by people with no formal training in assessment" (Salvia & Ysseldyke, 1988, p. 160). Since it is understood that "one of the basic assumptions underlying psychoeducational assessment is that the person who uses tests is adequately trained to administer, score and interpret them" (Salvia & Ysseldyke, 1988, p. 160), people who are untrained may make grave errors in all facets of the testing process. As well, even though training may ensure the adequate scoring of the instrument, an examiner who has years of experience might interpret the instrument differently from someone who has just recently learned to administer tests of this sort. Even though the instrument used is standardized, training is necessary to administer, score and most importantly, interpret it accurately. Analyzing an individual test is different from analyzing a group test in that a group test has standard answers from which an examinee

choices. It is from those multiple choice answers, one obtains a score. With an individual test, there are several elements at play. First, the examiner has the opportunity to view the emotional state of the examinee, evaluate their motivation towards answering the questions, and assess the answers that are given. A highly different answer might be questioned and the examinee has the opportunity to explain the response, which could be highly creative, and not a choice given in the examiner's guide. Thus, since individual intelligence tests are most frequently used for making educational placement decisions, it is imperative that the information gained through testing is accurate and useful.

A second major disadvantage is that since these tests are used to measure intelligence, one must understand what intelligence is. "No one has seen a thing called *intelligence*. Yet, there are many tests of this thing that no one has seen, and assessors are regularly required to assess it" (Salvia & Ysseldyke, 1988, p. 197). One must realize that assessment using intelligence tests is only a sample of behavior (Salvia & Ysseldyke, 1988).

A third disadvantage is that external influences can affect the outcome of testing. External influences constitute two possible variables, (a) the manner in which the instructions are given such as describing the test as a "game" versus describing it as a "test" (Bridgeman, Strang & Buttram, 1974) and (b) examiner incentives in the form of verbal approval or disapproval during the testing

procedure (Galdieri, Barcikowski, & Witmer, 1972; Witmer, Bornstein, & Dunham, 1971). In the Galdieri study, examiner incentives in the form of verbal approval as opposed to no approval showed no significant difference in test performance. However, when comparing approval versus disapproval versus neutral behavior on the part of the examiner, the approval group performed significantly higher (Witmer et al. 1971). Bornstein (1968) found that subjects who received approval received significantly higher performance quotients. Exner (1966) found performances to be adversely affected by an aloof style of administration.

Giving instructions describing the test as a "game" or as a "test" yielded unexpected results; scores on verbal tasks with sixth graders were higher when they were given as part of a "game" when compared with "test" instructions. "A possible explanation of these findings is that motivation of students tends to be low for tasks which the student perceives to be academically unimportant" (Bridgeman et al., 1974, p. 286).

Internal influences involve the effect of anxiety on performance of the WISC (Oros, Johnson, & Lewis, 1972). Induced anxiety was found to depress the subjects level of performance. Thus it is important to administer the test correctly as influences other than the test, or those from the examiner, may interfere with the outcome of the testing situation. If the test is done correctly, the

results would be deemed valid. If it were administered incorrectly, the results would be invalid.

Technical Adequacy. On tests such as the Stanford-Binet and the Wechsler scales there are high reliabilities with Full Scale IQ coefficients all .90 and above for both scales. However, the Slosson Test is based on high correlations with the Stanford-Binet citing the Slosson as a parallel form of the Stanford-Binet. However, they are not parallel forms and the arguments for reliability are therefore invalid (Salvia & Ysseldyke, 1988).

Picture Vocabulary Tests such as the Quick Test and the Peabody Picture Vocabulary Test-Revised (PPVT - R) all have questionable reliabilities ranging from .60 -.96 for the Quick Test and .71-.89 for the PPVT-R. Since a standard of .90 is necessary to make a placement decision, these tests wouldn't be advisable for this purpose. "The PPVT - R has satisfactory reliability for screening purposes, the intended use of the test" (Salvia & Ysseldyke, 1988, p. 183). Screening is used to identify who may need remediation or a referral for additional assessment.

Reliability. The empirical reliabilities of mental tests are .94, .90 and .96 for the WISC-R and .95, .91 and .96 for the WISC-III for Verbal, Performance and Full Scale IQ's. The same can't be said for the subtest scaled scores. Vernon (1984) listed that the subtest reliabilities for the WISC-R ranged from .62 to .92. This span is quite wide which indicates that some subtests are more reliable than

others. Tittle (1978) stated the subtest scaled scores "have reliability coefficients as low as .57 in one instance, and the median coefficient across all ages and subtests is only .77" (p. 353b). Therefore subtest scores should be treated with caution.

"A coefficient of stability is provided by correlating scores from the same test of a group of individuals on two different occasions. If the responses of the individuals are consistent . . . then the correlation coefficient, and the reliability, are high" (Schumacher & McMillan, 1993, p. 229). This is important because one must know that an instrument used will have reliable scores. Stability studies have been accomplished, mostly over a three year period, with children from various ethnic backgrounds. It has been shown that the influence of three year's time on intelligence test performance was clinically insignificant (Elliott & Boeve, 1987; Oakman & Wilson, 1988; Webster, 1988). This means that a span of three years on the outcome of IQ testing didn't change scores that much. While these results have been obtained, others suggest that we should treat the results of testing with caution due to decreases over time of Verbal, Full Scale and subtest scores (Schmidt, Kuryliw, Saklofske, & Yackulic, 1989; Anderson, Cronin, Kazmierski, 1989; Bauman, 1991). Here, these researchers found that scores decreased over a function of time.

Two conditions that influence test reliability are scoring and the range of ability in the sample (Jensen, 1980). "The scoring of

many of the items in individually administered intelligence tests such as the Stanford - Binet and the Wechsler scales, requires a subjective judgment on the part of the tester as to whether the examinee passed or failed the item" (Jensen, 1980, p. 266). However, the scoring within the WISC-R is limited to the selected answers provided in the manual for which points are given. Thus, there seems to be a contradiction here with regards to subjective or objective judgement. The second condition, range, "is a joint function of the test and the group of persons to which it is given" (Jensen, 1980, p. 268). The range is an element of the culture within which the examinee belongs. The range of scores obtained from the sample is "the difference between the highest and lowest scores in the distribution . . . and can be misleading if there is an atypically high or low score" (Schimacher & McMillan, 1993, p. 206). Bearing this in mind, if, within the sample used to discover the reliability of an instrument, there are atypical scores, the range of scores could be misleading. Thus, the sample of individuals chosen must be representative of the population you are testing.

Ysseldyke, Algozzine and Thurlow (1992) have found that students with disabilities are often excluded from norming populations, yet some educators feel that standards set for students should be applied to all students. "There is little demonstrated validity for making decisions about students with disabilities on this basis" (p. 178).

Validity. Validity is the most important element in a test. The instrument should measure what it purports to measure (Anastasi, 1988). Yet, there are three main elements which could affect the validity of an instrument. First, the experience of the examiner may play a role, such that creative answers may not be given credit because the examiner has adhered to a rigid set of answers in the manual (Salvia & Ysseldyke, 1988). Second, performance by the examinee may be affected by anxiety, motivation or fatigue (Oros, Johnson & Lewis, 1972). A diligent examiner must take these elements into account during evaluation. Third, the use of an intelligence test is an artificial means of evaluating a student. The examiner is usually a stranger to the test taker and evaluates the student over a short period of time rather than in a classroom, participating in real-world problem solving activities (Gardner, 1983, 1991, 1993).

Summary of Review of the Literature

Purposes of Testing. There are different purposes for using group and individual tests with children. Group tests (according to the literature) are used as screening devices to identify those who require further assessment and to provide information as to a student's capabilities in the classroom. Individual tests are used to provide information regarding an individual's cognitive strengths and weaknesses.

Strengths of Intelligence Testing. According to the literature, advocates of IQ testing describe strengths dealing with the concept of intelligence and the advantages of group and individual testing. The testing of a sample of behaviors serves as an indicator of other areas of behavior, (Anastasi, 1988) whereby you may obtain a general understanding of function based on test performance.

Group tests may be administered to many people at the same time, and an extensive amount of training is not required for scoring. Group tests such as the Goodenough-Harris Drawing Test and the Otis-Lennon School Ability Test have adequate reliability for groups.

With individual tests, the examiner is able to establish rapport with the examinee and in doing so, is able to discern conditions such as fatigue or anxiety which could affect the outcome of the testing. Individual tests such as the WISC-R, WPPSI, WAIS-R and WISC-III have adequate reliabilities for use with individuals.

Limitations of Intelligence Testing. The literature describes problems surrounding IQ testing which encompass the concept of intelligence and the disadvantages of group and of individual tests.

There are many definitions of intelligence according to Spearman, Thurstone, Wechsler, Guilford and Gardner. With so many definitions, it would be difficult to decide which tests would be the one to give the "true" answer or score about a person's intelligence as each conveys the understanding of intelligence in a different way which demonstrates the complexity and difficulty with

which one might attempt to assess intelligence. As well, the composition of the tests has been called into question as the kinds of questions asked omit mechanical, musical and other aptitudes. Intelligence must be seen as something that changes with experiences and maturation (Anastasi, 1988).

Group tests limit the amount of rapport an examiner has with the student. Thus, knowledge of conditions such as fatigue and anxiety about the testing situation would be unavailable to the examiner. As well, responses which could be construed as original and creative would be missed as the examiner has no way of questioning the student about the responses given. This problem could have major implications for validity of the testing.

Individual tests require a substantial amount of training to administer them as the scoring and interpretation, together with the rapport necessary to adequately put the student at ease, are important attributes to an understanding of the responses. A major disadvantage is that there are many definitions of intelligence and tests are developed according to each definition. One must remember that using a test reflects only a sample of abilities. As well, influences such as anxiety may affect the outcome of testing. Tests such as the PPVT-R and the Quick Tests have questionable reliabilities for making decisions about individuals.

Intelligence tests have advantages and disadvantages whether they be group or individually administered. All in all, they do five

things. First, "the scores reflect the nature and amount of schooling the individual has received rather than measuring "capacity" independently of relevant prior experiences" (Anastasi, 1988, p. 328). Feuerstein (1979) stated an equivalent idea.

Second, assessment is imperfect. An instrument only samples a limited range of behaviors to arrive at a score. Thus the score is imperfect as well.

Third, errors transpire during the testing situation with respect to responses and the past experiences needed to formulate the responses. Further, conditions such as fatigue or anxiety could affect the examinee's performance.

Fourth, no single battery of tests will tell us everything we need to know about a student. Thus assessment utilizing multiple forms including the use of classroom observations is necessary.

Fifth, one must look at the whole child to determine a course of action. "Students with academic and behavioral problems are assessed because somebody thinks there's something wrong with them" (Ysseldyke & Algozzine, 1990, p. 349).

CHAPTER III

METHOD

Population

For the purposes of this study, the population chosen consisted of psychologists who are members of the Manitoba Association of School Psychologists (MASP), who work with children, and the Child Guidance Clinic (CGC). Permission was previously obtained from the MASP (see Appendix A) and the CGC (see Appendix B) to acquire a list of names. Both groups incorporate psychologists who practice psychology with children.

This survey was mailed to a total of 132 psychologists in Manitoba who were from CGC and MASP. Since there were many psychologists who had a mailing address at 700 Elgin Ave., Wpg., (the principle establishment of Child Guidance Clinic psychologists), 49 surveys were mailed to the CGC at that address in one box. On the front of the box was a note; "surveys enclosed: please distribute". The surveys mailed were enclosed in their own envelope, therefore, the secretary would do the distributing. As well, there was a group of 17 who were listed as school psychologists but not members of CGC or MASP. The researcher received their names by mail with those from MASP through the MASP contact person. These members also received a survey. The respondents were given three weeks from the time of mailing to respond to the survey.

Instrument

A cover letter to the survey respondents to the Manitoba Association of School Psychologists (see Appendix C), to the Child Guidance Clinic (see Appendix D) and a questionnaire developed by the researcher (see Appendix E) was developed based on the need to discover more indepth views of Manitoba psychologists regarding their views on (a) which tests are used by Manitoba psychologists; (b) the purposes for which Manitoba psychologists use intelligence testing; (c) the perceived strengths about intelligence testing with children; and, (d) the perceived limitations about intelligence testing with children. Part I of the questionnaire covered demographic information and included 6 questions. Part II asked specific questions about intelligence testing and included seven open-ended questions.

Demographic information included whether they used intelligence tests with children in their practice, the type of practice they had (whether they had a private practice, worked for an organization or both) and the number of years they had practiced as psychologists. It was specified in the covering letter that it would take approximately one hour to complete the questionnaire.

Pretesting. The purposes of the pretest were to determine (a) whether the questionnaire had face validity and (b) the clarity of directions and questions. The subjects selected for the pretest were not included in the study and were selected through simple random sampling. The pretest group consisted of three professional

educators from the Lord Selkirk School Division No. 11. The subjects were given the pilot survey in person to complete and directions were given to make comments regarding the questions and the format of the instrument. They all found the survey to be clear and easy to understand. Although the pretesting group was small, consisting of three people, the information gained was invaluable.

Procedure

Permission was obtained from the Ethics Committee of the Faculty of Education at the University of Manitoba for the researcher to conduct the study. The study was conducted by mail in the spring of 1995, with the respondents given clear information within a covering letter as to the importance of the study, their importance as a respondent, a short description of the questionnaire and assurance that their responses would be anonymous and confidential. Anonymity was guaranteed as there were no identifying marks on the surveys mailed out. As well, there were separate envelopes for the return of the survey and for the results request form (see Appendix G), if they desired to send it. Respondents were assured that their cooperation was voluntary. A reminder letter, (see Appendix F) encouraging each participant to complete the survey, was mailed in mid June, 1995, three weeks after the initial mailing to all respondents as the sample size was below 70%. Separate

envelopes were enclosed for the return of the surveys and the results request form.

Analysis

Demographic data in Part 1 was organized in frequency tables, and responses to each open-ended question in Part 2 were grouped (by the researcher) into categories that reflected the dominant themes in each question. The categories were not pre-determined but developed qualitatively as the researcher accumulated data and organized it. They were determined to be dominant themes through a grouping of similar responses. A title was designated for each category which reflected a description of each section. All responses were organized and entered within the appendices found at the end of this study. Results may be found in Chapter IV and the Discussion in Chapter V.

CHAPTER IV

RESULTS

Of the 132 surveys initially mailed, 4 were returned to the researcher with a stamp indicating the respondent had moved. (No other addresses could be obtained). Therefore, only 128 were truly mailed. Of the 71 surveys returned, 3 respondents described themselves as administrators who do not use intelligence tests at this time and chose not to answer the questionnaire. Another respondent, an administrator, indicated that he administered tests only in rare circumstances, and chose to answer the questions. This leaves 68 usable surveys.

In total, 132 surveys were mailed and 71 surveys were returned. This represents a return rate of 54.4%. It is stated by Gay (1981) that, "If your percentage of returns is not at least 70%, the validity of your conclusions will be weak" (p. 164). However, this return rate is not completely out of line. Five previous studies regarding psychological assessment practices had similar response rates: Hutton, Dubes and Muir (1992) 39%; Goh, Teslow and Fuller (1981) 54.8%; Wade and Baker (1977) 50.1%; Brown and McGuire (1976) 47.3% and Lubin, Wallis and Paine (1971) 44%. Nonetheless, with my return rate at 54.4%, my conclusions will be weak.

One reason for a poor return might have been the time of year. The surveys were mailed at year end, late May and early June, a busy time. A second reason could be that those who are truly interested

were those who responded as the bulk of the responses arrived soon after the first mailing. A reminder letter was mailed three weeks later and only 6 surveys were returned subsequent to that mailing. "People who have a particular interest in the subject matter or the research itself are more likely to return mail questionnaires than those who are less interested. This means that mail surveys with low response rates may be biased significantly in ways that are related directly to the purposes of the research" (Fowler, 1993, p. 41).

Part 1 Results

Demographic Information

Data regarding gender and the number of years spent as a practicing psychologist are presented in Table 1.

Table 1

Number of Years as a Practicing Psychologist

	Male	Female	Totals
0 - 5 years	2 (6.9%)	12 (30.8%)	14 (20.6%)
6 - 10 years	6 (20.7%)	7 (17.9%)	13 (19.1%)
11 - 15 years	8 (27.6%)	8 (20.5)	16 (23.5%)
16 - 20 years	4 (13.8%)	4 (10.3%)	8 (11.8%)
21 - 25 years	7 (24.1%)	6 (15.4%)	13 (19.1%)
25 + years	2 (6.9%)	2 (5.1%)	4 (5.9%)
<u>Totals</u>	<u>29 (42.6%)</u>	<u>39 (57.4%)</u>	<u>68 (100%)</u>

There were 29 (42.6%) male and 39 (57.4%) female respondents with more females responding than males and a low response rate for the 25+ category.

Table 2 presents the type of practice each psychologist had.

Table 2

Type of Psychological Practice You Have

<u>Practice</u>	<u>N</u>	<u>%</u>
Private Practice	2	2.1 %
I work for an organization (Hospital, School Division)	62	91.2%
I work for an organization and have a private practice	4	5.9%
Total	68	100%
Other	3	4.4 %

There were 2 who practiced privately and 62 who worked for an organization. Four of those who said they worked for an organization stated that they worked for a school division. None said they worked for a hospital. There were 4 who worked for an organization and had a private practice. The total was 68. The respondents who chose to answer the "other" category were included within the total number of 68 and chose to specify other work they did, such as teaching a psychology course at the University of Manitoba; part time faculty member at Brandon University; research

and consultation in environmental psychology and work at the school board level. Table 3 illustrates the percentage of the work day spent on testing.

Table 3

Percentage of Work Day Spent on Testing

	<u>N</u>	<u>%</u>
0 %	1	1.5%
1 - 25%	40	58.8%
26 - 50 %	18	26.4%
51 - 75 %	5	7.4%
No response	4	5.9%
Total	68	100%

The percentage ranged from 0 - 75%, with the bulk of the respondents (n= 41) indicating they spent between 0 - 25% of their time testing. Eighteen respondents indicated that they spent between 26 - 50% of their time testing and 5 said that they spent 51 - 75% of their time testing. One respondent indicated that no percentage of the work day was spent testing. Yet the respondent answered subsequent questions detailing the individual intelligence tests used. Four respondents didn't answer the question. Nine of the 68 respondents answered using phrases and sentences describing the length of time spent. These were translated into percentages.

Table 4 indicates the highest degree held by each respondent. The question in the survey listed only B. Ed., M. A., M. Ed. and Ph. D. However, this was an incomplete list as some respondents added in B. A., B. Sc., M. Ps., B. A. T. and M. Sc.

Table 4

What Degrees Do You Hold?

Degree	N	%
M. A.	33	48.5%
M. Ed.	15	22.1%
Ph. D.	14	20.6%
M. Sc.	3	4.4%
M. Psychology	1	1.5%
B. A. T.	1	1.5%
B. A.	1	1.5%
Total	68	100%

There were 33 (48.5%) respondents who indicated they had M. A. degrees, with 15 (22.1%) respondents saying they had M. Ed. degrees. There were 10 (14.7%) respondents who said they had 2 or 3 degrees. The other 58 (85.3%) respondents listed only 1 degree, the highest attained.

Table 5 shows the number of people who hold a teacher's certificate. There were 18 (26.5%) who did and 50 (73.5%) who did not.

Table 5

Do You Hold a Teacher's Certificate?

	<u>N</u>	<u>%</u>
Yes	18	26.5%
No	50	73.5%
Total	68	100%

Table 6 shows whether the person used intelligence tests with children.

Table 6

Do You Use Intelligence Tests With Children?

	<u>N</u>	<u>%</u>
Yes	67	98.5%
No	1	1.5%
Total	68	100%

There were 67 (98.5%) people who said that they did. The person who said no stated that they assessed people with mental disabilities and deemed intelligence tests unfair to use with such a population.

Part 2 Results

This next section will discuss the responses with the open-ended part of the survey. Results of each question and sub-question will be discussed in order, according to the dominant themes present within each set of answers. The themes will be described for each question.

Question #1. Do you use group intelligence tests with children?

	N	%
Yes	0	0%
No	67	98.5%
No Response	1	1.5%
Total	68	100%

One respondent said that the reliability was poor and one didn't answer the question. None indicated that they used group tests.

Question #1a. Which group intelligence tests do you use?

One respondent indicated that group tests were used while in training. Another respondent said, "No, other than screening tests such as the DIAL-R." This was an inappropriate response as the person indicated that they do not use group tests, yet the response indicates that the DIAL-R is used for screening purposes. Further,

Sweetland & Keyser (1984) stated that the DIAL-R is not for group use. No one else answered the question.

Question #1b. For what purposes do you use group intelligence testing with children?

One respondent noted that a group tests might be used to screen young adolescents entering high school, for a) special needs, b) admission requirement for private schools. No other respondents answered the question.

Question #2 Do you use individual intelligence tests with children?

	N	%
Yes	64	94.1%
No	0	0%
No Response	4	5.9%
Total	68	100%

There were 64 (94.1%) respondents who said they used individual intelligence tests with children. Four (5.9%) respondents didn't answer this question. Three of the four answered questions 2a and 2b. One respondent didn't answer parts of question 2 but answered subsequent questions in the survey. Thus, the true number and percentage of those who do use intelligence tests was 67 out of 68 respondents (98.5%).

Question #2a. Which individual intelligence tests do you use with children?

The frequencies of the usage of these tests is progressive from the most used to the least used and are listed in Table 7.

Table 7

Individual Intelligence Tests Used

	<u>N</u>	<u>%</u>
WISC - III/ WISC-R	66	(98.5%)
WPPSI - R	30	(44.8%)
McCarthy Scales of Children's Abilities	18	(26.9%)
Stanford-Binet IV	15	(22.3%)
WAIS - R	14	(20.9%)
Raven's Progressive Matrices	14	(20.9%)
TONI II	13	(19.4%)
Woodcock - Johnson Psychoeducational Battery	5	(7.5%)
Kaufman Assessment Battery for Children (K - A B C)	5	(7.5%)
Leiter International Performance Scale	4	(6.0%)
Developmental Test of Visual - Motor Integration (VMI)	3	(4.5%)
DTLA (Detroit Test of Learning Aptitude)	3	(4.5%)
Bender Gestalt Test of Visual Motor Integration	2	(3.0%)
Carrow Visual Abilities Test	1	(2.0%)
DAP (Draw - A - Person)	1	(2.0%)
Goodenough - Harris	1	(2.0%)

Table 7 continued

	<u>N</u>	<u>%</u>
Hiskey - Nebraska Test of Learning Aptitude	1	(2.0%)
Non - Verbal Test of Cognitive Skills	1	(2.0%)
Peabody Picture Vocabulary Test (PPVT)	1	(2.0%)
Slosson	1	(2.0%)
WIAT	1	(2.0%)
Connor's Teacher and Parent Scales	1	(2.0%)
CBC	1	(2.0%)
Child Development Checklist	1	(2.0%)
<u>Vineland</u>	1	(2.0%)

The bulk of the answers revolved around the top 7 tests listed in Table 7. There were 66 (98.5%) who said they used the WISC-III/WISC-R; 30 (44.8%) who said they used the WPPSI-R; 18 (26.9%) who used the McCarthy Scales of Children's Abilities; 15 (22.3%) who used the Stanford-Binet; 14 (20.9%) said they used the WAIS-R; 14 (20.9%) used the Raven's Matrices and 13 (19.4%) said they used the TONI-II. The other 22 tests listed represented between 7.5% to 2.0% of the total responses.

This pattern of tests which includes the Wechsler Scales (WISC-III and WPPSI-R) as the most used tests, was repeated in the following four studies. In the Hutton, Dubes and Muir (1992) study, the WISC-R was used by 97%, the WAIS was used by 74% and the WPPSI was used by 74% of the respondents. The Goh, Teslow and

Fuller (1981) study listed the WISC-R as the most often used intelligence test. Brown and McGuire (1976) listed the WISC as the most used test and in the Lubin, Wallis and Paine (1971) study the WAIS was the most used, while the WISC was the 7th most used. Some tests listed such as the Leiter International Performance Scale and the VMI are not intelligence tests but are tests of perceptual functioning. Therefore, some tests used as intelligence tests, are not considered to be intelligence tests.

Question #2b. For what purposes do you use individual intelligence tests with children?

This question represents the first in which the answers from 67 (98.5%) respondents were categorized into 16 dominant themes. (One respondent didn't answer the question.) There were 109 distinct responses to this question. (Actual responses for this question can be found in Appendix H.) Each of the subsequent questions will also be grouped according to themes with actual responses recorded in subsequent appendices. Table 8 lists the categories with the number of entries for each category and their corresponding percentages.

The themes of responses for this question indicated that the majority of reasons for using individual intelligence tests with children were:

Table 8

Purposes for Using Individual Intelligence Tests

Category	N %
Identification of Learning Potential	27(25.0%)
Placement	12(11.0%)
Assessment of Student Strengths and Weaknesses	11(10.1%)
Program Planning or Program Justification	10(9.2%)
Identification of LD and ADD	10(9.2%)
Identification of Emotional Difficulties	9(8.3%)
Assessment of Intellectual Skills or Processes (Learning Style)	9(8.3%)
Miscellaneous Comments	6(5.5%)
Funding	4(3.7%)
Help Others to Understand or Accept the Child and His or Her Behavior	3(2.8%)
Screening for Indepth Assessment	2(1.8%)
Examination of Modified Program Effects	1(0.9%)
Identification of Gifted	1(0.9%)
Identification of Mental Retardation	1(0.9%)
Identification of Syndromes	1(0.9%)
Neurological Problems	1(0.9%)

The largest category of responses for this question is the Identification of Learning Potential. Typically, in the literature this

means that they are testing to acquire a score and that score will show the learning potential of the student. Out of 109 distinct responses, there were 27(25.0%) which referred to learning potential. Reasons in this category spoke of the capabilities of students, cognitive potential or aptitude. The second category was Placement with 12 (11.0%) of responses mentioning that they used intelligence testing for assessing students for placement in special programs. The third grouping with 11(10.1%) of the responses was Assessment of Student Strengths and Weaknesses. Typically this means that they are using the tests to look for unusual patterns of scores suggesting particular areas of strength or weakness. The fourth category encompassing 10(9.2%) of the responses was Program Planning or Program Justification. Here, respondents mentioned that they used intelligence testing to provide assistance with program planning in schools and with individual educational plans. The fifth category was the Identification of LD and ADD in students. Here, there were 10(9.2%) of the respondents who mentioned they used IQ testing for this purpose. (This is the same percentage as for category 4). The focus was on the diagnosis of specific difficulties the child had, namely, LD and ADD. The Identification of Emotional Difficulties was the sixth purpose for using intelligence testing with children. There were 8.3% of the responses which mentioned the use of IQ testing for an emotional assessment. The seventh purpose listed was for Assessment of Intellectual Skills or Processes (Learning Style).

This would include a comparative look at reasoning skills as part of the assessment. Here, there were 9(8.3%) of the responses who mentioned assessment of learning style and skills as a purpose. (This is the same percentage as the sixth category.) The eighth category encompassing 6(5.5%) of the responses was the group of miscellaneous comments. These remarks included those which didn't fit any other category, some of which included "to find out what factors may be underlying "reluctance to write", "to assist older students in making realistic vocational choices" and "to develop a relationship (although there are other ways)". The ninth category was Funding with 4(3.7%) of the responses saying they used intelligence testing for special needs funding purposes. The tenth reason for using intelligence testing was to Help Others to Understand or Accept the Child and His or Her Behavior. Here, there were 3(2.8%) of the responses reporting that they used IQ tests to assist teachers/parents in understanding a child. The eleventh purpose for using intelligence testing was as a Screening for Indepth Assessment. Here, 2(1.8%) of the responses reported that IQ testing was used as a starting point for further assessment. The last 5 groupings had only 1 response or .9% of the responses mentioning that IQ testing was used for these purposes. In order, the twelfth purpose was for Examination of Modified Program Effects, the thirteenth reason was for Identification of the Gifted; the fourteenth rationale was for identification of Mental Retardation; the fifteenth

reason was for Identification of Syndromes and the sixteenth was for identification of neurological problems.

Question #3. If you do not use intelligence tests with children, please explain why.

For this question, there were only 8 (11.9%) respondents who answered. This statistic displays a discrepancy however, as 67 (98.5%) of the respondents said they they used intelligence tests. Perhaps there are times in which IQ tests would not be used. There were a total of 11 distinct responses grouped into seven groups of issues. Table 9 lists the categories, the number of entries in each, and their corresponding percentages. (Responses may be found in Appendix I).

Table 9

Why Psychologists Don't Use Intelligence Tests

Category	N	%
Tests Not Used With Some Clientele	3	33.3%
Programming Reasons	2	22.2%
Referrals Are For Emotional Behavioral Problems	2	22.2%
Authentic Assessment Preferred	1	11.1%
Lack of Time	1	11.1%
Previous Tests Available	1	11.1%
Use of a Variety of Instruments	1	11.1%
Total	11	100%

The first issue concerned the clientele the psychologists saw. There were 3 (33.3%) reasons given including the (a) increasing age of clients to include those between 16 - 18 years, (b) the fact that IQ tests weren't used with all children referred and (c) that IQ tests weren't used with very young children or children of another culture.

For the second category, Programming Reasons were given. There were 2 (22.2%) responses which mentioned that (a) IQ tests were used on a limited basis as most programming decisions didn't require a detailed assessment and another (b) reported using IQ tests judiciously and focusing on programming issues rather than IQ scores.

The third category had 2 (22.2%) responses that indicated that many students were referred for social - emotional problems rather than intellectual concerns. For the fourth, fifth, sixth and seventh categories, there was 1 response each or 11.1%. The fourth referred to a preference for authentic assessment; the fifth cited a lack of time; the sixth focused on using previous test results and the seventh mentioned the fact that IQ tests weren't the only instruments used.

Question #4. If you do not use intelligence tests with children, what other assessment procedures do you use?

These answers represented 18 (26.9%) respondents. This statistic signifies another discrepancy. There were 67/68 (98.5%) respondents who said they used intelligence tests with children. Yet, there were 18 of that 67 who now answer this question as if they

didn't use them. Perhaps there was a misinterpretation of the question, or, the wording of the question was misunderstood, or, they answered with the understanding that on those occasions where they wouldn't use IQ tests, these would be the assessment procedures they would use. As well, the specific answers given reflect a conflict. The question particularly asks for other assessment procedures. Some of the answers provided included intelligence tests among other test instruments. Other responses were more consistent with the answers expected. They reflected various assessment procedures used as well as general comments regarding assessment. (Actual responses may be found in Appendix J). Table 11 represents the test instruments mentioned. In this table, the number of times each of these tests is mentioned is noted. There were 8/18 (44.4%) respondents in total who mentioned using the tests listed in Table 10. Therefore, some respondents said they used more than 1 test as an alternate assessment procedure.

Table 10
Tests Used

Tests	Times Mentioned
Bender Gestalt	5
Beery	3
Thematic Apperception Test	2
WAIS-R	1
McCarthy Scales of Children's Abilities	1
Tests of Nonverbal Intelligence (TONI)	1
Stanford - Binet IV	1
Wide Range Achievement Test (WRAT 3)	1
Peabody Individual Achievement Test (PIAT)	1
Woodcock - Johnson - R	1
Draw - a - Person Test (1. 2 Quantitative and SPED)	1
Piers - Harris Children's Self-Concept Scale	1
House - Tree - Person	1
Animal Wishes/3 Wishes	1
Coopersmith Inventory	1
Achenbach Behavior Checklists	1
Children's Apperception Test	1
Rorschach	1
Vineland	1
SIB	1

Table 10 continued

Tests	Times Mentioned
CTAB	1
ICAP	1
AAMD Adaptive Behavior Scale	1
PIMRA	1
REISS	1
Leiter International Performance Scale	1
Hiskey - Nebraska	1
Perkins - Binet	1
Wechsler Memory Scale	1
Visual Motor Inventory (VMI)	1
API - MR	1
CDI	1
Total	39

There were 13 (72.2%) respondents who mentioned using alternate assessment procedures (as the question asked) and each respondent mentioned more than 1 procedure. The assessment procedures used refer to 12 different methods of learning more about a student. They are listed here in Table 11 in order of most mentioned to least mentioned.

Table 11

Other Assessment Procedures Used

<u>Procedure</u>	<u>x mentioned</u>	<u>%</u>
Developmental/Social Tests	7	14.3%
Classroom Observation	7	14.3%
Assessment Information		
from Specialists	6	12.2%
Family History / Assessment	4	8.2%
Parental Interview	4	8.2%
Structured Interview	4	8.2%
Teacher Reports	4	8.2%
Variety of Assessment		
Procedures Used	4	8.2%
Reasons Why Other		
Assessments Used	4	8.2%
Drawing	3	6.1%
Authentic Assessments	2	4.1%
<u>Total</u>	<u>49</u>	<u>100%</u>

There were two alternate assessment procedures mentioned the most: the use of Developmental and Social Tests and Classroom Observation. Out of 49 distinct responses, there was a tie with 7 (14.3%) respondents who mentioned these procedures. The third category with 6 (12.2%) responses cited the importance of assessment information from specialists. Examples of specialists mentioned

were pediatricians, community professionals, resource teachers, speech and language pathologists and social workers. The fourth, fifth, sixth, seventh, eighth and ninth categories all cited 4(8.2%) responses each. The fourth referred to interviews with families as alternate procedures. The fifth category mentioned the use of parental interviews. This was a separate category from family interviews as specific mentions of parents was involved. The sixth category referred to a structured interview with the child. The seventh cited teacher reports as an alternate assessment procedure. The eighth group cited a variety of assessment procedures used. The general idea here was that IQ tests were used along with various inventories. The ninth group referred to comments made as to why they chose alternate assessment procedures. This group referred to intelligence tests as not being appropriate for children of ESL background and those who were developmentally delayed. As well, one response specifically mentioned a preference not to use IQ tests due to biases, labeling and doubts regarding validity. The tenth group referred to drawings. There were 3 (6.1%) responses here. The eleventh group spoke of the use of authentic assessments. There were 2(4.1%) responses here.

Question #4a. Why do you use other assessment procedures?

The following answers represented 21 (30.9%) respondents. This statistic presents a problem as there were 67(98.5%) who said that they used intelligence tests with children. Two reasons why

answers were given to this question could be that (a) the respondents misinterpreted the question and answered whether or not they said they used intelligence tests in the previous question, or (b) they might use alternate assessments at some time or another. (Actual responses may be found in Appendix K).

For this question, there were 40 distinct responses in total falling into 6 categories. Each category will be reported separately. Table 12 displays the categories, the number of distinct responses and their corresponding percentages.

Table 12

Reasons for Other Assessment Procedures

<u>Category</u>	<u>Number of Responses</u>	<u>%</u>
Need to Get a Complete		
Picture of Student	18	45.0%
Planning/Programming Reasons	6	15.0%
Behavioral Reasons	4	10.0%
Technical Reasons	3	7.5%
Need for Interviews with		
Parents Teachers, Child.	2	5.0%
<u>Other</u>	<u>7</u>	<u>17.5%</u>

For this question, there were 40 distinct responses in total. In the first category, there were 18(45%) of the responses which spoke to the need of getting a complete picture of the student. For the second category, there were 6(15.0%) of the responses which cited planning

or programming as reasons for using alternate assessment procedures. Behavioral Reasons was the third category which was cited 4 times or 10.0% of the time. Here, students were referred for behavioral reasons more than for academic ones. As well, there was a need for information on motivation and attitude, which would be accomplished through the use of other assessment procedures. Technical reasons was the fourth mentioned reason. Here, 3(7.5%) of the responses cited this as the reason. There were responses which spoke of wanting to find true measures, to provide some measure of reliability and validity and that normed tests only gave clues which must be confirmed through the use of alternate assessment. The fifth reason was a need for interviews with Parents, Teachers and the Child. For the last category, there were 7(17.5%) of the responses which didn't fit the other categories. These responses referred to using alternative assessment as being congruent with systemic perspective, providing valuable information and that intelligence is a complex area.

Question #5. What do you perceive to be the strengths or positive aspects of the intelligence tests you use with children?

Group Tests

There were 13 distinct responses from 9(13.2%) respondents. Table 13 shows the 8 categories of the answers given for this question. (Responses may be found in Appendix L).

Table 13

Strengths of Intelligence Tests (Group Tests)

Category	Number of Responses	%
Positive Use	3	23.0%
Other Comments	10	76.9%
Total	13	100%

For the category Positive Use, there were 3(23.0%) of the responses which stated that they used group tests and gained helpful results from them by giving an indicator in a short time. However, there were 10 (76.9%) responses which noted other comments regarding the use of group inventories. Some of these were decidedly negative in nature where psychologists feel unfavorably toward group tests because of their limited value. One response mentioned that group tests weren't used, that tests which could be given to a group (such as the TONI) were used instead. A strongly worded response mentioned that group tests were only good if you wanted to save time and money and didn't worry about the validity of the results.

There is an interesting point to make regarding the responses to this question. The question specifically asked for strengths or positive aspects of group tests, yet, out of the 13 distinct responses,

there were only 3 (23.0%) which related to positive issues. The other 10 (76.9%) were decidedly negative regarding the use of group tests.

Individual Tests

The answers in this section represented 67 (98.5%) respondents and are classified into 15 categories. All respondents answered this question except one. The exception was the person who assessed those with mental disabilities. There were 188 distinct responses. Table 14 lists the categories, the number of responses in each, and their corresponding percentages. The categories are organized from the most to the least number of responses per category.

Table 14

Strengths of Intelligence Tests (Individual Tests)

Category	# of Responses	%
Assessing Cognitive Skills		
Cognitive Processes and		
Learning Style	50	26.6%
Observations of Individual		
Behavior Can Provide		
Encouragement to Expand		
Answers	21	11.2%
Comparative Purposes	19	10.1%
Program Planning	19	10.1%

Table 14 Continued

Category	# of Responses	%
Providing Understanding of		
Problems for Others	14	7.4%
Easy to Interpret and Administer	12	6.4%
Efficient - Get a Large Amount		
of Information in a Short Time.	12	6.4%
Identifying Emotional Problems		
and Reactions	11	5.9%
Technical Advantages	11	5.9%
Identify Learning Potential	7	3.7%
Predicting Academic Success	7	3.7%
Funding	2	1.2%
Grade Retention Decisions	1	.5%
Learning Disabilities	1	.5%
Screening for Special Needs	1	.5%
Total	188	100%

Psychologists perceived that intelligence tests are very good at assessing cognitive skills, cognitive processes and learning style. This was the group with the highest number of responses at 50(26.6%). The second highest group included observations of individual behavior - where it can provide encouragement to expand on

answers. Here, there were 21(11.2%) responses. Psychologists surveyed found that intelligence tests were used for comparative purposes as well as program planning. There was a tie with these two groups of responses with 19(10.1%) responses each. Providing understanding of problems for others was the fifth most mentioned strength of intelligence tests. There were 14(7.4%) of the responses which thought so. There were 12(6.4%) of the responses which said that intelligence tests were easy to interpret and administer. There were 12(6.4%) of the responses which said that there was a large amount of information gained in a short time; that the instrument were efficient. There were 11(5.9%) responses which said that intelligence tests were good at identifying emotional problems and reactions and 11(5.9%) who found technical advantages with intelligence tests. A tie was noted with those who found that IQ tests identify learning potential and those who found that IQ tests were good at predicting academic success with 11(5.9%) responses. Funding was only mentioned by 2(1.2%) responses as a strength of IQ testing. The last 3 categories were mentioned at the same rate, 1(.5%) responses. It was found that at least one psychologist used IQ tests for grade retention decisions, learning disabilities and as a screening for special needs as the lowest strength of IQ tests.

Question #6. What do you perceive to be the limitations or negative aspects of the intelligence tests you use with children?

Group Tests

These answers represent the responses of 9(13.2%) of the respondents. The 26 distinct responses are grouped into 8 categories. Table 15 lists the categories, the numbers of distinct responses and their corresponding percentages. (Actual responses may be found in Appendix M.)

Table 15

Limitations of Intelligence Tests (Group Tests)

Categories	# of Responses	%
Can't Observe Subjects	6	23.1%
Doesn't Examine Process		
Skills / Just Facts	5	19.2%
Limited Value for Classrooms	4	15.4%
Technical Issues	4	15.4%
Lack of Diagnostic Value	3	11.5%
Inappropriateness	2	7.7%
Not a Good Predictor		
of School Performance	1	3.8%
Susceptibility to		
Outside Influences	1	3.8%
Total	26	100%

For this question, there were 26 distinct responses. The highest number of responses referred to the fact that they can't observe subjects, that they had no opportunity to observe emotional reactions to the testing. Here, there were 6(23.1%) of the responses which described concerns regarding not being able to watch the child take the test. The second group of answers referred to a problem with the tests themselves, that they don't examine process skills or creativity, but limit the questions to facts. There were 5(19.2%) of the responses which mentioned this concern. The next two groups of responses reflected concerns regarding the limited value group tests had for classrooms regarding (a) strategies for use by teachers, (b) the temptation to teach to the test and (c) the fact that teachers may treat children with low IQ's differently due to low expectations. There were 4(15.4%) of the responses which mentioned these concerns. The fourth grouping had the same number of distinct responses as the third group. Here, the category involved technical issues such as problems with reliability, validity, norms and standardization procedures. The fifth category referred to a lack of diagnostic value such as the inability to acquire individual differences about strengths and weaknesses. There were 3(11.5%) of the responses which mentioned these concerns. The sixth category related to the inappropriateness of mass screening and labelling. There were 2(7.7%) of the responses which cited these concerns. The seventh and eighth categories were both mentioned 1(3.8%) times each. The

seventh group specified that group tests were not good predictors of school performance, and the eighth stated that the tests were susceptible to outside influences such as nervousness of the child in a strange situation which in turn could affect the validity of the measure.

Individual Tests

For the answers to this part of the question, regarding the limitations of individual intelligence tests used with children, there were 21 dominant themes which emerged. There were 65(95.6%) respondents with a total of 121 distinct responses. Table 16 lists the categories, the number of distinct responses for each category and their percentages.

Table 16

Limitations of Intelligence Tests (Individual Tests)

Category	# of Responses	%
Doesn't Provide		
Enough Information	22	18.2%
Testing Poses Cultural		
Problems for ESL Children	14	11.6%
Other Variables		
Affect Testing	12	9.9%

Table 16 Continued

Category	# of Responses	%
Others Don't Understand		
Limitations	9	7.4%
Testing Can Be		
Unsympathetic to Child	8	6.6%
Not Related to Academic		
Performance	7	5.8%
Definition of "IQ" or		
"Intelligence" Poses Problems	6	5.0%
Testing Must Be Done		
Cautiously	5	4.1%
Labelling is Seen as		
Permanent Description	5	4.1%
Misuse of Scores	4	3.3%
Negative for Individual		
Tested	4	3.3%
Not Sure What Some		
Subtests Measure	4	3.3%
Report Writing is		
Time-Consuming	4	3.3%

Table 16 Continued

Category	# of Responses	%
Inadequate Measure for Low Functioning Individuals	3	2.5%
IQ Tests Need to Improve Norming Population	3	2.5%
Screening	3	2.5%
Testing Can't Generalize to Other Aspects	3	2.5%
Labelling Can Be Helpful	2	1.7%
Children Become Unjustly Treated	1	0.8%
Results of Testing Show How Child Functions at That Time	1	0.8%
Testing Can Be Helpful	1	0.8%
Total	121	100%

There were 22(18.2%) of the responses which stated that intelligence tests don't provide enough information to get a total view of the student. For example, 7(31.8%) of the 22 responses

mentioned the fact that these tests only measure "intelligence" as narrowly defined and ignored Gardner's other intelligences. There were 6(27.3%) other responses which cited a need for other assessment instruments such as classroom observation, interviews with the teacher and student to obtain a "whole" view of a child.

The second category concerned the cultural problems for ESL children that testing poses. There were 14(11.6%) of the responses which indicated that there is cultural discrimination with respect to intelligence tests. Specifically, there were 6/14(42.9%) of the responses in this category which spoke of the difficulty of obtaining reliable and valid results with First Nations students. The other 8 (57.1%) comments referred to general cultural bias in testing by language and culture.

There were 12(9.9%) of the responses in the third category which mentioned that there were many variables which affect testing. All 12 responses cited anxiety as an issue which may affect the validity of the results.

The fourth category stated that people don't understand the limitations that IQ testing presents. There were 9(7.4%) of the responses which commented in this manner. Of those nine comments, there were 8(88.9%) which specifically mentioned that the results could be misinterpreted. Of the eight, there were 3 (37.5%) which mentioned misinterpretation of scores and the low expectations one might have of students who received low scores.

The fifth category said that testing could be unsympathetic to the child because of two things: (a) the types of questions asked and (b) the format in which the test is given. Out of a total of 8 distinct responses, there were 4(50.0%) of the responses in this category which mentioned that there were questionable items in the tests, which had the potential to make the child feel "dumb". As well, the tests utilized oral responses whereas some children would do better with written answers. There were 3/8 (37.5%) of the responses in this category which mentioned the format of the tests themselves; that validity was compromised because of the standardized nature of the instrument. For example, an examiner is not supposed to ask why or how a child answered in a certain way or is unable to accomodate a child's language needs without ruining validity. There was 1/8 (12.5%) of the responses which mentioned that results used for placement purposes were not used with the proper objective in mind.

There were 7/121 (5.8%) of the total responses which said that intelligence testing was not related to academic performance. This was the sixth category. Of those 7 responses, there were 6 (85.7%) which reported that the critical area of performance was in a classroom environment and that a direct connection to the classroom and specific subject areas was not always apparent. The other response referred to the fact that tests measured traits that were not easily modified via training such as Feuerstein's LPAD.

The eighth category had 5/121 (4.1%) of the responses. Here, concerns were that the definition of "IQ" or "intelligence" posed problems. There were 3/5 (60.0%) of the comments in this category which revolved around the definition as being limiting; to think we could truly measure intelligence and /or communicate to parents and teachers the "intelligence" of the child. The last two comments discussed the terminology of IQ. One comment referred to the "semantic baggage" the term "IQ" carried and another comment cited that "IQ" implied something static or an unchanging trait.

The seventh category dictated that testing must be done cautiously. There were 6/121 (5.0%) distinct responses in this category. Of those 6, there were 3 (50%) which asserted that too many psychologists were presumptuous and dogmatic in test usage and interpretation and made claims that were very weak. The other 3 (50%) responses supported the need to use testing cautiously, with wisdom and judgement and to remember that it is an artificial situation.

The ninth category concerned labelling as a permanent description of a student. There were 5/121 (4.1%) distinct responses for this category. All five responses mentioned that the results were often interpreted as being permanent indicators or fixed entities.

The tenth category was similar to the ninth as it pertains to labelling but the focus was different. Here, there were 4/121 (3.3%) responses which referred to the misuse of scores. The focus in this

section was that people got number bound, that people focused on the IQ figure, labelled children from those numbers and didn't provide a service because of assumed limitations.

The eleventh grouping had 4/121 (3.3%) responses and revolved around the fact that testing was negative for the individual tested. The main focus here was that testing was done just to find out a "number", with no purpose for assessment. There were 2/4 (50%) which followed this thinking. The other 2/4 (50%) responses referred to the fact that poor results could really impact an individual's perception of self, that testing might be harmful not helpful.

Category twelve which had 4/121 (3.3%) responses spoke of subtests. All four responses mentioned difficulty with interpreting subtests because it was not certain what some subtests measure. The validity and reliability of individual subtests was in question unless the deviation was quite large.

The thirteenth category with 4/121 (3.3%) responses stated that report writing was time-consuming. All four responses were consistent in message.

The fourteenth group with 3/121 (2.5%) responses cited testing as an inadequate measure for low functioning individuals. All three responses were similar.

The fifteenth category referred to the need for IQ tests to improve the norming population. There were 3/121 (2.5%)

responses and all spoke for the need for intelligence tests to broaden their range of items and include more people with disabilities in trial populations. There was 1 specific mention of the Wechsler Scales being normed on an American population base, not a Canadian one.

The sixteenth category mentioned the fact that IQ tests were screening measures and were only a starting point, they didn't provide definitive answers. All 3/121 (2.5%) responses were similar.

The seventeenth category mentioned that testing couldn't generalize to other aspects, that intelligence tests were very specific in what they measured and did not measure all aspects of intelligence. There were 3/121 (2.5%) responses in this section and all were similar.

The eighteenth category was a surprising one; that labelling could be helpful. There were 2/121 (1.7%) of the responses, yet each was distinct. One response mentioned that you could have a clearer picture of what you're dealing with, and the other response mentioned that a label could be taken away and used to the benefit of the child.

The nineteenth, twentieth and twenty - first categories all had 1/121 (.8%) responses in them. The 19th refers to children becoming unjustly treated through misperceptions of teachers or parents; the twentieth mentioned that the results of testing showed how the child functioned at that time and the 21st mentioned that testing could be

helpful as an intellectual assessment and could provide an explanation when no progress had been made by the child.

Question #7. Do you have any other comments you would like to make about intelligence testing with children?

There were 39/68 (57.4%) respondents to this question with 49 distinct responses grouped into 9 categories. Table 17 lists the categories with the number of distinct responses for each group and their corresponding percentages. (Actual responses may be found in Appendix N).

Table 17

Other Comments About Intelligence Tests

Category	# of responses	%
Psychologists Must Use		
Tests Responsibly	13	26.5%
Intelligence Tests Have		
Their Place	12	24.5%
Assessment Has Many Parts	11	22.4%
Intelligence Tests Can be Misused	4	8.2%
Intelligence Tests Can be Ineffective	3	6.1%
Validity of Results is Affected		
by Many Factors	2	4.1%

Table 17 continued

Category	# of responses	%
The Public Needs to Be Informed		
of Advantages and Disadvantages	2	4.1%
Clinicians are Part of a School Team	1	2.0%
Intelligence Tests Have a Low Priority	1	2.0%
Total	49	100%

The first category said that psychologists must use tests responsibly. There were 13/49 (26.5%) of the responses which spoke of this topic. For example, two representative comments said that "the responsibility for what happens in intelligence testing resides almost solely with the tester," and "I always remember my report will be following this child throughout their school career and will be interpreted by many others".

The second category with 12/49 (24.5%) distinct responses referred to the fact that intelligence tests do have their place. There were 5/12 (41.7%) responses which cited intelligence testing as an efficient way to gain intellectual/behavioral and sometimes emotional information about children. Therefore, data other than IQ was available through testing. There were 2/12 (16.7%) responses which mentioned the helpfulness of IQ tests in making general

programming decisions about children. There were 1/12 (8.3%) responses which spoke of de-emphasizing actual IQ scores and focusing on strengths and weaknesses and 4/12 (33.3%) which mentioned general comments such as "can be useful if used appropriately and positively".

The third category referred to the fact that assessment had many parts to it. There were 11/49 (22.4%) responses which focused on this topic. All responses in this section referred to the importance of a wholistic assessment which included observations in class and a review of classwork and interviews with teachers/parents/student.

The fourth category said that intelligence tests could be misused. There were 4/49(8.2%) distinct responses here. All responses spoke to this topic and mentioned ways in which tests were misused. Two representative comments were that IQ tests had been "oversold" and "overused" by psychologists and that psychologists may be partially blamed for encouraging the mystique regarding IQ testing. Another comment referred to the need to follow through after testing beyond the reporting of the results. A fourth said that testing should not be used for entry into specific programs.

The fifth category had 3/49 (6.1%) responses and said that intelligence tests could be ineffective because they were of no value unless they improved the situation for the child. A second comment pointed to the fact that clinicians must be aware of the limitations of

tests and the third stated that in an ideal world, we wouldn't bother testing.

The sixth category concerned the results of testing which might be affected by many factors. There were 2/49 (4.1%) responses here. Both comments cited influences such as auditory comprehension and low self-esteem. As well, it was specified that the test is one measure on one day.

The seventh category with 2/49 (4.1%) responses mentioned that the public needed to be informed of the advantages and disadvantages of testing, as well as the way that information was provided to parents.

The eighth and ninth categories each had 1/49 (2.0%) responses. The eighth referred to clinicians working as part of a school based team not as "outside experts". The ninth referred to the fact that intelligence tests were low on the priority list for psychologists.

CHAPTER V

DISCUSSION

This study was an investigation involving school psychologists from the Manitoba Association of School Psychologists and the Child Guidance Clinic, to discern their views regarding the purposes, strengths and the limitations of intelligence testing with children. The responses helped to illuminate the perspective from which Manitoba psychologists regard the use of intelligence tests. They also aided in recording the specific tests in practice as well as the extent to which they were used.

In the demographic section, a question was asked whether the respondent held a teacher's certificate. Table 5 (in the results section) shows that only 26.5% did and that 73.5% didn't. This result is disturbing because psychologists who have no experience in the classroom (where children spend their time during the school day), are making decisions regarding the students they see. They need to understand that students may behave differently in the classroom from the office in which they might see them. Understanding how a child copes within a classroom setting is very important and without that knowledge, the assessment would be incomplete.

In question #1, the respondents were asked if they used group intelligence tests, which ones they used, and for what purposes they used them with children. Responses to this question were interesting. There were no respondents who said they used group

intelligence tests. However, in question #1a, one respondent mentioned using the DIAL-R for screening purposes, yet, according to Sweetland and Keyser (1984), that test is not for group use. For question #1b, one respondent noted that group tests might be used for screening purposes for special needs students or admission to private schools. Anastasi (1982), Salvia and Ysseldyke (1985) and Wright and Piersell (1987), stated several criticisms regarding group administered instruments such as: standardization samples, reliability, predictive validity, lack of opportunity to establish rapport and conditions such as anxiety which could affect scores.

From the literature, one can see that group tests aren't highly thought of and the respondents have indicated that are not used. Thus, one may conclude that the respondents have read the literature and chose not to use group instruments for the above reasons. Use of group administered instruments for screening purposes has been described as a concern by Hartsough, Elias and Wheeler (1983).

Question #2 had three parts as well: if individual intelligence tests were used by the respondents, which individual tests were used, and for what purposes individual intelligence tests were used with children.

Individual Tests Used

The majority of respondents, 98.5%, said they used individual intelligence tests. As well, there was a variety of intelligence tests

listed in Table 7, including individual, group tests and tests which could be used with special populations such as the Hiskey-Nebraska Tests of Learning Aptitude for use with deaf or hard-of-hearing children. It is interesting to note that some group inventories were included in a list which specifically asked for individual tests. Why this was done is a mystery. Perhaps they use group instruments and believe they are individual tests. These data suggested that psychologists obtained information about a child through the use of various inventories. This variety in testing could also be seen in three previous studies of psychological assessment: Brown and McGuire (1976), Goh, Teslow and Fuller (1981) and Hutton, Dubes and Muir (1992). All three studies listed a wide range of tests used by school psychologists and all had Wechsler tests listed and the most used test. The present study indicated that intelligence testing continues to be dominated by the use of the Wechsler scales, however, other assessment procedures were used as well.

Purposes of Using Individual Tests

This study identified 16 purposes for using individual intelligence tests. The purposes most often mentioned was for the identification of learning potential. There were interesting inconsistencies within the psychologists' answers when referring to Table 8 though, as 25% said that the purpose most often given for using intelligence tests was for the identification of learning potential and 11% said that they used them for placement in special programs.

However, in Table 14, when asked to list strengths of intelligence tests, only 3.7% pointed to the fact that the identification of learning potential was a strength of intelligence tests and a reason to use them. Therefore, they do not see the identification of learning potential as a strength of the test and in Table 8 they said that they used them for that very reason.

The data suggested that intelligence tests were used less often for funding, to help others understand the child and for identification of gifted students, mental retardation and syndromes.

For the most part then, intelligence tests were used to for academic planning, to identify those who would do well in school. This is however, a poor reason for using intelligence tests because one must remember that all academic subjects are not covered in intelligence tests (Sternberg, 1992; Wexler - Sherman, Gardner & Feldman, 1988), and that the predictive validity of intelligence tests has been called into question by Gardner (1984) and Wiggins (1989). Tests of this nature omit a relationship with real-world problems (Sigel, 1974; Hatch & Gardner, 1990), and foster a decontextualized nature during testing situations (Hatch & Gardner, 1990; Mcloughlin, 1985; Travers, Beck & Bissell, 1982).

More importantly, "standardized measures seldom indicate exactly *what* or *how* a child should be taught" (Smith, 1980). This is important to note as a test must not be used just to test. It should

lead somewhere. There should be follow - up that is useful for the psychologist, the teachers and most importantly, the student.

It is noteworthy that psychologists use intelligence tests in a limited way to identify those with mental disabilities. Cole (1981) and Garner (1985) speak to this practice. Cole stated that the identification will not necessarily ensure the education needed, and Garner emphatically stated that individuals labelled handicapped might be penalized on tests of intelligence.

Intelligence tests were mentioned by 8.3% of the respondents as being used to identify emotional difficulties. This is a low percentage but a significant one because the tests (according to the data) aren't used here for academic purposes. This leads the researcher to posit that perhaps the manner in which one responds during testing including the posture and tone of voice might indicate potential problems. Thus the test itself may be merely the vehicle by which the psychologist might deduce the emotional health of the client.

Question #3 asked: If you do not use intelligence tests with children, please explain why. For this question there were only 8 (11.9%) respondents. As stated in the results section, there were 67 (98%) of the respondents who said that they did use intelligence tests, which posed a statistical discrepancy. There seemed to be a problem with the questionnaire itself where the layout of the questions could have been clearer.

There were seven different reasons given that psychologists gave for not using intelligence tests. There were three of those eight responses which focused on not using tests with some clientele: such as those who were 16 - 18 years of age; not using them with all children referred and those of a young age or of another culture. These responses suggested that intelligence tests were not used routinely by some psychologists; that they were selectively used, and that if not warranted, other procedures were utilized. Specifically, it was noted that for very young children, or children of another culture, intelligence tests weren't used. This finding was an encouraging one as testing with young children could pose validity concerns, because of their manual dexterity and language development. Thus timed tests would be affected the most. As well, using conventional tests with those of another culture has been discussed considerably in the literature (Feuerstein, 1979; Gardner, 1983, 1991; Mercer, 1972, 1973, 1977; Mercer & Brown, 1973).

The second category, programming reasons, discussed the fact that tests are not used because programming decisions didn't require that level of detail, and the fact that the focus was on programming issues rather than scores. These two reasons again reflected the fact that psychologists did not view intelligence tests as the most important procedure.

The third category mentioned that many students were referred for emotional-behavioral problems rather than intellectual

ones. This reason tended to suggest that referrals were for non-academic concerns. Therefore, the focus in the classroom would be on the emotional and behavioral health of students.

The last four categories focused on (a) the preference for authentic assessment, (b) the lack of time to do intelligence tests, (c) the fact that previous test results were already available, and (d) that intelligence tests weren't the only instruments used. The two middle reasons here seem related, as the use of previous test results could be due to a lack of time.

The first and last categories here seem to be related. Authentic assessments are those which would be conducted within the classroom setting in context with the work normally conducted in that setting. The last category mentioned there were other instruments which could be used, which suggested that psychologists utilized procedures and instruments other than intelligence tests. Although there was only one response of this nature, it explained somewhat the answers to question four, where there were many other tests mentioned, as well as alternate assessment procedures. The conclusion that can be drawn from this is that psychologists realize the limitations of intelligence tests and so recognize the value of assessment within the classroom. Their conclusions regarding a student would be more valid if a thorough assessment with the child were done.

Question #4 asked: If you do not use intelligence tests with children, what other assessment procedures do you use? This question produced highly interesting data as the question specifically asked for alternate assessment procedures. Out of 18 (26.9%) respondents to this question, there were 8 (44%) who mentioned using other tests. Added to this incongruity, there were 4 of those tests which were intelligence tests.

Nonetheless, the tests mentioned most were fine motor tests such as the Bender Gestalt and the Beery. There were 3 different types of achievement tests used as well. The bulk of the tests used were those which related to behavior and personality. The mention of these tests complemented the results of question #3 where the number of psychologists who received referrals for the emotional/behavioral problems of students was mentioned as a reason why intelligence tests weren't used exclusively.

Psychologists mentioned using 12 different alternate assessment procedures. Two procedures mentioned at the top of the list were (a) the use of developmental/social tests and (b) classroom observation. There seemed to be a dichotomy here again, as the use of tests, albeit developmental and social ones, were used as alternate assessment procedures. Classroom observation, as well as the other procedures, mentioned could fall into a collective group described as comprehensive assessment, as it included much more information than any single test could provide.

Since there were more respondents (72.7%) who mentioned alternative assessment procedures than tests, (44.4%), this suggests that Manitoba psychologists are attempting to address the limitations of intelligence tests by using a variety of tests and procedures as well as focusing on the manner of testing which includes observations of the child during the testing as well as within - classroom assessment with work samples. Alternative assessment, that which is performance - based, is gaining in popularity and has been written about extensively. Gardner (1983, 1984, 1987, 1991, 1993); Hatch and Gardner (1988, 1990); Smith (1980); Wexler-Sherman, Gardner and Feldman (1988); Ysseldyke, Algozzine and Thurlow (1992) all discuss authentic assessment which includes performance - based tasks and problems solving ability in real - world situations.

In this question, the answers confirmed what had been said already; that intelligence tests were not the only instruments or procedures used, and that a wholistic approach gave more information. The reason for using other assessment procedures, mentioned, 45 % of the time, was a need to get a complete picture of a student. This finding was significant because it suggested that psychologists needed to see beyond the numbers and scores on tests to truly understand a student. They needed to use a wholistic methodology whereby students were appreciated from a wide variety of perspectives. Gardner (1983, 1984, 1987, 1991, 1993) wrote regarding the importance of assessment in context, that is, within the

venue in which students learn and interact. Planning and programming is the next step, and may be accomplished through a thorough understanding of the student. The technical reasons confirm this, as true measures are understood as a kin to authentic measures and reliability and validity are seen as measures gained through alternate assessments. As well, normed tests gave clues which needed confirming by authentic assessments. Interviews with parents, teachers, and child were also seen as important.

Through all this, one may see that assessment is seen as multidimensional, where scores are ignored, and the process of gaining information was endorsed.

Question #5 asked: What do you perceive to be the strengths or positive aspects of the intelligence tests you use with children?

Group Tests

Results concerning the strengths of group intelligence tests were minimal. The respondents who commented on this aspect did so with many reservations as to their use of group inventories. There were only 5(38.5%) of the distinct responses which talked about positive aspects of uses. The other responses were decidedly negative. Since the question specifically asked for positive aspects, this result is interesting. It leads the researcher to believe that perhaps there are few positive things to say in regards to group intelligence tests. The literature is scarce in mentioning positive aspects as well. One comment, which is an element of both the

responses in the survey and with the literature, related to the use of group instruments to tests large numbers of people. If this is an advantage, it is a slim one. A lack of rapport with students and poor validity were issues mentioned.

Individual Tests

In contrast, the responses to this section of question #5 represented 67/68 (98.5%) respondents with 188 distinct responses, all of which were positive. Manitoba psychologists perceived strengths of intelligence tests which were relative to the usefulness of them relating to: (a) teacher assistance, (b) technical adequacy and (c) process, (or the manner in which the tests were conducted).

Strengths mentioned in the literature included: the concept of intelligence, culture fairness tests and the advantages of group and individual tests. Respondents to the present study indicated a positive awareness regarding the strengths specified above. Statements regarding a total package of understanding the student including behavioral aspects were mentioned often. The ability to establish rapport with the student to get a sense of behavior, attitude and effort was specified as an opportunity for structured observation of the student. Anastasi (1988) suggested that the abilities and behaviors noticed during testing were those which were most likely to be present during school learning and an effective predictor of performance in other activities of daily life.

Further, the usefulness of these tests for teacher assistance, in forming diagnostic impressions, and to help establish reasonable expectations of the school and parents were issues involving the recommended uses of intelligence tests. As well, technical adequacy comments were extensive leaving no doubt as to the belief by the respondents that the tests they chose to use had value in their internal structure. Standardization of the tests, normative data, predictability, organization and educational importance were issues at the forefront within this area. The process of conducting tests of this sort underscored the qualitative information one may obtain during application.

Question #6 asked: What do you perceive to be the limitations or negative aspects of the intelligence tests you use with children?

Group Tests

The remarks were fairly decisive in that they referred to three main issues: 1. the inability to observe students; 2. the uselessness for educators and 3. technical concerns. This information suggested that psychologists valued the ability to observe students during tests, as this was the response most often mentioned. A complementary point centered on the process of completing the test to give clues as to thought processes. Other responses indicated concerns regarding a lack of diagnostic value, an inappropriateness of mass screening and labelling, as well as susceptibility to unknown outside influences which could affect the outcome. Limitations of group tests discussed

in the literature communicate similar concerns. Thus, respondents to this study indicated an understanding of those limitations. Limited value for the classroom, lack of opportunity for observation during testing and norms and standardization procedures were comments from the respondents. Lack of rapport, standardization and technical adequacy were issues mentioned in the literature.

Individual Tests

Manitoba psychologists perceived issues such as inappropriate conclusions and usage as limitations affecting intelligence testing. Limitations of individual intelligence tests mentioned in the literature specified the concept of intelligence, the limited nature of the questions (specifically the lack of inquiry into mechanical, musical and artistic aptitudes), cultural fairness, external influences which could affect the results of testing, and the technical adequacy of some instruments.

Respondents to the present study indicated an understanding of those limitations. Major issues mentioned were interpretation and diagnostic limitations, technical and cultural concerns (specifically First Nations students), definitions of intelligence, and information regarding cautious use of these tests.

A comparison of these two lists yields two conclusions. First, Manitoba psychologists recognized the inadequacies inherent in using intelligence tests. Second, they made it very clear that inappropriate usage and conclusions affect the student, the school

and the parents. Various influences such as fatigue and anxiety may affect the outcome. Schools may place too much emphasis on the results of one test and parents may be upset about a low score. Psychologists advised caution and recommended the use of data obtained from other sources in conjunction with the information from tests.

Cultural concerns were also mentioned as a factor in limiting the usefulness of tests as well as the definition of intelligence and the term "IQ". Comments mentioning cultural bias and the "semantic baggage" of IQ and its implications of a static individual trait determining future performance were also cause for concern.

Question #7 asked: Do you have any other comments you would like to make about intelligence testing with children?

Manitoba psychologists recommended caution when utilizing intelligence tests and advocate that it is the responsibility of the psychologist to use the test wisely, to use it as part of a process, and to inform parents carefully of the outcome of the testing procedure. These statements reflected conservative employment of intelligence tests, where the examiner was scrupulous, compassionate and viewed assessment as a compilation of information from many sources.

Psychologists made it known that they were very aware of the many factors such as low self-esteem and auditory comprehension

which can affect tests results. Added to that, is the understanding that an intelligence test is one measure on one day.

There were many strengths and weaknesses, approaches to testing and cautions described by psychologists which summarized their opinions and judgement regarding intelligence tests. They recommended use in an appropriate fashion with testing playing a small part in drawing out pertinent information about students and an awareness of the limitations, including unrealistic expectations of school and parents. One comment seemed to summarize the entire feeling surrounding intelligence testing: "I always remember my report will be following this child throughout their school career and will be interpreted by many others.

CONCLUSION

Packaged tests are still being used to assess students. Although more of an authentic assessment approach is being conducted, where psychologists obtained information about a child through the use of various inventories, tests are still used as a determining factor. Specifically, intelligence testing continues to be dominated by the use of the Wechsler Scales.

Purposes for using intelligence tests related to learning, to the capabilities of students, the manner in which they learn and the amount of learning achieved. For the most part then, intelligence tests were used to assess the academic characteristics of students. However, intelligence tests were not used routinely by some

psychologists; they were used selectively, and if not warranted, other procedures were utilized. Therefore, psychologists did not view intelligence tests as the most important procedure. This implies that psychologists are realizing the limitations of the sole use of intelligence tests and are employing various methodologies such as authentic assessment as part of their study of a child.

Since many students have been referred for emotional and behavioral problems, some assessment concerns were not academic. The focus in the classroom would be on the emotional and behavioral health of students. Further, Manitoba psychologists are attempting to address the limitations of intelligence tests by using a variety of tests and procedures as well as observations of the student within the classroom. Assessment is seen as multidimensional, where the process of gaining information is endorsed.

There were few positive things to say in regards to group intelligence tests and the opposite was true for individual tests where observation of the student was possible.

There seemed to be a pattern within the answers from all respondents; 1. caution in using the tests and 2. wholistic assessment. Psychologists use intelligence tests as part of the process of assessment, not as an end in themselves. They focused on strengths and weaknesses and used a variety of assessment procedures and used the tests as a tool to understand and help children, parents and educators.

Limitations

Those who did use intelligence tests for policy reasons, but disliked their use, may or may not have chosen to respond. It is important to know why some psychologists chose not to use intelligence tests in their practice. For this reason, specific questions in the survey addressed this issue. However, psychologists may have chosen not to respond to these questions. Therefore, the possibility of obtaining a biased information either promoting or discouraging the use of intelligence tests might have resulted. The return rate was poor thus there is not as much confidence about the results as there would have been had the rate been higher. "People who have a particular interest in the subject matter or the research itself are more likely to return mail questionnaires than those who are less interested" (Fowler, Jr., 1993, p. 41). Thus generalization to the whole population of psychologists in Manitoba is limited. As well, because of the return rate, the researcher may have low confidence regarding the results.

Further limitations of this study revolve around the fact that the study was confined to psychologists in Manitoba only. Further investigation in other provinces would give a more global view across Canada as policies might differ provincially regarding the importance given to tests of this nature.

RECOMMENDATIONS

Recommendations can be made regarding the direction of future studies of psychological assessment practices. Future investigations should focus on four things. First, it would be important to understand the perceptions that Canadian teachers have of intelligence tests and their use, as it was noted in the responses that some teachers focus on the scores. This is a problem because the scores themselves have limited value. What psychologists are attempting to address is the process of assessment whereby the manner in which the child responds during the test and within the classroom is important. Second, future research comparing the amount of time spent testing, tests used and the use of other assessment procedures with the present results would advance knowledge regarding assessment practices here in Manitoba. Third, it would be enlightening to study whether psychologists utilize the test itself as a vehicle by which to deduce the emotional health of the client as the results seem to suggest. Fourth, it would be useful to gather data to permit comparisons among other regions in Canada.

References

- Anastasi, A. (1982). Psychological Testing (5th ed.). New York: Macmillan.
- Anastasi, A. (1988). Psychological Testing (6th ed.). New York: Macmillan
- Anderson, P. L., Cronin, M. E. & Kazmierski, S. (1989). WISC-R stability and re-evaluation of learning-disabled students. Journal of Clinical Psychology, 45, 941-944.
- Bannatyne, A. (1974). Diagnosis: A note of recategorization of the WISC scaled scores. Journal of Learning Disabilities, 7, 272-274.
- Bauman, E. (1991). Determinants of WISC-R subtest stability in children with learning difficulties. Journal of Clinical Psychology, 47, 430-435.
- Bornstein, A. V. (1968). The effects of examiner approval and disapproval upon the performance of subjects on the performance scale of the Wechsler Intelligence Scale for Children. Dissertation Abstracts, 28, 3047 - 3048.
- Bridgeman, B. , Strang, H. R. , & Buttram, J. (1974). "Game" versus "test" instructions for the WISC. Journal of Educational Measurement, 11, 285 - 288.
- Brown, W. R. & McGuire, J. M. (1976). Current psychological assessment practices. Professional Psychology, 7, 475-484.
- Carroll, J. B. & Horn, J. L. (1981). On the scientific basis of ability testing. American Psychologist, 36, 1012-1020.

- Cattell, R. B. (1950). Culture Fair Intelligence Test: Scale 1.
Champaign, IL: Institute for Personality and Ability Testing.
- Cattell, R. B. (1962). Handbook for Culture Fair Intelligence Test: Scale 1. Champaign, IL: Institute for Personality and Ability Testing.
- Cattell, R. B. & Cattell, A. K. S. (1960). Culture Fair Intelligence Test: Scale 2. Champaign, IL: Institute for Personality and Ability Testing.
- Cattell, R. B. & Cattell, A. K. S. (1963). Culture Fair Intelligence Test: Scale 3. Champaign, IL: Institute for Personality and Ability Testing.
- Cole, N. S. (1981). Bias in testing. American Psychologist, 36, 1067-1077.
- Dean, R. S. (1977). Patterns of emotional disturbance on the WISC-R. Journal of Clinical Psychology, 33, 486-490.
- Elliott, S. N. & Boeve, K. (1987). An investigation of ethnic differences over time. Educational and Psychological Measurement, 47, 461-465.
- Exner, J. E. (1966). Variations in WISC performance as influenced by differences in pretest rapport. Journal of Genetic Psychology, 74, 299 - 305.
- Feuerstein, R. (1979). The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments and techniques. Baltimore: University Park Press.

- Fox, L. H. (1981). Identification of the academically gifted. American Psychologist, 36 (10), 1103 - 1111.
- Fowler, F. J. , Jr. (1993). Survey research methods. (2nd ed.). Newbury Park, California: Sage Publications.
- Freides, D. (1978). Review of WISC-R. In O. K. Buros (Ed.), The Eighth Mental Measurements Yearbook. Highland Park, NJ: The Gryphon Press.
- Galdieri, A. A. , Barcikowski, R. S. , & Witmer, J. M. (1972). The effect of verbal approval upon the performance of middle- and lower-class third-grade children on the WISC. Psychology in the Schools, 9, 404 - 408.
- Gardner, H. (1983). Frames of mind: The theory of multiple intelligences. New York: Basic Books.
- Gardner, H. (1984). Assessing intelligences: A comment on 'testing intelligence without i.q. tests'. Phi Delta Kappan, 65 (10) , p. 699-700.
- Gardner, H. (1987). Beyond the iq: Education and human development. Harvard Educational Review, 57 (2), p. 187-195.
- Gardner, H. (1987a). An individual-centered curriculum. In The schools we've got, the schools we need. Washington, DC: American Association of Colleges of Teacher Education.
- Gardner, H. (1991). The unschooled mind. New York: Basic Books.
- Gardner, H. (1993). Multiple intelligences: The theory in practice. New York: Basic Books.

- Garner, J. B. (1985). Intelligence testing or testing intelligently, School Psychology International, 6, 235-238.
- Gay, L. R. (1981). Educational research: Competencies for analysis and application. (2nd ed.). Columbus: Charles E. Merrill.
- Goh, D. S. , Teslow, C. J. & Fuller, G. B. (1981). The practice of psychological assessment among school psychologists, Professional Psychology, 12, 696-706.
- Hale, R. L. & Saxe, J. E. (1983). Profile analysis of the Wechsler Intelligence Scale for Children-Revised. Journal of Psychoeducational Assessment, 1, 155-161.
- Haney, W. (1981). Validity, vaudville, and values. American Psychologist, 36 (10), 1021 - 1034.
- Harris, D. (1963). Children's drawings as measures of intellectual maturity. Orlando, FL: Harcourt Brace Jovanovich.
- Hartsough, C. S. , Elias, P. , & Wheeler, P. (1983). Evaluation of a non-intellectual assessment procedure for early screening of exceptionality. Journal of School Psychology, 21, 133-142.
- Hatch, T. & Gardner, H. (1988). New research on intelligence. Learning, 17 (4), p. 36-39.
- Hatch, T. & Gardner, H. (1990). If Binet had looked beyond the classroom: The assessment of multiple intelligences. International Journal of Educational Research, 14(5), p. 415-429.

- Hutton, J. B. , Dubes, R. & Muir, S. (1992). Assessment practices of psychologists: Ten years later. School Psychology Review, 21, 271-284.
- Jensen, A. R. (1980). Bias in mental testing. New York: Macmillan.
- Kamphaus R. W. & Platt, L. O. (1992). Subtest specificities for the WISC-III. Psychological Reports, 70, 899-902.
- Kaufman, A. S. (1979). Intelligent testing with the WISC-R. New York: Wiley-Interscience.
- Kaufman, A. S. (1992). Evaluation of the WISC-III and WPPSI-R for gifted children. Roeper Review, 14(3), 154-158.
- Kaufman, A. S. , & Kaufman, N. L. (1983). Kaufman Assessment Battery for Children: Administration and scoring manual. Circle Pines. MN: American Guidance Service.
- Kramer, J. H. (1993). Interpretation of individual subtest scores on the WISC-III. Psychological Assessment, 5, 193-196.
- Lubin, B. Wallis, R. R. & Paine, C. (1971). Patterns of psychological test usage in the United States: 1935-1969. Professional Psychology, 2, 70-74.
- McCabe, D. & Hilmo, J. (1985). Pictures speak louder than test scores. Academic Therapy, 20, 333-338.
- McCluskey K. W. & Walker, K. D. (1986). The doubtful gift. Kingston: Ronald P. Frye & Company.

- McLoughlin, C. (1985). Evaluating the ecology of the young child: Internal and external environments. In C. McLoughlin & D. Gullo (Eds.). Young children in context (pp. 277-309). Springfield, IL: Charles Thomas Publishers.
- Mercer, J. R. (1972). I. Q. : The lethal label. Psychology Today, 6, 44-47, 95-97.
- Mercer, J. R. (1973). Labeling the mentally retarded. Los Angeles: California University Press.
- Mercer, J. R. (1974). Latent functions of intelligence testing in the public school. In L. P. Miller (Ed.), The testing of black students: A symposium for the american educational research Association. Englewood Cliffs, NJ: Prentice-Hall.
- Mercer, J. R. (1977). The struggle for children's rights: Critical juncture for school psychology. School Psychology Digest, 6, 4-19.
- Mercer, J. R. , & Brown, W. C. (1973). Racial differences in I. Q. : Fact or artifact? In C. Senna (Ed.). The Fallacy of I. Q. (pp. 56-113).
- Myers, D. G. (1981). Aren't all children gifted? Today's Education, 70, 30 - 33.
- Oakman, S. & Wilson, B. (1988). Stability of WISC-R intelligence tests: Implications for 3 year reevaluations of learning disabled students. Psychology in the Schools, 25, 118-120.
- Oros, J. A. , Johnson, J. J. , & Lewis, M. L. (1972). The effect of induced anxiety on the Wechsler Intelligence Scale for Children. Psychology in the Schools, 9, 388 - 392.

- Otis, A. & Lennon, D. (1982). The Otis-Lennon School Ability Test.
San Antonio, TX: The Psychological Association.
- Reschley, D. J. (1979). Nonbiased assessment. In G. Phye & D.
Reschley (Eds), School psychology: Perspectives and issues.
New York: Academic Press.
- Robinson N. M. & Chamrad, D. L. (1986). Appropriate uses of
intelligence tests with gifted children. Roeper Review, 8, 160-163.
- Rourke, B. P. , Bakker, D. J. , Fisk, J. L. & Strang, J. D. (1983). Child
neuropsychology: An introduction to theory, research and
clinical practice. New York: Guilford Press.
- Ryan, A. M. & Sackett, P. R. (1987). A survey of individual
assessment practices by psychologists. Personnel Psychology, 40,
455-488.
- Salvia, J. & Ysseldyke, J. E. (1982). Assessment in special and
remedial education. Boston: Houghton Mifflin.
- Salvia, J. Ysseldyke, J. E. (1985). Assessment in special education
(3rd ed.). Boston: Houghton Mifflin.
- Salvia, J. & Ysseldyke, J. E. (1988). Assessment in special remedial
education (4th ed.). Boston: Houghton Mifflin
- Schmidt, H. P., Kurliw, A. J., Saklofske, D. H. & Yackulic, R. A.
(1989). Stability of WISC-R scores for a sample of learning
disabled children. Psychological Reports, 64, 195-201.
- Schumacher, S. & McMillan, J. H. (1993). Research in Education
(3rd ed.). New York: Harper Collins.

- Short-DeGraff, M. A. & Holan, S. (1992). Self-drawing as a gauge of perceptual-motor skill. Physical and Occupational Therapy in Pediatrics, 12, 53-68.
- Short-DeGraff, M. A. , Slansky, L. & Diamond, K. E. (1989). Validity of preschoolers self-drawings as an index of human figure drawing performance. Occupational Therapy Journal of Research, 9, 305-315.
- Sigel, I. (1974). When do we know what a child knows? Human Development, 17, 201-217.
- Smith, C. R. (1980). Assessment alternatives: Non-standardized procedures. The School Psychology Review, 9(1), p. 46-57.
- Sternberg, R. J. (1984). Testing intelligence without i.q. tests. Phi Delta Kappan, 65 (10), 694 - 698.
- Sternberg, R. J. (1992). Ability tests, measurements, and markets. Journal of Educational Psychology, 84, 134-140.
- Sweetland, R. C. & Keyser, D. J. (Eds.). (1984). Tests: A comprehensive reference for assessments in psychology, education and business. (Supplement). Kansas City: Tests Corporation of America.
- Thorndike, R. & Hagen, E. (1986). Cognitive abilities test. Chicago, IL: The Riverside Publishing Company.
- Tittle, C. K. (1978). Review of WISC-R. In O. K. Buros (Ed.), The Eighth Mental Measurements Yearbook. Highland Park, NJ: The Gryphon Press.

- Travers, J. R. , Beck, R. , & Bissell, J. (1982). Measuring the outcomes of day care. In J. R. Travers & R. J. Light (Eds.), Learning from experience: Evaluating early childhood demonstration programs (pp. 109-162). Washington, DC: National Academy Press.
- Vernon, P. A. (1984). Wechsler Intelligence Scale for Children - Revised. Test critiques (Vol. 1). Kansas City: Westport.
- Wade T. C. & Baker T. B. (1977): Opinions and use of psychological tests. American Psychologists, 32, p. 874-882.
- Wagner, R. K. , & Sternberg, R. J. (1984). Alternative conceptions of intelligence and their implications for education. Review of Educational Research, 54, 179-223.
- Webster, R. E. (1988). Statistical and individual temporal stability of the WISC-R for cognitively disabled adolescents. Psychology in the Schools, 25, 365-372.
- Wechsler, D. (1974). Manual for the Wechsler Intelligence Scale for Children -- Revised. New York: Psychological Corporation.
- Wechsler, D. (1992). Manual for the Wechsler Intelligence Scale for Children -- Third Edition. New York: Psychological Corporation.
- Wexler - Sherman, C. Gardner, H. & Feldman, D. H. (1988). A pluralistic view of early assessment: The project spectrum approach. Theory into Practice, 27(1), p. 77-83.
- Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70 (9), p. 703-713.

- Witmer, J. M. , Bornstein, A. V. & Dunham, R. M. (1971). The effects of verbal approval and disapproval upon the performance of third and fourth grade children on four subtests of the Wechsler Intelligence Scale for Children. Journal of School Psychology, 9, 347 - 356.
- Wright, D. & Piersel, W. C. (1987). Usefulness of a group-administered ability test for decision making. Journal of School Psychology, 25, 63-71.
- Ysseldyke, J. E. , & Algozzine, B. (1990). Introduction to special education. (2nd ed.). Boston: Houghton Mifflin.
- Ysseldyke, J. E. , Algozzine, B. , Richey, L. & Graden, G. (1982). Declaring students eligible for learning disability services: Why bother with the data? Learning Disability Quarterly, 5, 37-44.
- Ysseldyke, J. E. , Algozzine, B. & Thurlow, M. L. (1992). Critical issues in special education. Boston: Houghton Mifflin.

Appendix A
Permission Letter Mailed to the
Manitoba Association of School Psychologists

April 26, 1995

Dear Mr. Alper:

I am writing to request a list of the names of all who are members of the Manitoba Association of School Psychologists to participate in a research project to study the opinions of Manitoba Psychologists regarding intelligence testing with children.

"Intelligence tests are psychometric devices - in practice, sets of standardized questions and tasks, for assessing an individual's potential for purposeful and useful behavior, at least in those aspects of it which one agrees to designate as intelligent" (Wechsler, 1974). I am carrying out this survey in partial fulfillment of the requirements for the degree of Master of Education in the Faculty of Education at the University of Manitoba.

This survey intends to collect information concerning the following questions:

1. What intelligence tests do Manitoba psychologists use?
2. For what purpose(s) do psychologists in Manitoba use intelligence testing?
3. What do Manitoba psychologists perceive as the strengths or positive aspects of intelligence testing?
4. What do Manitoba psychologists perceive as the limitations or negative aspects of intelligence testing?

In reporting the results, I will be describing the incidence, frequency and distribution of the responses to the questionnaire. The identity of individual respondents will not be revealed and individual responses will remain confidential. All members of the Manitoba Association of School Psychologists will receive a survey.

If you have any questions about the survey, feel free to contact me at 338-7510 (work) or 663-1107 (home). Results of the study will be available by returning a request form in a separate envelope from that used to return the survey. A contact person at the Faculty of Education is my advisor, Dr. Gerry Bravi. He can be reached at 474-9018. Your assistance in this project is greatly appreciated.

Yours truly,

Donna L. Johnson

Appendix B
Permission Letter Mailed to the
Child Guidance Clinic

April 26, 1995

Dear Mr. Kircher:

I am writing to request a list of the names of all who are members of the Child Guidance Clinic to participate in a research project to study the opinions of Manitoba Psychologists regarding intelligence testing with children. "Intelligence tests are psychometric devices - in practice, sets of standardized questions and tasks, for assessing an individual's potential for purposeful and useful behavior, at least in those aspects of it which one agrees to designate as intelligent" (Wechsler, 1974). I am carrying out this survey in partial fulfillment of the requirements for the degree of Master of Education in the Faculty of Education at the University of Manitoba.

This survey intends to collect information concerning the following questions:

1. What intelligence tests do Manitoba psychologists use?
2. For what purpose(s) do psychologists in Manitoba use intelligence testing?
3. What do Manitoba psychologists perceive as the strengths or positive aspects of intelligence testing?
4. What do Manitoba psychologists perceive as the limitations or negative aspects of intelligence testing?

In reporting the results, I will be describing the incidence, frequency and distribution of the responses to the questionnaire. The identity of individual respondents will not be revealed and individual responses will remain confidential. All members of the Child Guidance Clinic will receive a survey.

If you have any questions about the survey, feel free to contact me at 338-7510 (work) or 663-1107 (home). Results of the study will be available by returning a request form in a separate envelope from that used to return the survey. A contact person at the Faculty of Education is my advisor, Dr. Gerry Bravi. He can be reached at 474-9018. Your assistance in this project is greatly appreciated.

Yours truly,

Donna L. Johnson

Appendix C
Letter Mailed to Manitoba Association of School Psychologists
Survey Respondents

May 18, 1995

Dear Colleague:

I am writing to request your participation in a research project which studies the opinions of Manitoba Psychologists regarding intelligence testing with children. "Intelligence tests are psychometric devices - in practice, sets of standardized questions and tasks, for assessing an individual's potential for purposeful and useful behavior, at least in those aspects of it which one agrees to designate as intelligent" (Wechsler, 1974). I am carrying out this survey in partial fulfillment of the requirements for the degree of Master of Education in the Faculty of Education at the University of Manitoba.

This survey will cover the opinions of psychologists concerning which tests Manitoba psychologists use with children as well as the strengths, limitations and purposes for which they use intelligence testing. Although much has been written regarding the use of intelligence tests, we don't know how practicing psychologists view the use of intelligence tests and their strengths and limitations.

In reporting the results, I will be describing the incidence, frequency and distribution of the responses to the questionnaire. The identity of individual respondents will not be revealed and individual responses will remain confidential. All members of the Manitoba Association of School Psychologists will receive a survey. This survey will take approximately one hour to complete and can be returned in the addressed, stamped envelope provided by May 31, 1995. Your cooperation is voluntary.

If you have any questions about the survey, feel free to contact me at 338-7510 (work) or 663-1107 (home). Results of the study can be obtained by returning the enclosed request form in a separate envelope from that used to return the survey. A contact person at the Faculty of Education is my advisor, Dr. Gerry Bravi. He can be reached at 474-9018. Your assistance in this project is greatly appreciated.

Yours truly,

Donna L. Johnson

Appendix D
Letter Mailed to Child Guidance Clinic
Survey Respondents

May 18, 1995

Dear Colleague:

I am writing to request your participation in a research project which studies the opinions of Manitoba Psychologists regarding intelligence testing with children. "Intelligence tests are psychometric devices - in practice, sets of standardized questions and tasks, for assessing an individual's potential for purposeful and useful behavior, at least in those aspects of it which one agrees to designate as intelligent" (Wechsler, 1974). I am carrying out this survey in partial fulfillment of the requirements for the degree of Master of Education in the Faculty of Education at the University of Manitoba.

This survey will cover the opinions of psychologists concerning which tests Manitoba psychologists use with children as well as the strengths, limitations and purposes for which they use intelligence testing. Although much has been written regarding the use of intelligence tests, we don't know how practicing psychologists view the use of intelligence tests and their strengths and limitations.

In reporting the results, I will be describing the incidence, frequency and distribution of the responses to the questionnaire. The identity of individual respondents will not be revealed and individual responses will remain confidential. All members of the Child Guidance Clinic will receive a survey. This survey will take approximately one hour to complete and can be returned in the addressed, stamped envelope provided by May 31, 1995. Your cooperation is voluntary.

If you have any questions about the survey, feel free to contact me at 338-7510 (work) or 663-1107 (home). Results of the study can be obtained by returning the enclosed request form in a separate envelope from that used to return the survey. A contact person at the Faculty of Education is my advisor, Dr. Gerry Bravi. He can be reached at 474-9018. Your assistance in this project is greatly appreciated.

Yours truly,

Donna L. Johnson

Appendix E

Survey with Manitoba Psychologists

Intelligence Tests: A Perspective from Psychologists

PART 1

Demographic Information

1. For how many years have you been a practicing psychologist?

Circle the appropriate selection.

0 - 5 years

6 - 10 years

11 - 15 years

16 - 20 years

21 - 25 years

25 years +

2. Please circle.

Are you : Female

Male

3. What type of psychological practice do you have?

Please circle.

(a) Private practice

(b) I work for an organization (e.g. Hospital, School Division,
etc.)

(c) I work for an organization and have a private practice.

Please specify the organization.

(d) Other (Please specify)

4. What percentage of your work day is spent on testing?

5. What degrees do you hold?

Please circle.

B. Ed.

M.A.

M. Ed.

Ph. D.

5a. Do you hold a teacher's certificate?

Please circle. Yes

No

6. I use intelligence tests with children.

Please circle.

Yes - continue on.

No - skip to Part 2 questions #3.

PART 2

This section asks specific questions about intelligence testing. Please answer as completely as possible.

1. Do you use group intelligence tests with children?

Please circle : Yes

No - skip to # 2

1a. Which group intelligence tests do you use?

1b. For what purposes do you use group intelligence testing with children?

2. Do you use individual intelligence tests with children?

Please circle: Yes

No - skip to #3

2a. Which individual intelligence tests do you use with children?

2b. For what purposes do you use individual intelligence tests with children?

3. If you do not use intelligence tests with children, please explain why.

4. If you do not use intelligence tests with children, what other assessment procedures do you use?

4a. Why do you use other assessment procedures?

5. What do you perceive to be the strengths or positive aspects of the intelligence tests you use with children?

Group Tests _____

Individual Tests _____

6. What do you perceive to be the limitations or negative aspects of the intelligence tests you use with children?

Group Tests _____

Individual Tests _____

7. Do you have any other comments you would like to make about intelligence testing with children?

Thank you again for participating in this study. Your answers have been very helpful.

Appendix F

Reminder Letter to All Psychologists

354 Jacques Avenue
Winnipeg, Manitoba
R3W 1N6
December 6, 1994

Dear Colleague:

I am writing to remind you of the survey you received concerning the intelligence tests you use, and your opinion of the strengths and limitations of intelligence testing as well as the purposes for which you use intelligence testing with children.

If you have returned your completed survey, thank you very much for your time and consideration. If you have not completed and returned your survey, please take the time to comment on this important issue. Another copy of the survey is enclosed with an envelope to be used for reply. Requests for a report of the results can be made by mailing the enclosed Results Request form and sent in a separate envelope from that used for the survey.

Thank you for giving this your attention.

Sincerely,

Donna L. Johnson

Appendix G
Results Request Form

354 Jacques Avenue
Winnipeg, Manitoba
R3W 1N6

Results Request Form

To obtain a copy of the results of the study Intelligence Tests: A
Perspective from Psychologists, please complete and return this form
in the addressed, stamped envelope provided.

(PLEASE PRINT)

Name: _____

Address: _____

Appendix H

Comments from Part 2 of Survey

Question #2b. For what purposes do you use individual intelligence tests with children?

This question represents the first in which the answers from 67 (98.5%) respondents were categorized into dominant themes. There were 109 distinct responses to this question. The themes of responses for this question indicated that the majority of reasons for using individual intelligence tests with children revolve (in order from most predominant reason to the least predominant reason) around:

1. Identification of Learning Potential
2. Placement
3. Assessment of Student Strengths and Weaknesses
4. Program Planning or Program Justification
5. Identification of LD and ADD
6. Identification of Emotional Difficulties
7. Assessment of Intellectual Skills or Processes (Learning Style)
8. Miscellaneous Comments
9. Funding
10. Help Others to Understand or Accept the Child and His or Her Behavior
11. Screening for Indepth Assessment

12. Examination of Modified Program Effects
13. Identification of Gifted
14. Identification of Mental Retardation
15. Identification of Syndromes
16. Neurological Problems

1. Identification of Learning Potential

- to assess whether language delay is related to global deficit.
- when academic performance is lower than would be expected
- to obtain some type of "normed" indication of how they have developed compared to others
- to assess global abilities
- to get a piece of the puzzle forming a respectful picture of a child
- assessing intelligence
- to develop appropriate, reasonable learning expectations
- to get an indication of overall cognitive functioning level
- to assess range of intellectual functioning
- to obtain a rough measure of overall intellectual capacity
- to determine cognitive potential
- compare achievement and IQ
- learning potential/ability
- to assess potential
- to determine intellectual potential
- to identify potential learning ability (predict school performance)
- indicator of their learning potential

- to determine whether child is achieving to potential
- academic aptitude evaluation
- to verify whether the child is working in class at his potential
- to get an estimation of a child's potential for working in a white middle class curriculum
- ability - achievement discrepancies
- ability estimates

Chronic poor school performance in order to rule out low global ability

- ability-achievement discrepancy
- to determine their ability level re: placement
- IQ tests are excellent predictors of academic and social/emotional performance
- high correlation between IQ scores and performance

2. Placement

- assist in placement decisions
- to "classify" child for program eligibility
- to determine special education placement
- assessing children for special academic programs
- to provide info re: placement
- program placement recommendations
- to determine their ability level re: placement
- placement for special needs
- to access special services

- assess special education services
- determination of need for extra school supports (resource, special education)
- to get an estimation of a child's potential for working in a white middle class curriculum

3. Assessment of Student Strengths and Weaknesses

- helping teachers understand students' strengths and weaknesses
 - looking for processing skills and weaknesses
 - to determine strengths and weaknesses
 - to assess specific strengths and weaknesses, talents
 - to help provide explanation for learning difficulties
 - diagnostic
 - to look for unusual patterns of scores suggesting particular areas of strength or weakness
 - to help in understanding learning difficulties
 - identification of factors contributing to learning problems
 - to apply knowledge of cognitive strengths and weaknesses to classroom expectations, programming
 - to supply a profile of strengths and weaknesses in order to plan programming
- ### **4. Program Planning or Program Justification**
- to assist with developing programming plans with the school team
 - to assist school in programming
 - to assist in program planning

- to try to build skills through appropriate programming
- long term educational planning especially when school achievement is not commensurate with previous intelligence tests and/or school staff's and/or parents beliefs about the child's ability
- to apply knowledge of cognitive strengths and weaknesses to programming
- to empirically back up an individual educational plan
- to supply a profile of strengths and weaknesses in order to plan programming
- level of functioning, potential and style of learning for programming, remediation
- program planning

5. Identification of LD and ADD

- secondary insight into ADD sensitive areas
- general and specific L. D.
- find learning disabilities
- possibility of attention deficit
- for diagnosis of ADHD, L. D. (or rather part of the diagnostic process)
- to help diagnose learning disabilities
- as evidence of a learning disability
- to diagnose L. D.
- chronic poor school performance in order to rule out a learning disability

- can show scatter - a sign of L. D.

6. Identification of Emotional Difficulties

- anxiety
- identify cognitive style, risk behavior, frustration tolerance
- to observe child's response to a structured testing situation
- to assess a student's understanding of social cause and effect
- to see how children solve problems, handle frustration (ability to concentrate, persistence)
- as part of emotional assessment
- I use the WISC as a "projective"
- identify area of strength for self-esteem programming
- IQ tests are excellent predictors of social/emotional performance

7. Assessment of Intellectual Skills or Processes (Learning Style)

- comparative look at reasoning skills as part of assessment
- learning style
- to obtain a comparison of visual-motor vs verbal skills
- to assess indications of learning style
- language functioning
- identify cognitive style, risk behavior frustration tolerance
- to observe child's response to a structured testing situation, challenges and various factors integrated into the test e.g. auditory, visual, attentional, processing
- better understanding of learning style

- level of functioning, potential and style of learning for programming, remediation

8. Miscellaneous Comments

- to find out what factors may be underlying "reluctance to write"
- confirmation or rejection of previous hypotheses developed about a child
- to assist older students in making realistic vocational choices
- to assist classroom teacher
- facilitate need for community supports (Child Development Clinic, MATC, family doctor, 1:1 or group therapy, SMD, Rehab Center for Children)
- to develop a relationship (although there are other ways)
- individual intelligence tests are one piece of a Psychological Assessment

9. Funding

- for administrative purposes - mainly to support applications for funding
- occasionally need norm-referenced tests for funding
- to obtain funding for children in special needs - "proof"
- funding for special needs

10. Help Others to Understand or Accept the Child and His or Her Behavior

- showing students with learning disabilities that they are not "dumb" or "stupid" and that they are as good or better than their peers in most areas
- assist teachers and parents to understand/accept the child's individuality
- long term educational planning especially when school achievement is not commensurate with previous intelligence tests and/or school staff's and/or parents beliefs about the child's ability

11. Screening for Indepth Assessment

- used as part of a battery of tests, observations, interviews
- a starting point for determining areas in need of further assessment (e.g. language weaknesses, distractibility, visual-motor difficulties)

12. Examination of Modified Program Effects

- reassessment

13. Identification of Gifted

- to help diagnose giftedness

14. Identification of Mental Retardation

- to help diagnose mental retardation

15. Identification of Syndromes

- tool as part of diagnosis of other difficulties

16. Neurological Problems

- evidence for neurological influences

Appendix I

Question #3. If you do not use intelligence tests with children, please explain why.

There were only 8 respondents (11.9%) to this question.

1. Tests Not Used With Some Clientele

- during the past 3 years (approximately) the age of our clients has changed to include those over the age of 16 - 18.
- I do not use intelligence tests with all children referred to me.
- I do not with Kindergarten and grade 1 students, or children of another culture.

2. Programming Reasons

- I use them on a very limited basis believing that most/ many program planning decisions do not require that level of detailed assessment
- I do, but judiciously. We do not report IQ scores per se, but focus on programming issues.

3. Referrals Are For Emotional - Behavior Problems

- many in my caseload are referred for social-emotional problems; student may not be sufficiently stable emotionally to participate.
- for some of my referrals it is obvious that there is no intellectual concern. It is emotional or behavioral.

4. Authentic Assessment Preferred

- Where possible, we prefer longitudinal, authentic assessment (where intelligence testing is simply one small piece of a process).

5. Lack of Time

- no time

6. Previous Tests Available

- previous test results available

7. Use of a Variety of Instruments

- I use them but they are not the only instruments that I use.

Appendix J

Question #4. If you do not use intelligence tests with children, what other assessment procedures do you use?

These answers represent 21(30.8%) respondents and are organized according to Tests Used and Assessment Procedures Used.

Tests Used

- adult intelligence tests - WAIS-R and Stanford-Binet Fourth Edition
- Observation, interviews with child teachers and parents, checklists, analysis of patterns of difficulties, responses to intervention efforts, adaptive behavior tests.
- classroom observation, informal assessment of reading and math skills, interviews, record review, history, checklists
- specialized tools - Leiter, Hiskey-Nebraska; Perkins-Binet; Wechsler Memory Scale; VMI; WAIS- R as a neurological instrument; Tests of adaptive behavior (Vineland, SIB, CTAB, ICAP)
- screening tests for mental illness and dementia (ie. PIMRA, REISS); academic tests (WRAT 3; PIAT); Draw-a person (Quantitative and SPED) APT-MR. CDI etc.
- McCarthy Scales for Children; use a combination of nonverbal measures with other cultures; Woodcock Johnson Revised - Cognitive.
- I use intelligence tests as well as Learning Behaviours Scales, social emotional measures (Coopersmith Inventory); behavioral-academic;

self-esteem inventory, and academic tests to have a good understanding of the child's functioning.

eg. Connor's Questionnaire, Achenbach Behavior Checklists, Piers Harris Self-Concept, Berry / Bender Etc.

-AAMD Adaptive Behavior Scale; Bender; Connor's Scale;
Personality of Self-Concept;

Other Assessment Procedures Used

1. Developmental/Social Tests

- Vocational, social skills, communication tests, developmental (ie. Insite Developmental checklist)
- developmental history
- projective tests, behavior checklists
- psychomotor tasks; questionnaires; projectives.
- rating scales (Connors or ADDES)
- in addition to intelligence tests, I also use achievement tests, rating scales to gather additional information about children.
- developmental

2. Classroom Observation

- At times I will use classroom observation
- observation
- In assessing children's learning needs, it is important to consider classroom functioning (work skills, attention, effort, social-emotional functioning) and background . Classroom observation
- observation

- other assessment procedures such as observation
- classroom observation
- I also use class observation

3. Assessment Information from Specialists

- pediatrician
- assessment information from other specialists (reading, resource, speech-language, medical, etc.).
- collaborating with other specialties occurs with all standard intelligence tests.
- consultation with other CGC staff, community professionals, family doctor ; Language Interpreter services.
- neurological cum file
- Information from other CGC clinicians if available (reading, SSW, psychiatry).

4. Family History/ Family Assessment

- family assessment; structured clinical interview
- interview with family
- family information
- family members' interviews

5. Parental Interview

- parent interview
- interviews with parents and child are all important
- parent teacher interviews
- Parent interviews

6. Structured Interview

- clinical interview with child
- clinical interview
- interview
- child interview

7. Teacher Reports

- teacher reports
- teacher interviews
- interview with teacher
- discussion with C.T./ other school personnel

8. Variety of Assessment Procedures Used

- other objective instruments
- I use IQ tests as a component of a battery of procedures.
- We use segments of various inventories only occasionally.
- I use other procedures along with IQ tests

9. Reasons Why Alternate Assessments were Used

- other children are of an ESL background, or are a recent immigrant/refugee where individual intelligence tests may be totally inappropriate.
- In general, we prefer not to use intelligence tests due to biases, labelling, dangers of misinterpretation and doubts about validity.
- some children are developmentally delayed and do not respond to standardized intelligence tests.
- 1:1 intervention, small group and behavioral programming

10. Drawing

- drawing
- drawings
- drawing

11. Authentic Assessments

- "portfolio" type assessments
- Most often we strive towards authentic assessment.

Appendix K

Question #4a. Why do you use other assessment procedures?

The following answers represented 21/68 (30.9%) respondents and are organized according to six areas: 1. Need to Get a Complete Picture of Student; 2. Planning/Programming Reasons; 3. Behavioral Reasons; 4. Technical Reasons; 5. Interviews with Parents, Teachers, Child; and 6. Other.

1. Need to Get a Complete Picture of Student

- to get a complete picture of how the student is functioning at home, school and in the community.
- an intelligence test only gives you part of the total picture
- to provide addition information
- to look for consistencies/inconsistencies
- to have a comprehensive view of the child's functioning levels, difficulties, problems, and strategies the child uses to learn, process and master.
- tests are only a "snapshot" and should not be seen to be a complete picture of a child. Individual subtests should not be used to overly justify any diagnosis over another.
- avoid labelling at a very young age.
- Observation over time, work samples, portfolios, creativity inventories, and other assessment tools (achievement, diagnostic and "personality") can all sometimes provide certain information -- if used humbly and wisely.

- I never limit intellectual assessments to just one measure.
Assessment should include (in my view) a classroom observation,
- I want to know how a child functions in all spheres: physical, emotional, social and intellectual. I also like to see how a child functions in context of the classroom. Testing is an artificial situation (but nonetheless part of assessment). It shows potential.
- tap more relevant areas
- I use intelligence tests with children and I also use other assessment procedures to look at the whole child. I do class observation- give out questionnaires about social-emotional -behavioral functioning, administer visual motor tests; sometimes the Peabody Picture Vocabulary Test is used to measure the child's level of receptive vocabulary. I also use WIAT achievement test and the Incomplete Sentences (projective).
- other skills assessments e.g. visual-motor, writing.
- broader look at the difficulties than solely intra-individual.
- intelligence devices should never be used on their own.
- to get more than a biased, socially-constructed, one-shot measure!
- to obtain info not available from intelligence tests.
- for a more thorough picture of what is and has been happening in child's life.

2. Planning/Programming Reasons

- to help me plan interventions
- not necessarily always for effective program planning/intervention

- more relevance to programming
- quick grade levels for rough program planning.
- to match child with tool.
- progress monitoring.

3. Behavioral Reasons

- children are rarely referred to me for one reason-- usually have academic delays, emotional-behavioral problems, family issues that are imparting on their progress.
- need information on motivation, attitude.
- often the reason for referral has nothing to do with academic functioning. I only use intelligence tests when I think they will give information relevant to the problem in the referral.
- social-emotional information.

4. Technical Reasons

- wanting to find true measures
- to provide some measure of reliability and validity.
- normed tests only give clues which must be confirmed

5. Interviews with Parents, Teachers, Child

- interviews with child, parent and classroom/resource teachers, behavior checklists, developmental information and school history review.
- I also base my recommendations on interviews with parents and teachers and child.

6. Other

- in order to do my job.
- More congruent with my systemic perspective
- provide valuable information
- fairness
- intelligence is a very complex area
- years of training and experience
- I specialize in assessment of people with mental disabilities.

Appendix L

Question #5. What do you perceive to be the strengths or positive aspects of the intelligence tests you use with children?

The following answers are organized according to 2 dominant themes: 1. Positive Use and 2. Other Comments.

Group Tests

There were 13 distinct responses from 11 (16.1%) respondents.

1. Positive Use

- I review the results of group IQ tests done by schools in the cum files before seeing a child for an assessment or counseling. The results often relate to my individual testing and add weight to the results I obtain. I can adjust my expectations in counseling somewhat after reviewing group IQ results.
- Strengths: Quick, allows large numbers to be tested in very short time. Gives indicator, catches gross problems early.
- I can help teachers adjust their expectations of students, since they don't often look at the results and wonder why child isn't performing.

2. Other Comments

- I don't. Only if you consider saving time and money as being more important than validity of results.
- I never give group tests. I use tests that can be given to a group (eg. TONI)

- very limited value, but it is fast and provides some information re: trends for classroom, school.
- screening device
- Only used as screening with very young group of children.
- I never use group tests as these are not appropriate for the mentally disabled.
- by using a group format, you miss out on interpersonal interactive ingredients, creative mistakes, analysis of miscues, and subjective impressions.
- very few - we view group inventories as particularly suspect in terms of validity.
- I guess you could use group tests to see the general level of functioning of a whole class but the purpose would be for classification and I don't see the point of doing that.
- I used it only in training (to learn how to administer them, to score them) and really never saw how the results were used.

Individual Tests

These responses represent 67 (98.5%) of the respondents to this question and are organized according to fourteen topics: 1. Assessing Cognitive Skills, Cognitive Processes and Learning Style; 2. Observations of Individual Behavior - Can Provide Encouragement to Expand Answers; 3. Comparative Purposes; 4. Program Planning; 5. Providing Understanding of Problems for Others; 6. Easy to Interpret and Administer; 7. Efficient - Get a Large Amount of

Information in a Short Time; 8. Identify Emotional Problems and Reactions; 9. Technical Advantages; 10. Identifying Learning Potential; 11. Predicting Academic Success; 12. Funding; 13. Grade Retention Decisions; 14. Learning Disabilities; 15. Screening for Special Needs.

1. Assessing Cognitive Skills, Cognitive Processes and Learning Style

- describes the ways the child learns best so that the teacher can utilize these methods
- further confirmation of observations and impressions - sometimes surprising results however!
- indicates intellectual strengths/weaknesses
- clarifies specific strengths/weaknesses
- provides information to confirm or generate new hypotheses.
- to assess learning style and strengths and weaknesses.
- the verbal - performance clusters have utility in understanding learning style as well as indicating recommended follow-up.
- I use the Wechsler Verbal Scales with ESL children not as a measure of "intelligence" but as an indication of their competence in using English for thinking, self-expression and problem-solving.
- they provide a measure of a child's ability and (hopefully) confirm what has already been observed by the classroom teacher and school - based clinicians.
- they also provide a basis for further consultations.

- however for kids who "deviate from the norm" substantially and who may therefore require individualized/specialized instructional programs, information about their strengths and weaknesses, cognitive functioning and learning styles can be helpful.
- Gives insight into specific strengths and weaknesses that this person has.
- Sometimes provides clues in the test profile which changes one's original hypothesis.
- can be specific about area of learning difficulty.
- can also be specific about child's strengths to build on to generate self-esteem, baseline to build success experiences.
- I can identify and confirm areas of a child's strengths or weaknesses.
- It gives us information about how a child approaches problems.
- helps determine preferred learning style.
- The non-verbal tasks/subtests along with the verbal tasks/subtests indicate cognitive functioning strengths and problems.
- helps identify cognitive weaknesses that will need further assessment for possible remediation (eg. speech/language, vision, psychiatry).
- adds another perspective to describing a child's abilities.
- They provide one piece of information in the very complex task of assessing a child's learning needs and ability and can work well with assessment tools from other disciplines: S/C. reading, Resource.
- identify strengths and weaknesses

- enables teacher to focus on problem areas, or avoid them depending on style.
- profile of strengths and weaknesses
- I can also point out the strengths of the child (sometimes teachers are totally amazed by the child's good test results).
- It can be an important piece of information combined with many other sources of information.
- aids in proper diagnosis, so that relevant literature can be brought to bear on the child's problems. (eg. nonverbal learning disability has implications for both math achievement and social skills).
- normed referenced IQ tests provide guidelines for evaluating both mental development and learning difficulties.
- can assess verbal-performance split.
- division of certain tests into various factors (visual, auditory) permits better understanding of child.
- The WISC-III is a kind of simple indicator of general information, auditory attention, abstract thinking, vocabulary and comprehension measures.
- gives more focused areas of abilities in the form of index scales, is applicable to a wide age range of children.
- provides info about verbal, performance functioning in several areas - auditory, visual, abstract reasoning, long and short term auditory memory, social judgement quantitative - visual

discrimination, visual - perception, visual-motor, motivation, sequencing and spatial abilities, horizontal and vertical scanning.

- the Wechsler Scales provide a comprehensive measure of both verbal and nonverbal abilities.
- The profile of subtest scores proves useful for interpretation purposes relating this information.
- they constitute one or more measures of ability (as do achievement tests or achievement period) but they are not the definitive measure of ability.
- WISC-III shows relationship between verbal and non-verbal skills.
- WISC-III - levels at Verbal/ non-verbal abilities, child's memory (auditory, visual), reasoning, general knowledge.
- scatter of subtests - really important in my work; helps to show areas to do remediation or strengths to use in working with child.
- helpful in forming diagnostic impression
- I use it as an indicator of child's level of functioning to make further decisions about how to go assess his/her strengths and weaknesses through portfolio approaches, classroom based assessments.
- occasionally can be used to confirm specific or general diagnosis for the purposes of placement.
- It provides information that can later be used for recommendation (re: learning style).
- provide evidence for advocacy issues

- covers verbal strengths, perceptual spatial abilities
- Once you become familiar with a good test and its limitations, it becomes easier to see individual strengths and weaknesses in the patterns of the subtest scores.
- provides objective information about student
- gives info re: related abilities, eg. attention, confidence level.
- can provide a good deal of useful information.

2. Observations of Individual Behavior - Can Provide Encouragement to Expand Answers

- the individual nature (time to observe child) on an instrument I have used with lots of kids.
- testing of limits permissible with the tests permits opportunities to break down the tests into smaller pieces and observe child better.
- give me an opportunity to observe the child with the instrument.
- we can see how child approaches tasks and verify what teachers have observed.
- on 1:1, I can ensure child is given opportunity to answer questions to show his/her ability and is comfortable.
- If you observe behavior during the test you can discover a great deal about the child.
- provides valuable indications about how a student takes tests.
- I do not assess until some rapport has been established so I get the most accurate results.

- the children usually enjoy the experience especially the performance items.
- provides opportunity for rapport
- able to use them as diagnostic tools for observing approach to tasks; tolerance to frustration, work habits etc.
- offers time to understand the child - as opposed to group tests.
- shows what a child can do in a structured interview type session.

There's probably not a lot of useful information to be gained about children who function in the "average" range - beyond observational data pertaining to problem-solving strategies. ability to handle frustration, etc. - which can be obtained from other sources

- the Wechsler Performance Scales enable me to pick up qualitative information about the student's feelings and attitudes eg. approaches to a problem, self-guidance, self-monitoring, anxiety levels when doing certain tasks, persistence etc.
- way to interact with the child in a structured format where observations can be particularly helpful.
- individual testing also gives you unique opportunities to observe the child in an uninterrupted problem - solving situation.
- opportunity for personal observation of child's handling of challenging cognitive tasks and those "easy" for him/her
- observations in the classroom or one - on - one chats with a child will never give you as much information in a limited time period

- an IQ assessment essentially provides a structured observation period to assess a wide variety of behavior. Although this is not necessarily the intended purpose of an IQ test, it is undoubtedly a valuable one.
- what's important to me is the child's method of approaching a task and the various subtests provide opportunities for observing the child's strategies.
- they provide a wide variety of tasks and therefore the clinician has the opportunity to observe the child's approach and success with them.

3. Comparative Purposes

- comparison "inter" and "intra" individual differences.
- they provide an indication of strengths, either in relation to the student's own skills or in relation to the population.
- provides a standardized content from which to observe child's behavior, reactions to success/failure, attention, etc.
- you compare a child with his/her age group in a test situation, which is standardized.
- standardization allows for comparison across children
- the fact its a standardized instrument by which children can be compared - particularly interesting in observing their behavior and emotional reactions during testing.
- provide empirical data
- acts as a comparison to a standardized sample.

- using a standardized test is obviously a strength when you need to provide "hard" evidence for your recommendations or to communicate with other professionals.
- standardized and objective
- Provide normative data, profile of student.
- helpful to contrast with same age peers
- gives you a comparison of individual to group
- normative data relevant for interpretation to student / parents / school/ Dept. of Education
- they give a comparison with a larger population than the classroom.
- normative data provides points of comparison which are standardized.
- achievement testing using the WIAT can be directly compared to ability level when using the WISC-III or WPSSI.
- can assess if strengths and weaknesses are statistically significant.
- ability to assess the child's strengths and weaknesses relative to same-aged peers.

4. Program Planning

- provides info to school personnel for ways of teaching on a day- to-day basis.
- child can be taught at his/her level - some kids are frustrated because of too high expectations

- provides context from which helpful suggestions can be made for school program.
- helps to identify the most appropriate means of teaching the child.
- this can provide a foundation for programming strategy, positive expectations, success, hope, planning, etc.
- helps in verification of ADD symptomology (Distractibility Factor)
- verbal and performance scales (Wechsler) are very useful and have clear implications for individualized programming as well as for later vocational choices.
- aid special programming
- they should not be relied upon as the only source of info re: planning for therapeutic change for the child.
- I can get an indication of the child's strengths and weaknesses to aid with programming.
- provides a piece of information about the child that may be useful when planning programs.
- very useful info when developing a program for the student.
- provides useful/valuable information especially pertaining to educational programming as many of the areas are relevant to school progress/achievement.
- Provides information for developing school programming that emphasizes the child's cognitive strengths (learning style).
- They are useful to identify long-term programming needs.
- provide basis for individual help recommendations.

- helpful in identifying strengths to tap into for accommodating programs.
- Well researched implications for use. Gives us a place to start in determining a programming direction.
- helpful in planning strategies

5. Providing Understanding of Problems for Others

- helps understand learning difficulties
- as part of the process (not a definitive, god-like technique), they can sometimes give us information, and ideas how better to help our students and clients.
- gives concrete evidence to explain why a child may be having problems so that teachers and parents can better understand that child and meet his/her needs.
- provides a profile of strengths and weaknesses, additional info about the child to help teachers, parents, children understand child better.
- some children need encouragement to expand on answers - what you first get is not all the child knows - useful info for teacher.
- the tests usually give a variety of scores. I tend to use percentiles and "approximate age equivalents" in presenting test results to teachers and parents because these are easy to understand and practical. eg. a grade 5 student with Verbal scores in the 6 - 7 years range will not understand much of his grade 5 novel, even if he is able to decode adequately.

- I have learned to interpret scores in a way which makes sense to teachers and parents and has direct implications for the teaching/learning situation (more a function of my own experience than of the tests themselves.)
- provide information to parents; suggestions dealing with behavior.
- help establish reasonable expectations: of the student, of the school and for the parents.
- to help teachers understand their students, adjust their expectations for them, and develop appropriate programs (eg. vocabulary development, social problems-solving etc.).
- a good screening for language disorders or visual-perception problems so that speech/language pathologist and occupational therapist can be involved to provide further assessing and therapy if need be.
- the student often does well in areas not probed by the teacher in the classroom.
- helps at times to modify unhelpful expectations of student and significant others.
- provides an "allowable, safe in" with the psychologist when there are other concerns that need to be addressed (eg. behavior, family situation, personal issues , - too threatening for family to sign consent.)

6. Easy to Interpret and Administer

- easily understandable.
- user friendly - kids/ examiner
- WISC-III is easy to administer and holds children's interest.
- TONI- performance scale score - good quick screening tool.
- McCarthy - better for use with children 7 and under
- efficient way to get an overview of child
- user friendly.
- not labelling devices. We have the freedom to use the test usefully versus needing to label kids.
- Kid friendly in building a rapport.
- most are relatively non-threatening
- most children seem to respond positively to the structure of packaged tests; it seems to put them at some ease with me
- the WISC and WPPSI have been updated and I have found the changes to be appropriate and useful.

7. Efficient - Get a Large Amount of Information in a Short Time

- ie. obtain a large amount of information in a short time.
- comprehensive
- WISC-III provides rich clinical information.
- McCarthy - breaks down further than the WISC-III; more factors than just verbal and performance.
- IQ indexes for perceptual speed.
- fairly thorough, tests wide range of differing abilities.

gives good overview of the students' learning styles, strengths and weaknesses

- under optimum conditions eg. interesting items, short time period for each subtest
- gives information about the individual that cannot be gleaned from other sources.
- variety of tasks and variety of models of intelligence represented
- also the quality of answers gives you additional valuable info
- lots of info re: abilities gleaned in relatively short period of time.
- time effective method to assess a variety of abilities

8. Identify Emotional Problems and Reactions

- provides information on how they deal with frustration, pressure (ie. timed tasks) the processes they use to achieve answers.
- provides a standardized content from which to observe child's behavior, reactions to success/failure, attention, etc.
- allows you (tester) to work with student and get a sense of behavior, attitude, effort (etc) and make comments/judgements about validity and reliability.
- gives you impression of individual's approach to a task
- close look at child's approach to learning
- good way to assess how the child relates to the examiner, to verbal questions and how he/she problems solves non-verbal items.
- to assess clinically how the client deals with specific tasks.

- useful tool for eliciting clinically useful observations (eg. distractibility, memory strategies; reaction to non-success.
- allows for assessment of individual traits such as perseverance, ability to cope with stress, self-evaluation of work efforts motivation etc.
- in addition to acquiring information about student's cognitive profile, one also has an opportunity to view how the student approaches tasks (ie. sequentially, simultaneously, etc.) how they respond to their successes and failures and how motivated a child can be when the environment is conducive to their learning needs.
- observing the child in relation to the test materials eg. attention span, concentration, interest, coping resources, affect, self-esteem, add to the assessment process.

9. Technical Advantages

- well researched, standardized and normed.
- The scales are well normed.
- well normed
- good norms
- a normed instrument with proven validity/ reliability.
- good validity and reliability
- good reliability and validity compared to other standardized measures
- TONI- in my experience it correlates well with WISC-R/ WISC-III
- stability over time (for most children)

- well established standards/reliability
- it's a stable instrument - for academic potential verification - used over along period of time with increased reliability and normative samples.

10. Identifying Learning Potential

- efficient way to find out ability and potential to help teachers get a different perspective on child as they work with classroom behaviors,
- it's the best I know of for assessing potential reliably.
- In most cases it serves to support one's clinical judgement regarding a child's potential.
- fairly good predictor of the amount of struggling or ease a child will work through the curriculum with.
- can identify gaps in learning or potential learning problems
- can identify gaps in learning or potential learning problems
- give you a better sense (impression) as to what the student is capable of so as to make appropriate expectations.

11. Predicting Academic Success

- verbal tests in particular do give a good idea of classroom success.
- good at predicting academic success
- some correlation with ability to problem solve and with specific academic activities.
- WISC-III seems to have good validity in predicting academic achievement.

- good predictability re: future achievement and long-term program planning.
- intelligence tests are good predictors of school performance, regardless of what they actually measure.
- scores seem to correlate with school achievement and can account for learning problems or rule out reasons for learning problems.

12. Funding

- to provide documentation for funding.
- useful for funding

13. Grade Retention Decisions

- aids in decision- making re: retention (grade).

14. Learning Disabilities

- can sometimes "pick up" specific learning problems.

15. Screening for Special Needs

- excellent screening for special needs students.

Appendix M

Question #6. What do you perceive to be the limitations or negative aspects of the intelligence tests you use with children?

The following answers will be organized according to dominant themes.

Group Tests

These answers represent the responses of 9 (13.2%) individuals and are organized into 8 categories: 1. Can't Observe Subjects/ No Emotional Reactions Available; 2. Does not Examine Process Skills; 3. Limited Value for Classrooms; 4. Technical Issues; 5. Lack of Diagnosis - No Subtests; 6. Inappropriateness; 7. Not a Good Predictor of School Performance; 8. Suseptibility to Outside Influences.

1. Can't Observe Subjects/ No Emotional Reactions Available

- does not build on cooperation
- unable to observe "inter" subject behaviors.
- observations with respect to "how" a child does a task is extremely important. You don't get this info in group tests
- also emotional reactions will be overlooked.
- may not always be accurate, since I did not see the child take the test.
- It doesn't provide a lot of information, as it is difficult to monitor exactly what is going on for each child

2. Does not Examine Process Skills

- does not examine process skills, creativity
- does not get to 'depth' of learning but rather the focus is on facts

- don't get secondary factors that help in diagnostic impression.
- the type of abilities measured are sometimes very broad
- no test can test all the abilities implied.

3. Limited Value for Classrooms

- teacher may treat children with lower IQ's differently - ie too low expectations
- limited or 'no' value re: classroom strategies
- encourages a narrow curriculum
- temptation to teach to the test

4. Technical Issues

- typically don't get an impression of attitude and effort which are major factors that affect reliability and validity.
- noise and interference in group can invalidate test.
- Some difficulty with norms and standardization procedures.
- issues about validity

5. Lack of Diagnosis - No Subtests

- you don't get to individual differences which programming is based on.
- fairly blunt, not likely to catch minor problems.
- the tests are often not classified into subtests which could provide specific information on strengths and weaknesses.

6. Inappropriateness

- mass screening - inappropriate testing for testing's sake - overuse
- labelling (and negative self-fulfilling prophecy)

7. Not a Good Predictor of School Performance

- are not always good predictors of school performance.

8. Susceptibility to Outside Influences

- very susceptible to environmental influences, such as nervousness of the child in a strange situation.

Individual Tests

The following were organized according to 21 dominant themes:

1. Doesn't Provide Enough Information;
2. Testing Poses Cultural Problems for ESL Children;
3. Other Variables Affect Testing;
4. Others Don't Understand Limitations;
5. Testing Can Be Unsympathetic to Child;
6. Not Related to Academic Performance;
7. Definition of "IQ" or "Intelligence" Poses Problems;
8. Testing Must Be Done Cautiously;
9. Labelling is Seen as Permanent Description;
10. Misuse of Scores;
11. Negative for Individual Tested;
12. Not Sure What Some Subtests Measure;
13. Report Writing is Time-Consuming;
14. Inadequate Measure for Low Functioning Individuals;
15. IQ Tests Need to Improve Norming Population;
16. Screening;
17. Testing Can't Generalize to Other Aspects;
18. Labelling Can Be Helpful;
19. Children Become Unjustly Treated;
20. Results of Testing Show How Child Functions at That Time;
21. Testing Can Be Helpful; and represented 65 (95.6%) of the respondents.

1. Doesn't Provide Enough Information

- important not to make wide - ranging recommendations on the basis of an intelligence test only!
 - my main concern would be making "pronouncements" about a youngsters capabilities based solely on one instrument. I think intelligence tests can provide useful information when used in conjunction with data obtained from other sources - classroom observation, portfolios, examinations, perhaps other assessment tools depending upon the nature of the information one is seeking about the child's functioning.
 - Information from them must not be viewed in isolation. A variety of tests need to be used, as well as information from home and school to get a good "total" view of an individual's attention. Personality factors, attention, motivation, social-emotional issues all impact on an individual's functioning.
 - shouldn't be the only test given or used to make diagnosis / recommendations.
 - too much influence is often given to one test or part of a test to justify people's beliefs about what a child needs or what a child's ability is or learning strengths and weaknesses without regard for what the test really measures or was originally designed for.
- Intelligence tests do give some insight into behavioral or social -emotional problems, but they have to be supplemented with other assessment procedures for a fuller picture (eg. systematic observation,

social skills assessments) This is a limitation rather than a "negative aspect".

- danger of value judgements being made based on ability level.
- I'd like to see more correlations between factor scores and achievement problems in reading, math etc.
- they often claim to measure more definitively than they do eg. vocabulary on WISC-III is a very limited range of vocabulary. As well in discussing it with a speech and language pathologist, found that their tests start with simple concepts. She feels the words were at a higher level of difficulty than necessary.
- The Verbal Scale provides only a limited assessment of language abilities, an area that is more thoroughly assessed by assessment instruments within the battery of tests used by speech-language pathologists.
- WISC-III misses areas of intelligence - measureable on Woodcock - Johnson eg. music ability etc.
- WISC-III does not provide good measure of Freedom from Distractibility factor.
- intelligence tests do not pinpoint learning and developmental disorders to the degree that other psychometric instruments (eg. MMPI) identify problems.
- non global. Addresses only a fraction of identified aptitudes and abilities.

- there are other types of intelligence which standardized instruments like the WISC-III do not tap into.
- these tests only measure "intelligence" as narrowly defined - eg. ignores Gardner's other intelligences, creativity factors etc.
- intelligence tests do not assess a child's creativity, ability to relate with other individual or specify a child's interests or concerns in life. Thus the total child is not assessed through an intelligence test.
- instrument isn't sensitive enough to diagnose specific learning disabilities.
- don't provide info on other "intelligences" eg. social, artistic, divergent thinking etc. therefore they can't stand on their own.
- must not be used alone - must collect more information through assessment, observation, parent input.
- intelligence is very multifaceted and testing identifies only certain aspects of cognitive functioning.
- at times, school request learning potential assessments for the purpose of applying for funding. It is at these times when observations, interview with the teacher, and examples of student's classroom work would be more essential in providing information regarding the student's abilities and/or limitations for funding purposes rather than an IQ score.

2. Testing Poses Cultural Problems for ESL Children

- cultural discrimination - especially with aboriginal population (ie performance and verbal).
- difficulty obtaining reliable and valid results for children who have problems with the English language (ie some aboriginal children.)
- More research is needed on disparities between the Wechsler and Stanford-Binet in special populations, and on validation of interpretation. A Revised Wechsler Adult Scale is long overdue.
- I would like norms that are appropriate for inner city aboriginal children. I usually give a nonverbal test such as the TONI as the WISC-III is normed on mostly white American working population.
- test/retest scores with exceptional population not as reliable as with general normative group.
- verbally oriented subtests tend to be culturally biased in the uses with native children.
- cultural differences are not handled well by our usual standardized tests
- biased by language, culture, enrichment, mood, maturation, emotional problems disposition of the examiner, physical health.
- you have to consider the population you are working with. eg. I work in the inner city.
- culture bias remains a concern
- strong cultural bias especially when dealing with First Nations students.

- cultural and linguistic bias need to be taken into account when assessing minority and/ or ESL students. American bias in WISC-R and WAIS-R re Canadian students.
- cultural differences of the children - tests biased towards white middle - class type children.
- most of the standized tests are culturally biased and tend to depress the scores of minority students.

3. Other Variables Affect Testing

- sometimes test results are not necessarily a reflection of the child's ability (eg if child is anxious - or has difficulty concentrating)
- Psychologists have to be sensitive to issues affecting the child's functioning on the tests.
- possibility child had a bad day - testing unknowingly would be invalid. However, as a rule, look at other aspects.
- limited sample of behaviors on a given day. Susceptible to lots of factors such as student's mood, bad morning at home, fear of strangers, colds, which the examiner may not be aware of.
- anxiety for kids who don't like the focus to be on them.
- results inconclusive with emotionally disturbed population.
- the time limits, different cultural background, early social-emotional definitions, lack of language and other age and stage experiences are not "standardized" variables which affect all children
- some more and some less and there is no accomodation of those differences.

- if you are aware of the limitations of testing in general (performance anxiety etc), negative aspects are minimal.
- If a child is anxious, that is valuable information too for school functioning.
- The test may be interrupted (ie recess, a teacher walking in, noise.)

4. Others Don't Understand Limitations

- people don't realize that these don't provide a "magic answer".
- You have to be careful when discussing the results of the meaning of the intelligence tests. There are quite a few people who are misinformed about 'intelligence' tests. When discussing/interpreting the results of a test, I try to provide as much information as feasible as to the meaning or purpose of a test, the type of ability/characteristic/skill required to do the test, what the score means in practical terms and so on.
- parents often do not fully understand the test results - are upset if their child obtains a low score.
- results can be misinterpreted - often teachers do not understand concept of ranges - they want a number!
- other individuals (school staff, parents and psychologists) can focus too much on the results of one test in the total assessment process. (eg. for school placement services only).
- many people do not correctly understand the use for intelligence tests and the info they give you which can result in misinterpretations and or overuse.

- can be used in a narrow way i. e. given scores may lead to fewer expectations from teachers.
- some psychologists and professionals/parents receiving test data overestimate validity.
- with young children, reliability/validity is sometimes suspect.

5. Testing Can Be Unsympathetic to Child

- utilizes oral response - some children do better if written (visual)
- often you can't accomodate a child's needs without ruining validity of results (ie language)
- built into the standardization is the need to ask questions/present tasks which the person can't do. This has the potential to make him/her feel "dumb".
- IQ does not look at multiple intelligences - just a general G.
- although the tests are improved there are still individual items that I see as questionable re: degree of difficulty.
- in testing younger children, one question can make a huge difference in overall scores.
- I prefer to use it as a diagnostic instrument and I can't ask how or why to do something as it would interfere with validity. I guess standardization and diagnostics can't coincide very well. (I do ask after sometimes).
- tests results used for placement purposes are in my opinion, not used with the proper objective in mind.

6. Not Related to Academic Performance

- provide little information on dynamic performance
- not easily linked to performance on academic tasks
- focus is on the child when the milieu is critical. eg. classroom environment and teacher is not a part of the test only process
- very difficult and precarious to view the child as an individual with strengths and weaknesses instead of an IQ score but not translate those strengths and weaknesses into specific classroom potentials, recommendations as that is not the intent of the test.
- measures stable traits that are not easily modified via training. Not like Feuerstein's LPAD.
- only represent that child's performance in a 1:1 setting.
- direct connection to the classroom and specific subject areas not always apparent.

7. Definition of "IQ" or "Intelligence" Poses Problems

- limited definition of intelligence restricting intellectual freedom.
- the term "intelligence" is too global and I rarely use it to communicate findings to parents or teachers.
- the term "IQ" carries a lot of "semantic baggage" and it arouses anxiety unnecessarily. All an "IQ" is, is a score on a particular kind of test. In explaining test results I talk of scores not IQ's.
- the terms "intelligence" and "IQ" imply a static, unchanging individual trait which determines future performance in a wide variety of life situations. In explaining test results I interpret them as

"current functioning levels" and leave them open-ended ie. changes may occur, and often do occur (esp. for younger students.)

- used "softly" and flexibly, there are plusses, but it is naive, narrow and limiting to think we can truly define and or measure intelligence
- it's often difficult to find a very private place to give a test within a school.

8. Testing Must Be Done Cautiously

- dangers of "playing God" and really believing in this stuff.
- too many psychologists are presumptuous and dogmatic in test usage and interpretation.
- often psychologists over analyze and make claims that are very weak.
- I think we have to use testing cautiously and with wisdom/judgement.
- an artificial situation
- one-to - one support (sometimes not realistic since child must behave/interact in group setting)

9. Labelling is Seen as Permanent Description

- the scores are often interpreted as being more accurate than they really are or permanent indicators. Children are mislabelled and treated unjustly as a result. The tests and results should be used only for the child's benefit.
- It must not be used as the indicators of learning potential to categorize and group students and to label them.

- can lead to "labelling"
- danger of labeling or perceiving level of functioning to be a fixed entity.
- kids tend to get labelled throughout their school careers on the basis of one test.

10. Misuse of Scores

- easily misused ie. labeling; not providing a service because of limitations, self-fulfilling
- kids tend to be labelled by their IQ scores.
- too much reliance by other professionals eg. social workers, bureaucrats providing funding and sometimes teaching personnel on the actual IQ figure which you generate and little emphasis upon other data such as test behavior.
- people get number bound

11. Negative for Individual Tested

- poor results can really impact individual's perception of self
- self-fulfilling prophecy - different treatment that may be harmful not helpful.
- testing done just to find out a "number" no purpose for assessment.
- concerns about results being used only as verification of ability without follow- up in terms of remediation and expectations of improvement and change.

12. Not Sure What Some Subtests Measure

- Certain subtests, such as Object Assembly and Symbol Search, can be difficult to interpret in terms of what they actually measure, and, as such, it becomes difficult to actually derive the educational implication of significantly high or low scores on such subtests.
- individual subtests do not always have great reliability unless the deviation is quite large.
- Intra-subtest scores can't legitimately be used unless it is very large.
- usefulness and validity of the factors on the WISC (esp. Freedom from Distractibility) questionable.

13. Report Writing is Time-Consuming

- report writing after an assessment is completed is a time-consuming task.
- takes time that might be spent on other forms of assessment.
- time-consuming
- time - consuming to do and write up - time could be utilized more effectively other ways often.

14. Inadequate Measure for Low Functioning Individuals

- it's difficult to explain IQ 'drop' over time in some low functioning children
- too high a floor - inadequate for assessment of low functioning individuals.
- The whole child has to be considered, and many children with lower IQ have been mainstreamed with positive results.

15. IQ Tests Need to Improve Norming Population

- Intelligence tests need to broaden their range of items and include more people with disabilities in their trial populations.
- the Wechsler Scales are normed on an American population base, not Canadian. Some of the information subtest questions are too American based, although this has improved with the Third Edition of the WISC.
- norms may not fit population well - Hutterite kids, English as 2nd language kids

16. Screening

- only a starting point
- They can provide clues, but they don't provide definitive answers. they also can tend to be overused.
- the tests sample abilities but conclusions can be premature e.g. "digit Span score inactivates short-term auditory memory" may be too general based on 1 test.

17. Testing Can't Generalize to Other Aspects

- generalization to other tasks has shaky validity
- individual intelligence tests are very specific in what they measure, and do not measure all aspects of intelligence. There are even questions whether what they say they are measuring is actually what they are measuring.

- don't allow for assessing "teachability" i.e. demonstrating a task and then determining how child uses instruction to improve performance.

18. Labelling Can Be Helpful

- contrary to many of my peers, I don't think "labelling" is all bad - sometimes it's helpful to have some diagnostic information so that you can have a clearer picture of what you're dealing with, hence, develop appropriate supports and interventions for the child.
- I don't like to "label" children. But a label can be taken away and used to the benefit of the child.

19. Children Become Unjustly Treated

- misperception by teachers or parents.

20. Results of Testing Show How Child Functions at That Time

- there are always intervening variables (eg family issues) so that you never truly measure a child's ability. This is how they are functioning at this time in their circumstances.

21. Testing Can Be Helpful

- they are not the only thing used to assess children. However, without them no assessment is complete. In my experience, there have been times when children have been referred for counselling and no IQ was done. Later, when no progress has been made, we have done an intellectual assessment and the results have provided an explanation because we were guessing that the child was of normal IQ but actually not. This has happened quite often.

Appendix N

Question #7. Do you have any other comments you would like to make about intelligence testing with children?

The following responses are organized according to nine dominant themes: 1. Psychologists Must Use Tests Responsibly; 2. Intelligence Tests Have Their Place; 3. Assessment Has Many Parts; 4. Intelligence Tests Can be Misused; 5. Intelligence Tests Can be Ineffective; 6. Validity of Results is Affected by Many Factors; 7. The Public Needs to Be Informed of Advantages and Disadvantages, 8. Clinicians are Part of a School Team; 9. Intelligence Tests Has a Low Priority; and represented 39 (57.4%) respondents with 49 distinct responses.

1. Psychologists Must Use Tests Responsibly

- Hopefully the tester is better than the tests so a meaningful interpretation can be made.
- Extremely important as to how they are used, for what purposes and absolutely necessary follow-up communication between teacher and psychologist regarding working towards improvements or changes to benefit child.
- It is the responsibility of the psychologist to not "over sell" or over interpret the results of an intelligence test, and to see it as any other piece of the assessment process.
- it should not assume an over-large portion of a clinicians practice.
- The responsibility for what happens in intelligence testing reside almost solely with the tester. All testers should realize that the

possibility for error exists at all times. Suggestions should be guidelines to help children based on test results, not narrow classifications designed to permanently pigeonhole kids.

- If we get too rigid and narrow in our view of what psychologists do, we will overlook a greater contribution that we can make to the emotional well being of our clients.
- we're reluctant testers: at times it's necessary to run these things due to "societal rules" (ie for admission to treatment facilities, accessing supports, funding, etc.). In other words, to help some people, it has become socially acceptable and indeed mandatory to test. Caution should be exercised however!
- any measure has to be used with caution.
- I always remember my report will be following this child throughout their school career and will be interpreted by many others.
- I find myself feeling less and less comfortable talking about specific skill weaknesses and their relevance to functioning (eg. Johnny is a concrete thinker") based on performance of any particular subtest.
- I believe it is a helpful guide but it must be used carefully and not be over interpreted.
- after practicing psychology for nearly 20 years I believe intelligence test are here to stay. When used properly and appropriately, tests enhance the assessment process. The new WISC-III disappoints

however, (more of the same) and I hope future tests will become more sophisticated!

- Like any other tool, they can be used in an inappropriate fashion to do a great deal of harm, or they can be used sensibly to do a great deal of good. It depends on the person administering them to behave responsibly. The child's future may depend on it.

2. Intelligence Tests Have Their Place

- it has its place
- After doing a large number of individual tests what is more critical is the "differences" that show up. These are not related to scores but to the more subjective elements.
- Done with care, compassion and understanding, test results can serve as protection for children. Secondly, need areas can be identified early and hopefully addressed through recommended programming.
- Let's not throw the baby out with the bath water!
- I always try to de-emphasize actual IQ scores. Rather, I emphasize the strengths and weaknesses and the pattern of scores.
- can be useful if used appropriately and positively.
- They are an essential component of any assessment because it gives us an idea about what we have to work with and gives us an idea of what areas to concentrate on during remediation.

- I feel intelligence testing is an efficient way to gain intellectual/behavioral and sometimes emotional information about children.
- It is often not understood how valuable they are in lots of ways: if you know how to interpret them you can obtain a fair amount of data other than IQ: such as learning style, specific disabilities and even presence of anxiety.
- I've found them helpful for understanding why children behave and learn the way they do. At times, you need some objective measure to cut through parental denial and give children what they need.
- Psychologists need to retain and celebrate their expertise in this area but need to move to a more selective, differentiated use of such tests.
- Generally speaking, IQ tests are helpful in making general programming decisions with children.

3. Assessment Has Many Parts

- I think it is a useful and essential tool if used "intelligently". As such it is only one facet of the wider assessment process.
- am doing less "intelligence testing" now than I did when first beginning work as a school psychologist - a good thing. More emphasis now on how kids approach tasks rather than scores per se in order to produce relevant recommendation for teachers.
- The social context of testing needs to be understood. In the school I work in, children of all abilities are mainstreamed as much as

possible; there are no "special classes" so tests are not used to select or place children into such classes; academic, social-emotional and behavior problems are often intertwined and a variety of assessment tools are used in addition to or instead of "intelligence tests".

- The main purpose of testing is not to put diagnostic labels on students but to find out which of many factors are contributing to the child's problems (including academic, life-skills, social skills, counselling, behavior management) and general support for the student.
- I think that intelligence plays an important part in a holistic assessment. It is also important to observe in class, review class work and interview with teacher/parents/student.
- I view the testing process as both a quantitative and qualitative process. It provides a rich field for both types of data. The data are the pieces of the puzzle which together with parents, teachers and the child's observations contribute to a better understanding of the child's his/her needs.
- should not be administered unless examiner has enough background to utilize tests in conjunction with other information and therefore is able to judge meaningfulness, significance of differences. Not to be used in a black/white cookbook fashion.
- There are many better (curriculum-based) ways to determine how to improve instruction. I would help the intelligence test would

provide some additional information but be only part of the student's assessment.

- Seems to be a wave against testing lately. Although IQ testing is not always used for the right reasons that does not mean we throw out the baby with the bath water. Utilized correctly and with other assessments it is a valuable tool.

- Testing should be only a small aspect of drawing pertinent information about students. Often it's the clinical impressions/observations which are the most helpful in the assessment process. It is important to use testing as only a part of this process.

- I highly believe intelligence tests should only be administered in conjunction with others. Then the results of tests should be incorporated with ecological and historical info about the client so as to interpret them correctly. They are part of, NOT THE ASSESSMENT.

4. Intelligence Tests Can be Misused

- They have been 'oversold' and are 'overused' by many psychologists trapping many of us by unrealistic expectations and demands from teachers and parents.

- psychologists may be partially blamed for encouraging the mystique regarding IQ testing.

- I feel that school personnel request testing for IQ measure only and do not see the utility of other evaluating techniques to be just as

informative or if not more informative than IQ tests. When an assessment is requested, all too often, the school is interested in acquiring the IQ score, the recommendations and general impressions, but rarely has the time, resources, or energy to follow through on the suggestions. This is when I feel that maybe we should take a second look as to why we put our students through 2 hours of testing (which can be traumatic) if there is not real commitment to follow through beyond the reporting of the results.

- Testing should not be used for entry to specific programs.

5. Intelligence Tests Can be Ineffective

- In a more ideal world, only slightly more ideal, we wouldn't bother.
- Clinicians must be aware of the limitations of tests.
- must be used only under very special circumstances and is of no value if it doesn't improve the situation for the child.

6. Validity of Results is Affected by Many Factors

- When a child is not functioning well in school (subjects) or has a learning problem, his/her test scores on the intelligence tests will be depressed and will not be reflective of his/her actual potential/ability. Low self-esteem, auditory comprehension, discrimination etc will affect test results.
- it is one measure/one day at a certain time and performance can be influenced by many things.

7. The Public Needs to Be Informed of Advantages and Disadvantages

- there needs to be more public information regarding both the advantages and limitations of IQ testing.
- I have concerns in the way information is provided to parents. I feel there is great discrepancy in how information is framed and whether cautions are always outlined. Also reports and interpretation of results- especially subtest scores - can vary.

8. Clinicians are Part of a School Team

- Clinicians work as part of a school based team, not as "outside experts"

9. Intelligence Tests Have a Low Priority

- It is not a critical task and certainly comes lower on the priority list for psychologists, yet does have its usefulness, helps kids if used properly.