

**Educator's Perceptions
of
Educational Accountability**

by
William Joseph Kilfoil

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts in Education

Faculty of Education
Saint Mary's University
Halifax, Nova Scotia, Canada
February 1994

(C) William Joseph Kilfoil 1994



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

1. Acquisition - Bibliothèque nationale

2. Bibliothèque nationale - Acquisitions

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 0-315-90961-7

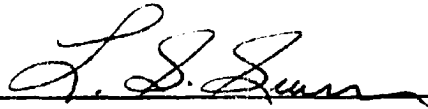
Canada

Saint Mary's University


Faculty of Education

The undersigned certify that they have read, and recommend to the Faculty of Education for acceptance, a thesis entitled EDUCATOR'S PERCEPTIONS OF EDUCATIONAL ACCOUNTABILITY submitted by WILLIAM JOSEPH KILFOIL in partial fulfilment of the requirements for the degree of MASTER OF ARTS IN EDUCATION.

Dr. T.S. Sullivan



Dr. T. Piper (Dean)



DATE:

MAR 04 1994

SAINT MARY'S UNIVERSITY

RELEASE FORM

NAME OF AUTHOR: WILLIAM JOSEPH KILFOIL


TITLE OF THESIS: PERCEPTIONS OF EDUCATIONAL
ACCOUNTABILITY

DEGREE: MASTER OF ARTS IN EDUCATION

YEAR GRANTED: 1994

Permission is hereby granted to SAINT MARY'S UNIVERSITY LIBRARY to reproduce single copies of this thesis and to lend or to sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive abstracts from it may be printed or otherwise reproduced without the authors written permission.



William Joseph Kilfoil

72 Sunnyholme Drive

Oyster Pond, Nova Scotia Canada

BOJ 1WO

DATE: MARCH 4/94

ACKNOWLEDGEMENTS

I would like to acknowledge the support and encouragement of a number of individuals during the preparation of this thesis. In particular I appreciate the confidence of the staff of the Halifax County Bedford District School Board who granted me the sabbatical leave that allowed me to complete this project.

In addition I acknowledge the assistance provided by the teachers and principals who allowed time in their busy schedules to complete and return the questionnaires which were a crucial part of the work.

The counsel and guidance provided by my thesis director, Dr. Terry Sullivan was essential to the completion of this study and is greatly appreciated. I also thank Dr. Bernard Davis for his assistance and patience with the statistical analysis.

Finally I thank my wife, Yvonne, for her encouragement, patience and tolerance during the preparation of this study and my children who were invaluable as data entry assistants.

ABSTRACT

The purpose of this study was to determine how teachers and administrators feel about the possibility of identifying and measuring specific educational outcomes and being held accountable for the measurement. With regard to these concepts, the study was designed to determine if there is any difference in attitudes and perceptions among the various demographic categories of educators. In addition, the relationship of these attitudes to the results of the Nova Scotia Achievement Tests and to each other is examined.

While there is a range of opinion (and a significant number of undecided) the study indicates that teachers and administrators generally believe that cognitive educational outcomes can and should be identified and measured and that the "results" of education deserve at least as much attention as the inputs and processes of education. The study also indicates that teachers and administrators (generally) feel that the school communities (the public) have a significant role to play in the educational process and should have access to information about school performance. Attitudes toward components of an accountability relationship were examined to determine their relationship to each other. For example, perceptions of identifiability, measurability, and accountability were found to be more strongly related than perception of community involvement and NSAT awareness. One of the most interesting results was that none of these attitudes appear to be related to the Nova Scotia Achievement Tests results achieved by the subject school.

TABLE OF CONTENTS

	Page
Chapter 1	Introduction
A. Introduction.....	1
B. The Purpose of the Study.....	2
C. Rationale	4
D. Significance of the Study	13
E. Definitions and Limitations.....	14
Chapter 2	Review of the Related Literature
A. An Historical Perspective	18
B. The Accountability Concept in Education...	34
C. Limitations of Standardized Tests.....	36
D. Alternative Assessment.....	40
E. Outcome Based Education.....	43
F. Indicator Systems.....	44
G. Criteria for Quality Assessment.....	50
H. The Nova Scotia Achievement Tests.....	52
I. Perceptions of Accountability.....	54
J. Summary.....	58
Chapter 3	Research Design and Methodology
A. Purpose of the Study.....	59
B. Conceptual Framework.....	60

C. Methodology.....	61
D. Research Instrumentation.....	62
E. The Pilot Study.....	67
F. Profile of Sample.....	69
G. Procedures.....	76
H. Ethical Safeguards.....	77

Chapter 4 Presentation and Discussion of Findings

A. Introduction.....	79
B. Community Involvement.....	83
C. Identifying Outcomes.....	94
D. Measuring Outcomes.....	101
E. Accountability.....	110
F. Awareness of NSATs.....	119
G. Utility of NSATs.....	126
H. Accountability for NSAT results.....	134
I. Relationships Among Indices.....	142
J. Relationships to NSAT results.....	147
K. Discussion of Findings.....	148

Chapter 5 Recommendations and Conclusions

A. Summary of Findings	153
B. Attitudes and Perceptions	154
C. Perceptions of NSATs	157

	D. Demographic Variables and Attitudes	158
	E. Relationships Among Indices	159
	F. Recommendations for Future Research	160
	G. Recommendations for Practitioners	162
References	163
Appendices	175
Appendix A	Permissions	176
Appendix B	Questionnaire	177
Appendix C	Pilot Questionnaire	178
Appendix D	Item Correlation Matrix	179
Appendix E	Nova Scotia Achievement Test Results	180

LIST OF TABLES

Table	Page
4.1 Responses to Community Involvement Cluster "A"	83
4.2 Correlation of Statements in Cluster "A" to IACI	86
4.3 Gender and IACI	87
4.4 Category Differences in Gender and IACI	87
4.5 Position and IACI	88
4.6 Category Differences in Position and IACI	88
4.7 Level Taught and IACI	90
4.8 Category Differences in Level Taught and IACI	90
4.9 Experience and IACI	91
4.10 Subject Taught and IACI	92
4.11 Category Differences in Subject Taught and IACI	92
4.12 Teaching License and IACI	93
4.13 Responses to Identifying Outcomes Cluster "B"	94
4.14 Correlation of Statements in Cluster "B" to IAIO	97
4.16 Gender and IAIO	98
4.16 Position and IAIO	98
4.17 Level Taught and IAIO	99
4.18 Experience and IAIO	99
4.19 Subject Taught and IAIO	100
4.20 Teaching License and IAIO	100
4.21 Responses to Measurability Cluster "C"	101

Table	Page
4.22 Correlation of Statements in Cluster "C" to IAMO	104
4.23 Gender and IAMO	105
4.24 Category Differences in Gender and IAMO	105
4.25 Position and IAMO	106
4.26 Category Differences in Position and IAMO	106
4.27 Level Taught and IAMO	107
4.28 Experience and IAMO	108
4.29 Category Differences in Experience and IAMO	108
4.30 Subject Taught and IAMO	109
4.31 Teaching License and IAMO	109
4.32 Responses to Accountability Cluster "D"	110
4.33 Correlation of Statements in Cluster "D" to IAA	112
4.34 Gender and IAA	113
4.35 Category Differences in Gender and IAA	113
4.36 Position and IAA	114
4.37 Category Differences in Position and IAA	114
4.38 Level Taught and IAA	115
4.39 Experience and IAA	116
4.40 Category Differences in Experience and IAA	116
4.41 Subject Taught and IAA	117
4.42 Teaching License and IAA	118
4.43 Category Differences in License and IAA	118
4.44 Responses to Awareness Cluster "E"	119
4.45 Correlation of Statements in Cluster "E" to IPA	121

Table	Page
4.46 Gender and IPA	122
4.47 Position and IPA	122
4.48 Level Taught and IPA	123
4.49 Category Differences in Level Taught and IPA	123
4.50 Experience and IPA	124
4.51 Category Differences in Experience and IPA	124
4.52 Subject Taught and IPA	125
4.53 Teaching License and IPA	125
4.54 Responses to NSAT Utility Cluster "F"	126
4.55 Correlation of Statements in Cluster "F" to IPU	129
4.56 Gender and IPU	130
4.57 Position and IPU	130
4.58 Level Taught and IPU	131
4.59 Category Differences in Level Taught and IPU	131
4.60 Experience and IPU	132
4.61 Subject Taught and IPU	132
4.62 Teaching License and IPU	133
4.63 Responses to NSAT Accountability Cluster "G"	134
4.64 Correlation of Statements in Cluster "G" to IPAC	137
4.65 Gender and IPAC	138
4.66 Category Differences in Gender and IPAC	138
4.67 Position and IPAC	139
4.68 Level Taught and IPAC	139
4.69 Experience and IPAC	140

Table	Page
4.70 Subject Taught and IPAC	140
4.71 Teaching License and IPAC	141
4.72 Relationships Among the Indices	142
4.73 Stronger Relationships Among Indices	144
4.74 Weaker Relationships Among Indices	146
4.75 Relationship of Indices to NSAT Results	147

LIST OF FIGURES

Figure	Page
3.1 Gender of Respondents	70
3.2 Position Held by Respondents	71
3.3 Experience	72
3.4 Level Taught	73
3.5 Subject Taught	74
3.6 Teaching License	75
3.7 Return Rate by School	76

CHAPTER I

Overview and Purpose of Study

Introduction

For those involved in public school education, the last ten (or more) years have been characterized by change and uncertainty. In this climate of uncertainty, one thing is clear. There is a dramatic increase in public concern about the outcomes of schooling and the extent to which public school systems are held accountable by the public they serve.

Parents, business, industry, politicians, legislators, are demanding to know if public schools are doing what they purport to do. Are teachers teaching effectively, are students learning, and if so, where is the evidence? Demands include both the specific identification of the desired outcomes of learning and objective, quantifiable data to indicate the extent to which these outcomes are being met. Much of the literature refers to a "crisis in confidence" due mainly to uncertainty about the central mission and purpose of public school education.

Accountability almost always means assessment and assessing schools and school systems requires some form of standardization of assessment among the schools. Underlying this is the assumption that uniform testing will improve the quality of teaching and that students will ultimately benefit. There is also an assumption that the outcomes of education can be identified and measured and that schools have a direct responsibility for the measurement obtained. To a large extent, this study is about teacher's responses to these assumptions.

Purpose of the Study

The purpose of this study is to explore educator's perceptions of educational accountability, the role of standardized tests, and the association that exists among components of an accountability relation.

In particular, an investigation was carried out to determine educators attitudes toward several concepts included in an accountability relationship such as (1) public involvement in education, (2) identification of educational outcomes, (3) measurability of educational outcomes, and (4) accountability

for educational outcomes. In addition, the study was conducted to measure the strength of the relationships among these concepts. Attitudes were examined to determine the predispositions of teachers toward standardized tests of achievement in general and Nova Scotia Achievement Tests in particular the extent to which these attitudes are related to each other and to the NSAT results. To fulfil the purpose of this study, it includes the following:

1. An examination of the historical development of Standardized Tests of Achievement (NSAT's in particular) in view of recent testing reform and renewed public interest in educational outcomes.
2. A determination of attitudes (among teachers and administrators in one Nova Scotia School District) toward public involvement in education, identification of outcomes, measurability of outcomes and accountability.
3. A determination of perceptions among teachers and administrators in one Nova Scotia School district regarding awareness, utility and accountability for the results of Nova Scotia Achievement Tests.
4. A determination of the relationship among various

components of the accountability concept and the relationship to NSAT results.

5. A discussion of the implications that these attitudes and perceptions (and relationships) have for future use of standardized achievement tests.

Rationale

Renewed interest in identifying and measuring desirable educational outcomes is prevalent both locally and globally. In the small communities of Nova Scotia as well as international educational communities (particularly in North America) people are concerned about the "results" of public schooling. Professional literature, grassroots movements and the media provide ample evidence of the expanding interest in educational accountability.

In 1966 when James Coleman published his famous report on Equality of Educational Opportunity (Coleman et al:1966) he initiated a fundamental change in the way people perceive the quality of schools. Prior to his report, much of the focus had been on the inputs, resources and processes of education. The shift in interest to educational outcomes, largely begun by

Coleman et al, continues to influence educators and researchers today.

As educators and other stakeholders focus on the outcomes of education they have two distinct and formidable tasks; (1) they must more clearly define the goals and aims of education and the specific outcomes that are desired and (2) they must determine the extent to which these outcomes are being achieved.

Along with the requirement for public demonstration of the results of schooling, it is interesting that there exists a widespread attack on the use of standardized tests. There appears to be some inconsistency between the apparent mistrust of standardized tests of achievement and the increasing demand that they be used to insure accountability. Many educators and researchers have presented compelling arguments why we should not use standardized tests of achievement and others have defended their use.

Glickman (1990) summarizes much of the reaction to this debate...

we may disagree on the importance, validity, or priority of particular data, but who could argue that the schools and the public shouldn't know how their students are doing? To refuse to collect, disseminate and use data about students attitudes,

achievements and performances would be irresponsible (p.40)

Internationally, nationally and at the provincial level there is a great deal of material chronicling the perceived need for identifying educational outcomes and educational indicator systems.

The Nova Scotia Report of the Select Committee on Education (1992) cites the the United Nations World Report on Education (UNESCO)...

the content of schooling and the methods used in teaching students are too important to be left to teachers and must be brought in line through strongly interventionist policies by governments.

In fact the role of the teacher in systems of assessing students learning achievement is likely to emerge in the coming decade as a critical issue for both the status of the teacher within education generally and the credibility of the assessment practice itself. (p. 259)

United States

Although many educational reform movements were begun and developed in the United States, their influence has been felt throughout North America. In 1983, the report of The National Commission on Excellence in Education, "A Nation at Risk"

reported that the once unchallenged (American) preeminence in industry, science and technology was being "overtaken by competitors throughout the world". Passow (1990) cites the U.S. Department of Education observation that reports issued in 1983 had "created a tidal wave of school reform which promises to renew American Education" with "an extraordinary array of initiatives under discussion" and " a quantum increase in the variety of school activities involving leaders of the university, corporate and foundation communities"

In the 1960s and 1970s declines in SAT scores in the U.S. were widely publicized and provided much of the evidence that reformers needed to show that the education system was in trouble. Test scores began to be used more and more frequently to monitor educational progress and the effectiveness of educational systems. In the 1970s there was a widely held perception that large numbers of high school graduates were completing high school without the basic skills needed to function in society. Thus began MCT or minimum competency testing and the majority of states legislated some form of competency testing during this period.

In response to national concerns about the quality of education the NAEP (National Assessment of Educational Progress) began gathering data in 1969 and although the NAEP

was relatively low key during its early years, in recent years publications such as "The Nation's Report Card" in 1983 added considerably to its prominence. Comparisons among states was highlighted with the publication of the "wall charts" in 1983 which emphasized differences in "performance outcomes".

During the 1980s when many education reform movements were under way in the United States, there was an ever increasing demand from policy makers and legislators for data to refine and audit existing policy and to assist with the development of new policy. Edmond (1992) tells us that by 1988, " 43 states publicly reported data on school districts to state residents and 38 of those states include comparative data on individual school districts."

In 1984, the Education Commission of the States reported that the extensive activity at the state level almost always included these two themes, "more rigorous academic standards for students and more recognition and higher standards for teachers" (Pipho, 1986)

These two trends characterize reform during this period; a significant movement from local to state control and a significant increase in auditing and monitoring educational activity. Pressure continued at the national level to

encourage greater accountability at the state level and states continued to publish results and compare achievement.

Concerns about the results of public schooling continue into the present day. In 1991, President Bush's "America 2000" plan called for the monitoring of six national goals and placed great emphasis on accountability for reaching those goals. In 1993 the National Council on Education Standards (U.S.) stated that assessments and standards can become the cornerstone of education reform in the 1990s. Discussions concerning "outcome based" education, "measurement driven instruction", "authentic" assessment and accountability will continue for some years to come.

Canada

A recent report by the Economic Council of Canada (1992) records a number of indications that all is not well with the Canadian system of education. They cite enrolment percentages, dropout rates, literacy studies and international and inter-provincial comparisons as cause for concern. They acknowledge the difficult but vital role that teachers play in the educational process and the difficulty in evaluating teachers and teaching. They do, interestingly, offer a solution...

the quality (of teachers) can be assessed by the outcome of teaching - that is by the progress made by the students, provided proper allowance is made for the differential conditions that make teaching easier or more difficult...the progress of learning is the criterion that measures good teaching, and regular testing can help measure that progress (p.11)

In the conclusion of the same report, The Economic Council of Canada does not equivocate when it states:

Provincial authorities and school boards should provide continuous, comparable information on the performance of all schools in their jurisdiction on the basis of selected relevant educational indicators. Such information must be comparable across schools and through time, and must be made public (p.54)

The Council of Ministers of Education in Canada (CMEC) in 1988 approved the introduction of an education indicator system which is referred to as the School Achievement Indicators Program (SAIP) with the intention that the initial results (data gathered on 13 and 16 year olds) will be released in 1993. Edmond (1992) reports that the stated intention of the SAIP is " to compile the necessary information to provide a reliable profile of the comparative effectiveness of public schools in every province." (p.9)

After engaging in a process of consultation, reflection and inquiry for two years The Corporate Higher-Education Forum (of Canada) (1992) articulates three comprehensive recommendations. The first of these is "the establishment of national goals for education across Canada and the creation of indicators through which movement toward these goals can be measured and encouraged."(p.18)

Freedman (1993) states that "most knowledgeable Canadians agree that the quality of our k-12 system of education is seriously wanting" (p.6) and demonstrates the depth and breadth of the Canadian national concern regarding public education by citing studies from a variety of sources including; The Canadian Chamber of Commerce, The Business Council on National Issues, The Canadian Manufacturers Association, The Corporate-Higher Education Forum, The Conference Board of Canada, The Institute for Research In Public Policy, and The Science Council of Canada.

Nova Scotia

In March of 1992, The Report of the Select Committee on Education (1992) was tabled in the House of Assembly in the province of Nova Scotia. Section XVII of that report is entitled "Accountability and Standards". Among the

"Observations and Recommendations" documented in that section are the following from the section on Education Outcomes...

One of the difficulties we face in our Nova Scotia school system is that our curriculum is concerned with content and process but is not related to educational outcomes. We do not have specific and known standards against which to measure expected outcomes and desirable outcomes. In fact, it is questionable whether we know where we are. (p.264)

From the section on Assessments....

To talk about equal educational opportunities across the province, it is not sufficient to talk about the resource going into any particular system.....if results are not measured then it is absolutely impossible to tell whether an equal opportunity or access to education is being provided. (p.267)

To be truly accountable....the result of school wide achievement results and Board wide achievement test results should be mandated public information...

...one of the significant benefits of testing is to hold schools accountable. (p.268)

There is little doubt that at both the macro and micro level there is overwhelming interest in developing and measuring appropriate education outcomes and insuring that all government institutions (including public schools) are held accountable. One of the objectives of this study is to determine the status of this issue among teachers in Nova Scotia. How do teachers feel about the identifiability and measurability of educational outcomes? In what sense do

educators feel they should be held accountable? Does accountability imply assessment? Is there a legitimate role for standardized tests of achievement?

Significance of the Study

There is unquestionably a demand for greater accountability in public school education. When members of the public speak of accountability in education they (almost always) mean that they require clear evidence that students attending public schools are learning appropriately and that teachers are teaching effectively. It is important to know how educators (teachers and administrators) feel about the concept of accountability and to appraise their sense of responsibility for the outcomes of education. Do educators feel that the desired (cognitive) outcomes of education are identifiable, and if they are identifiable are they measurable, and if they are both identifiable and measurable to what extent are educators responsible for these outcomes. The significance of this study will be determined by the extent to which it contributes to the answers to these questions. In addition it will attempt to examine the role of standardized measures of achievement in this regard.

Definitions Limitations and Delimitations

Accountability is defined in a variety of ways. It requires that one determines; (1) who is accountable and who allocates this accountability, (2) to whom are they accountable, and (3) what are they accountable for?

The Oxford dictionary tells us to be accountable is "to be responsible for, or answerable for...". The Auditor General of the Province of Nova Scotia (1990) suggests that "accountability presumes the existence of at least two parties, one who allocates responsibility and one who accepts it, with the undertaking to report on the manner in which it has been discharged".(p.99)

With regard to accountability in education, Lessinger (1973) distinguishes among Performance Accountability, Professional Accountability and System Accountability. He concluded that Performance Accountability requires clearly defined specific objectives and "proceeds to close the gap between the learners entry level and the desired results". Instructional success is determined by direct measurement. Professional accountability is described as the extent to which educators know and use "skills, attitudes and techniques that research has shown to be reliable and valid in getting results". Under the heading

of System Accountability, Lessinger suggests that educational systems must "continually assess the achievement of pupils; relate their achievement to community goals, to the resources allocated...and to disseminate the findings to parents, teachers, taxpayers and other citizens of the community."

Accountability, for the purposes of this study, will be largely performance/system accountability. The following conjectures will be explored and accountability is defined in this context: (1) Accountability is inexorably linked with assessment (2) The outcomes of education are identifiable and measurable (3) Schools and school systems are responsible for identifying outcomes and monitoring progress toward these outcomes and (4) Schools and schools systems are responsible for reporting this progress to the citizens of school districts and are accountable to the citizens of the community which the schools serve.

There are many indicators of educational quality. In addition to academic goals, education has vocational, social, civic, cultural and personal goals. Judgements about the quality of schooling can be made by examining dropout rates, accessibility of programs, preparation for the labour market etc. Notwithstanding the importance of these indicators, this study will focus on student achievement. It is concerned with

the extent to which schools and school systems are held accountable for the academic (cognitive) achievement of students.

Outcomes

The outcomes that will be considered in this study are cognitive outcomes only. It does not include affective or psychomotor outcomes but only the "traditional" academic outcomes of schooling.

In this study, only standardized achievement tests will be considered. It will not be concerned with other types of standardized tests such as Standardized Aptitude Tests or Interest, Personality or Attitude Inventories. The reference to standardized achievement tests is not limited to multiple choice or norm referenced tests (Although the Nova Scotia Achievement Tests will be the particular test under study and they are both multiple choice and norm referenced). The study will include a brief exploration of alternate assessment including performance or "authentic assessment" and the discussion concerning alternatives to norm referencing.

This study will not take a position on the interesting debate

about whether we should administer standardized tests of achievement at all (although the results could influence that debate). It does assume that Standardized tests of achievement (not necessarily multiple choice or norm referenced) will be a reality that must be dealt with by public school educators in the foreseeable future. This study may offer some insights and provoke some thought and discussion with regard to the most appropriate manner in which to interpret and utilize the data obtained from these tests for accountability purposes.

The intention is to examine teachers perceptions of standardized tests in the context of accountability. This study is concerned only with the auditing of instruction and not the improving of instruction. Much has been written regarding the question whether standardized tests should "audit or improve", this study will deal with the "audit" side of this question only. Although clearly the diagnostic applications of standardized achievement tests and their use to inform curriculum and instruction are important topics, they do not fall within the scope of this study. Standardized achievement tests will be examined only for the purpose of assessing their function as an instrument for auditing or monitoring teaching and learning; and the extent to which they have utility in holding educators and educational policy makers accountable.

Chapter II

A Review of Related Literature

Testing Reform and Accountability - An Historical Perspective

Historical documents indicate that testing has been used in the United States as a policy tool since at least the 1870s. (Madeus and Kellaghan 1992). The best known of the early twentieth century educators concerned with testing was certainly R.L.Thorndike whose significant contributions to measurement technology are well documented.

Samuelson (1987) named Frederick Kelly as the inventor of the multiple choice test in 1914 and suggested that this development occurred partially in response to Starch and Elliot studies (1912, 1913) which showed that the marking of essay type questions was unreliable and unpredictable. The other obvious influences on the development of multiple choice tests were the growing numbers of students to be tested, the efficiency of scoring, as well the need for objectivity in the data collected. (Madaus and Tan, 1993)

Other significant developments during the early part of the twentieth century include Otis's development of a group

administered IQ test (used mainly with U.S. army recruits) and the emergence of the College Entrance Examination Board in 1926. It is interesting to note that these SAT initially had a writing component (a performance assessment) that was dropped in 1937 in order to allow greater variety as well as decrease the cost of the SAT's. (Madaus and Tan, 1993)

During the 1920's objective testing in the U.S. was flourishing and by the mid 1930's standardized testing had become a huge commercial enterprise and over half of the states in the United States had some form of statewide testing.

In 1932 Ralph W. Tyler, a professor at Ohio State university, in studying "conventional" verses "progressive" high schools may have been one of the first to view educational evaluation as an evaluation of program and not students. Tyler argued that a program should be judged simply by the extent to which it promoted student's mastery of the goals of the program. Popham (1988) states that " this objective based conception of evaluation, devised by Tyler in the 1930s, was destined to influence the view of subsequent generations of educators".

Madaus and Tan (1993) report that "the nature and magnitude of test use changed dramatically after World War II and each

succeeding generation witnessed an inexorable shift in the importance of testing as a major tool of educational policy." (p.53)

When the Russians launched Sputnik in 1957, this event led to heightened concern (some would say panic) about the quality of education in America. Popham (1988) reports "in the post - Sputnik soul searching of American policymakers, the schools became the splendid scapegoat candidates...it was clearly time to spruce up America's public schooling". The next year, in 1958 the U.S. enacted the "National Defence Education Act" which resulted in a flurry of activity to strengthen curricula (particularly in Mathematics and Science), improve the quality of teachers and develop new instructional resources. Madaus and Tan (1993) indicate the passage of this act "marks the emergence of testing as a tool in the national policy arena." (p.53)

The size and perceived importance of the NDEA projects led policymakers to fund evaluations of most of them. Educators found themselves without the tools to assess their own progress and abruptly became aware of the limitations of available measurement instruments. Worthen and Sanders (1991) report that "the resulting studies revealed the conceptual and methodological impoverishment of evaluation in that era" (p.4)

Popham (1988) reports that an important essay written in 1963 by Lee J. Cronbach entitled "Course Improvement Through Evaluation" unfortunately attracted little attention as "American educators of this era showed little interest in evaluation per se".

In 1965 the United States government enacted the Elementary and Secondary Education Act (ESEA). Up until this period, the financial support of schools (as well as their governance) was largely limited to state and local taxes rather than federal revenues. Popham (1988) reports that "such legislation emerged, in part, because of civil rights groups concerns as to how public schools were serving minority students". This bill (ESDA) provided extensive federal financial support for education and because of the magnitude of the federal funding involved, the debate over these expenditures extended to the U.S. Senate floor. Notably, Robert F. Kennedy (and others) insisted that each ESEA grant recipient file an evaluation report to determine if the expenditures had resulted in any real educational improvements. Concerns about evaluation (and accountability) suddenly became very important and many educators were left struggling to evaluate their own efforts and the need for new evaluation strategies become obvious. As summarized by Worthen and Sanders (1987) the best known contributors to the field of evaluation during this period

included Scriven (1967), Guba (1967), Stake (1967) and Stufflebeam (1968). This flurry of activity in the field of evaluation resulted in a variety of philosophical and methodological views but also some common ground. Worthen and Sanders (1991) reported " there was much agreement - evaluation is of necessity a multidimensional, pluralistic, situational and political activity that encompasses much more than simple application of the skills of the empirical scientist."(p.5)

In 1965 several states had testing programs in place but the results of these testing programs were never published on a statewide basis. Education was seen as a local concern, best controlled locally with little state interference. This would soon change. During the period that followed there occurred several developments that continue to influence testing programs today.

During the period of then President Johnson's War on Poverty James S. Coleman (et al) was commissioned (in 1964) to complete a report on equality of educational opportunity. His well known study which was published in 1966 minimized the effect of input and process variables on student achievement and forced educators to begin focusing on educational outputs and on the evaluation and assessment of outputs.

Following Coleman's report, educators had great difficulty accepting the notion that the variables they manipulated (the processes of education) had little effect on student achievement. Several studies contradicted Coleman's findings, notably Summers and Wolfe (1975) and serious efforts began to document the variables that affect student outcomes. At about the same time Jerome Bruner (1966) and Benjamin Bloom (1968) revived the idea that all students could learn if properly taught, and during the late seventies the Effective Schools Movement began to emerge.

During the sixties and early seventies, education reform was characterized by diversity, innovation and openness. Glickman (1990) suggests that during this period young people on the political left were attracted to education because they saw the opportunity to address social issues such as freedom, racial discrimination and winning the war on poverty. Schools and classrooms were open, innovative and unique - attempting to meet the needs of all social classes. There was a clear emphasis on processes and a minimum of emphasis on outcomes. About this time, student performance on SAT scores began to decline and this decline was given a great deal of media attention. Many people concluded that open education had been too permissive and as Glickman (1990) wrote "had sacrificed

the essential values of western civilization to meaningless relativism"(p.39) The widely held view was that the United States had lost its competitive edge in the world. In the latter seventies the U.S. political agenda had moved to the right and the operative word in education became "accountability". This accountability would be achieved through legislated policies. Education would no longer be the exclusive domain of the educators and local education authority. Legislative involvement in educational policies and practices grew and teacher morale deteriorated significantly. A general loss of respect for the profession and very little sympathy from the public added to this morale problem.

In 1968, seventy four different state testing programs were operating in forty two states but test results were never given to the public on a statewide basis or in any form that could be used to compare schools or districts.

Although it was conceived in 1963, the "National Assessment of Educational Progress" (NAEP) began operating in 1968. The goal of the NAEP was to provide a survey of knowledge, skills understanding, and attitudes at various grade levels in ten different subject areas. In addition to NAEP testing which was national in scope, by 1978 thirty three states had mandated minimum competency testing (MCT). Pipho (1978) describes MCT

as "legislative and legally mandated...which establish minimum levels of achievement required for all students..." This legislation usually required students to pass some state administered test to receive a high school diploma while in other states students had to pass a test for grade promotion. Bowers (1991) states that MCT laws were "often poorly conceived, established with inadequate funding and unrealistic time lines and showed little understanding of the complexities of the test development process." (p.56) At the state level MCT programs lost popularity (although many local Boards developed MCT requirements) and in 1986 only about half of the fifty states require students to pass a test to receive a high school diploma. (Bowers 1991). Apart from the obvious shift in control from local to state level, it is not clear what effect MCT has on students and schools.

Meanwhile, public concern for the quality of education increased and in 1983 a historic document from the National Commission on Excellence in Education (1983) was published entitled A Nation at Risk. This document was very critical of the condition of public school education and urged immediate reform. These "top-down" reforms were characterized by calls for higher standards, tougher course requirements, standardized curriculum and standardized testing. Glickman (1990) reports that following the release of A Nation at Risk,

accountability became the central tenet for a return to a traditional conception of schools. Now schools were supposed to show improvement in basic skills, stop fooling around with course electives and projects, and get back to drill, practice, homework, direct instruction and serious time-on-task schooling. (p.39)

Notable publications responding to the concern for American public school education during this era included; High School: A Report on Secondary Education In America (1983), by Ernest Boyer, Horace's Compromise: The Dilemma of the American High School (1984), by TheodoreSizer and John Goodlad's A Place Called School (1984). These works provide unflattering portraits of typical classroom routine and testimonials of what students don't know.

In 1985, following the publication of A Nation at Risk in the United States, the president of the Canadian Educational Researchers Association, Thomas Maguire (1985), reported on the significance (for Canadian educators) of accounts of the status of education in the U.S. He concluded that Canada is not yet a "nation at risk" for a variety of reasons including; (1) education is strongly entrenched as a provincial responsibility and (2) the timelag between the creation of ideas in the U.S. and their movement to Canada gives Canadians time to evaluate and select the best of these ideas. He did

however, make several important recommendations regarding the need for strong qualitative research to determine the influence that standardized testing programs have on educational practice and the need to develop instruments to assess higher order cognitive achievements.

Clarke (1993) reports that the philosophical basis for the reforms of this period was "the radical functionalist paradigm which included such concepts as competition, excellence, rewards and sanctions. Teachers were viewed as technicians and education was closely tied to the economy and the corporate sector." (p.7)

In May of 1984 the U.S. Department of Education (1984) reported that A Nation at Risk had resulted in a "tidal wave of reform" across the country. A "Survey of Education Commission of the United States" published "A Nation Responds" and gave an idea of the depth and breadth of the reform movements that were under way.

In 1984 the "Council of Chief State School Officers" (CCSSO) created the State Educational Assessment Centre and began (for the first time) the development of national education indicators of educational quality which called for student achievement test data to be collected by every state. Kaagan

& Smith (1985) reported on the state of development of national indicators of educational quality in the mid eighties. They define an educational indicator simply as a statistic that "provides information about the health of an educational system " (p.22) In order for an indicator to have meaningful policy implications, they say, it must be placed in a particular context. Establishing this context must be done by; (1) contrasting that statistic with an established standard or criterion level, (2) contrasting it with itself over time, or in two different systems or (3) contrasting it with other indicators in a cost benefit analysis. Kaagan and Smith go on to suggest that improved educational indicators could be used to monitor the quality of teaching staff, assess the impact of reforms and compare the U.S. system with other countries.

The basis of any indicator system are the statistics generated by information gathering processes including standardized achievement tests. Information is gathered over time and is available to policy makers as it becomes necessary.

In June 1985, at the annual convention of the National Education Association in the United States, seventy-five hundred delegates gave overwhelming approval to resolutions

which endorsed licensing examinations for new teachers and the dismissal of experienced teachers found to be incompetent. Wagner (1989) indicates that as both resolutions were "introduced by an NEA leadership long opposed to such measures" it represented a "victory for an accountability movement that had begun in the seventies". The emergence of this accountability movement Wagner suggests can be traced to three concerns (1) concern over the rising costs of education and in public services generally, (2) concern that the public school system is failing unacceptable numbers of youths (high dropout rates) and (3) the trend to use modern business as a model for school management.

Not everyone agreed with the movement toward accountability. Passow (1990) reports that legislated mandates increasing teacher and school accountability for student achievement are contributing to the "fragmentation of teachers, curricula, and teaching". He cites Ravitch(1985),Sizer(1985) and MacNeil(1987) whose work has found that many of the reforms meant to upgrade the quality of education have had the opposite effect.

In the latter part of the eighties, sympathy began to develop for the position that the top down mandates were inappropriate. In the United States there was some

decentralization of authority from the state to the local level. The National Governors Association (1988) declared that legislation had gone too far and some local control of education should be restored. The National Association of State Boards of Education (1988) supported this return to local (Board) control. The (U.S.) Office on Policy Research and Improvement (1988) suggested that "state and district policies should reflect greater freedom and accountability at the local and school level for decision making". No one was willing to abandon the notion of accountability, but perhaps it could be assured at the local level.

Clarke (1993) reports this period as one characterized by the re-professionalising of the teacher (giving them more authority to exercise their professional judgement but also making them more accountable for educational results) and more attention to the " social and economic contexts in which students learn." (p.8) The ideal was to "decentralize and democratize public schools" (p.8)

In 1991 President Bush's "America 2000" (1991) plan heavily emphasized accountability and monitoring of six national goals and data became available in 1991 when the National Education Goals panel produced its report. In the 1993 ASCD yearbook (Association for Supervision and Curriculum Development)

Madaus and Tan (1993) in their article "The Growth of Assessment", provide a comprehensive report on the emerging role of assessment in public school education. They mention (then) President Bush's stated intention to develop a national testing system geared to "world class standards" and the report the National Council on Education Standards and Testing (NCEST) which stated that "standards and the assessments linked to standards can become the cornerstone of the fundamental, systematic reform necessary to reform schools."

Wothen and Sanders (1991) list five trends that have characterized evaluation over the last twenty years. These trends include (1) the emergence of career opportunities in the field of evaluation, (2) the development of programs to prepare evaluators, (3) the institutionalization of evaluation in education, (4) the development of a distinct evaluation profession and (5) methodological developments in evaluation. As the cost of developing, administering and scoring tests increasingly became a factor, the role of technology became increasingly important. The dramatic decrease in the cost of "computing power" and in particular the development of the high speed optical scanner have had a significant impact on evaluation practices.

Fullen (1991) categorizes the reform (and accountability) movements in two waves, an "intensification wave" and a "restructuring wave". The "intensification wave" is characterized by increased government spending and control, reduced professional role for teachers, more specific definition and control of curriculum and greater demand for improved standardized test results. The more recent "restructuring wave" features more local autonomy and "collaborative cultures" and increased focus on outcomes. Fullen suggests that we should move beyond the regressive actions associated with the first wave and into the more appropriate actions associated with the second.

In spite of their falling into relative disfavour among educators, American polls indicate that public support for standardized testing is very strong. The 1992 Phi Delta Kappa poll of the public's attitude toward American public schools (Elam et al, 1992) revealed "overwhelming support for a national public school curriculum, for national goals and standards, and for a national testing program to measure progress toward these goals. Americans favoured standardized national tests by a margin of 77% to 14%, with only 9% undecided."

During the early 1990s educators continued to struggle with the accountability concept. Leblanc (1994) maintains that accountability is likely to be a more positive force in schooling when educators know why it has become an issue. She suggests that accountability trends in Canada have been the result of; (1) public perception of declining school effectiveness, (2) downturns in the economy, (3) a new global and knowledge based economy, and (4) public demand for a variety of services including inclusive education and technological skills.

Glickman (1990) maintains that the "open education movement gave educators choice but little responsibility...the accountability movement gave educators responsibility without choice." (p.41) He suggests that "open-accountable" schools of the nineties will embrace both pillars of reform; equal access to knowledge and public demonstration of results. The message to policymakers, he says is:

to hold schools accountable for achieving negotiated goals but not to legislate how they are to achieve such results ... schools that can not document success should reap the consequences of external systems of evaluation. (p.41)

The Accountability Concept in Education

Wagner (1989) states that " proponents of educational accountability hold the position that if educational institutions were less autonomous and more "accountable", greater harmony between their performance and the public interest would be served" but, he suggests the assumption that the nations schools could be better managed if those who teach in them were more accountable, is open to debate and subject to broad interpretation.

As reported in the introductory chapter, Lessinger (1973) distinguishes among performance, professional and system accountability. As indicated, this study is concerned with performance accountability to the extent that it explores attitudes regarding responsibility for the academic (cognitive) performance of students in public schools.

Wagner (1989) has conducted a philosophical inquiry into the nature of educational accountability. In formulating his definition of accountability he explains "standard usage holds that being accountable means, among other things, being obligated or subject to giving account, and in saying that a particular agent is accountable we imply that he is obligated to give a report, relation, description, explanation,

justifying analysis or some form of exposition..." The concepts of "obligation", "responsibility", and "entitlement" are highly relevant to any accountability relationship.

Sockett (1980) suggests that "to say an agent is accountable...is not merely to say he is able to deliver an account but that he is obliged to do so". Wagner goes much further in his development of the concept by exploring not only obligation, but also responsibility and entitlement... not only must there be a condition of obligation and responsibility to deliver an account, but since that account must be delivered to someone, the issue of entitlement must also be examined. The obligation to deliver an account, he suggests, may be either a legal or a moral obligation.

The argument of educational accountability proponents is based on the assumption that "educational excellences" can be defined in behavioral terms - that the behaviour can be observed and that the behaviour can be measured. The assumption is, that the preferred results of education can be identified, that these results can be expressed in behavioral terms, and that these behaviours are observable and measurable. Accountability proponents are focused on results and outcomes.

Wagner (1989), indicates that " the term "results" closely parallels what is generally meant when one speaks of educational "ends" or "outcomes", and these terms are used interchangeably in discussions of educational accountability." He goes on to suggest that these terms can be used in a prescriptive sense - to suggest a preferred state, or in a descriptive sense - to describe what has actually been achieved. Wagner makes the interesting observation that while accountability proponents are concerned with evidence of student performance, their demands always include a normative judgement - a pre-conception of what ought to be.

Limitations of Standardized Tests:

Madaus and Tan (1993) clearly state their view of the important role that testing will play in influencing educational policy in the next decade. They indicate "the main testing story of the last fifty years is the evolution (of testing) as a social technology, first to inform, and eventually to implement national education reform policy" (p.54)

In spite of the perception of increased importance, a number of researchers have found that, due to accountability pressures, testing has the potential to influence schooling in

a negative fashion. Teachers are forced to focus planning and instructional effort on test content and to focus more and more time to preparing students to do well on the tests (Smith and Rottenberg 1991, Kellaghan and Madaus 1991, Shepherd 1991).

Using standardized tests to hold teachers accountable clearly puts them in the "high stakes" category. Much has been written and researched about "high stakes" testing. The most serious problems are those concerned with the tendency to "teach to the test" because the potential consequences of not doing so are so great.

Lieberman (1991) maintains that policy discussions regarding testing for accountability purposes " have not dealt seriously with the harm that standardized testing may have already done ... ample evidence shows that while low level skills have improved, higher order skills have declined." (p.219) Making schools genuinely accountable, she says, will require that teachers be involved in developing alternate modes of assessment that measure what students know and are able to do.

Suarez and Gottovi (1992) document the concern that external assessments for accountability have taken precedent over all other forms of assessment. Their concerns also extend to

include narrowing of curriculum, less emphasis on higher order thinking, and over emphasis on academic achievement at the expense of intellectual, emotional and physical development.

Worthen (1993) noted that "the blue ribbon panel that issued "A Nation at Risk" noted that minimum competencies quickly became the maximum that schools attempted to attain, with the unintended consequence of lowering standards across the board." (p.445)

Many of the researchers conclude that time focused on test content has narrowed the curriculum, focused on lower order skills and neglected higher order thinking. More damaging is the conclusion that this narrowing is likely to be greatest in schools serving at risk students where there is even more pressure to improve scores (Herman, 1993).

In 1987 John Cannell, who was then a Medical Doctor in West Virginia, was surprised to learn that the (largely disadvantaged) students in his home state were scoring above the national average on a statewide assessment program. (Cannell, 1988) As all states began to teach to the test (due to accountability pressures) all scores improved. However, superficial changes in instruction to improve test performance do not result in meaningful learning (Shepherd, 1991).

When policymakers and others try to raise standards based on test results, "safety nets are strung up (in the form of exemptions, repeated trials, softening scores, tutoring for retests etc.) to catch those who fail" and further, "standards are determined by politically and economically acceptable pass rates, symbolic messages and appearances, and scarcely at all by analysis of competencies" (Glass and Ellwin, 1986).

Worthen and Spandel (1991) record seven criticisms levelled by opponents of standardized achievement tests, then they systematically answer (or at least expand on) each of the criticisms. On the charge that SATS narrow curriculum, for example, they suggest that educators should examine the decision making process (regarding curriculum content) and insure that it is not driven by the test. On the question of measuring only superficial (low level) skills and knowledge, they counter by simply stating that it does not have to be that way. "The notion that multiple choice tests can tap only recall is a myth. In fact the best multiple choice tests can and do measure students ability to analyze, synthesize information, make comparisons, draw inferences and evaluate ideas" (p.67) It is not entirely clear that multiple choice type tests should be summarily discarded.

Herman (1992) reports that " much of the research supporting the power of testing to influence schooling is based on traditional standardized tests and concludes that such tests have a negative impact on program quality."

Because of the limitations of "traditional" multiple choice standardized achievement testing, a brief examination of several other topics is warranted. These include (1) alternative assessment (2) outcome based education (3) indicator systems and (4) criteria for quality assessment.

Alternative Assessment

A great deal of attention has been given in recent years to increased efforts to develop alternatives to standardized testing. Entire issues of leading educational journals (Educational Leadership, Kappan, etc.) have been devoted to alternative assessment.

"Alternative assessment" is a generic term which is generally understood to include such labels as "direct assessment", "performance assessment", "authentic assessment", etc. All of these types of assessment are viewed as "alternatives" to traditional, multiple choice, standardized tests and all refer to " direct examination of student performance on significant

tasks that are relevant to life outside the school." (Worthen 1993, p.447)

Increasing criticisms of traditional standardized testing, demands for accountability and the negative consequences of high stakes testing, have influenced a number of alternative assessment proponents. While there is certainly a great deal of interest and support, some educators are concerned that such assessments are often initiated without adequate preparation. Worthen (1993) reports that "differences between proponents and opponents (of alternate assessment) have sparked vigorous debates...leaving many educators feeling rudderless as they attempt to chart an (assessment) course for their schools."(p.445)

Proponents argue that alternative assessment offers the opportunity to examine and judge a student's actual performance on significant relevant tasks (if we wish to determine if students can write, we should ask them to write). Sampling student performance (as on multiple choice tests) they contend, does not provide insight into how students would perform on "truly worthy intellectual tasks"

Opponents of alternative assessment practices are concerned that new assessment techniques cannot stand up to the same

level of criticism given to traditional tests. The objectivity, cost, and efficiency of alternative testing is being questioned. It is not clear if alternative testing (with a heavy dependence on teachers and uncertain technical quality) will be acceptable to various educational stakeholders who wish to assure accountability. Worthen (1993) lists several critical issues facing the future of alternative assessment. These include, (1) conceptual clarity, (2) mechanisms for self criticism and (3) support from well informed educators. Popham (1993), in an interesting article on circumventing the high costs of authentic assessment, suggests that "the most serious obstacle faced by proponents of authentic assessment is that their views will be dismissed as pie in the sky idealism by the very people that need to implement the proposed strategies."

Many educators, however, feel that authentic assessment is an idea whose time has come. Popham (1993) feels that "if authentic assessment becomes widespread, it will spawn a more appropriate instructional emphasis in our schools." Given the volume of literature devoted to the topic, it is an issue that will be debated extensively.

Outcome Based Education

As this study is concerned with perceptions about focusing on the outcomes of education, it may be appropriate to briefly examine an approach to teaching and learning that is based solely on outcomes.

Outcome Based Education can be defined in terms of four principles (1) all curriculum, instruction and assessment is focused on outcomes (2) students have a variety of ways and opportunities to learn and demonstrate learning (3) high expectations for everyone (all students will be able to significantly demonstrate learning) and (4) curriculum is designed back from where you want students to end up. (Brandt, 1992)

Bill Spady defines an educational outcome as a "culminating demonstration of learning" (p.67) and continues to explain that this demonstration of learning involves the substance, processes and settings of education. "Outcomes of significance require substance of significance, through processes of significance in settings of significance" (Spady and Marshall, 1991, p.67)

Spady and Marshall (1991) describe three forces that have resulted in a large educational following for this approach to teaching. Firstly, many schools and districts have recorded significant improvements in student learning using OBE. Secondly, public demand for improved student outcomes and educational accountability, and thirdly the perceived need for a new operating paradigm for education. "Successful outcomes are now both the starting point and the bottom lines of educational policy thinking in both Canada and the United States" (Spady and Marshall, 1991)

Indicator Systems

Many educators are including discussions of educational outcomes in the larger context of educational indicator systems. The underlying premise of indicator systems is that education is such a complex and diverse enterprise that it cannot be described by any single measure (such as achievement testing). An indicator system is, as the name implies, a system that will demonstrate the state of health of an educational enterprise. American accountability systems have been widely criticized because they have focused almost exclusively on student outcomes. An indicator system is not simply a student testing or assessment program. McEwan and

Chow (1990) describe an indicator system as " a much broader, more comprehensive undertaking which attempts to provide a more balanced picture of the operations and outcomes of schooling"

In 1984 (in the U.S.) the "Council of Chief State School Officers" (CCSSO) created the State Educational Assessment Centre and began (for the first time) the development of national education indicators of educational quality which called for student achievement test data to be collected by every state. Kaagan & Smith (1985) in the mid eighties, reported on the state of development of national indicators of educational quality. They define an educational indicator simply as a statistic that "provides information about the health of an educational system". In order for an indicator to have meaningful policy implications, they say, it must be placed in a particular context. Establishing this context must be done by; (1) contrasting that statistic with an established standard or criterion level, (2) contrasting it with itself over time, or in two different systems or (3) contrasting it with other indicators in a cost benefit analysis. Kaagan and Smith go on to suggest that improved educational indicators could be used to monitor the quality of teaching staff, assess the impact of reforms and compare the U.S. system with other countries.

Kaagan and Coley (1988) suggested that pressure to use indicator results to hold school systems accountable is premature largely because adequate funding is not available to ensure high quality for the measures that should be used and, they suggest, educators are reluctant to explore the critical relationships among inputs, processes and outcomes.

The topic of indicators of educational achievement has been receiving attention at the international level as well. The Organization for Economic Cooperation and Development is currently involved with the International Educational Indicators Project (IAEP). This project examined (1) enrolment, (2) outcomes, (3) functioning of schools, (4) costs and resources and (5) attitudes and expectations. Other international indicator projects include the International Association for the Evaluation of Educational Achievement (IEA) and the Third International Mathematics and Science Study. (TIMSS)

In Canada in 1988 the Council of Ministers of Education initiated an educational indicator system known as the School Achievement Indicators Project (SAIP). It consisted of three components; (1) participation, retention and graduation rates (2) expectations of and satisfaction with the education system and (3) achievement of literacy and numeracy among 13 and 16

year old students. This initiative is particularly important in Canada as there is no federal control over education and national information on education has been released only through the census, every four years (McEwan & Chow, 1990).

According to an School Achievement Indicators Program background paper (1991) released in December 1991, the stated objective of the SAIP is:

to provide a Canadian Information base that will assist ministries and departments of education to assess the performance of their education systems in relation to agreed upon criteria and to report indicators to the Canadian public, to assist each ministry and department to evaluate student achievement and to identify priorities in education, as well as to insure a high quality education for young Canadians. (p.1)

In developing the SAIP program the ministers of education recorded in a background paper (December, 1991) that they had encountered " resistance to large scale assessment from teacher groups across the country".(p.3) McConaghy (1991) comments that "even though teachers were to be involved in all aspects of the testing program, the Canadian Teachers Federation reaffirmed (in 1991) its opposition to standardized tests." (This study may shed some light on what it was that teachers were opposed to.)

In Alberta, Canada, twelve local school jurisdictions agreed in 1988 to work with the Alberta provincial department of education to develop Educational Quality Indicators (EQI) for that province. This project was guided by a four dimensional model of education consisting of partners (schooling, family, society), conditions (context, inputs, processes), outcomes (cognitive, affective, behavioral) and time (grades 3,6,9,12) (McEwan & Zatko, 1989). Three of the projects in the EQI program focused on accountability at the district level. These projects involved the development of a set of goals for learning, identification of outcomes, condition and process indicators, identification of ways to measure the indicators and interpreting the results. At a provincial meeting in January 1992, EQI partners recommended that any provincial indicator system should (1) focus on student outcomes, (2) be based on the goals of schooling and desirable personal characteristics, (3) represent a balance between cognitive, affective and behavioral outcomes, (4) provide corollary evidence (tests, surveys, interview records) and (5) provide multiple perspectives (student, staff, parents, stakeholders, public)

Edmond (1993) reports extensively on one indicator system developed in Winnipeg, Manitoba which focuses primarily of educational inputs and processes rather than outputs which

appears to be the focus of the SAIP initiative. He tells us that this distinction in focus reflects "the reality of confronting different levels of the education system...local school divisions must ensure a means to describe its local student population and organizational characteristics, while provincial and national policymakers are primarily concerned with results." He goes on to suggest that both of these perspectives would be represented in a complete indicator system.

Research is indicating that multiple indicators are necessary to measure the effectiveness of schools and schools systems and to reflect the complexities of schooling. Accountability implies a comprehensive, broad based integrated information system which provides a complete and inclusive picture of the health of education systems.

Bryk and Hermanson (1992) argue that the proper goal of indicator systems is to enrich public discourse. They suggest

indicators are not the exclusive domain of policymakers and system administrators seeking to exercise more effective external control over schools. Rather, these systems should be conceptualized as an effort in community education and crafted in ways that encourage a continuing, broad based public involvement with educational issues. (p.465)

Criteria For Quality Assessment

Much attention has been given in recent years to the articulation and documentation of appropriate assessment practices.

Herman (1992) contends that "educational assessment is in a process of invention, old models are being questioned and new models are in development", and continues to inquire about the components of good assessment. He lists some essential considerations including (1) the need to appraise the specific use and consequences of an assessment, (2) the need to appraise fairness (have all students had equal opportunity to learn and comprehend), (3) the need to assess the extent to which the results transferable and generalizable, (4) the need to assess the cognitive complexity of the assessment, (5) the need to assess the content quality and coverage, and (6) the need to assess meaningfulness and cost efficiency.

Quality assessment is inextricably linked to quality education. Writing in Education Canada, Common (1987) makes a number of critical points about evaluation and its relationship to the improvement of schooling; (1) evaluation is a critically important and complex dimension of public education, (2) what we deem quality in education is determined

not by the measurements we take, but in the judgements of worth we make about those measurements, (3) when evaluation is a public action, it becomes a political activity, (4) achievement tests are neither a complete indication of a student's performance or a teacher's performance and should not be used as the sole determinant of a teacher's accountability.

In 1993 a Joint Advisory Committee (Centre for Research in Applied Measurement and Evaluation) from the University of Alberta published a twenty one page booklet titled "Principles for Fair Assessment Practices for Education in Canada". Their documentation is divided into two parts, the first part deals with assessments carried out in classrooms by elementary and secondary teachers and the second part deals with standardized assessments developed external to the classroom.

With regard to standardized achievement tests, comprehensive guidelines are presented outlining what both test developers and test users should do in terms of four areas, developing and selecting methods of assessment, collecting and interpreting assessment information, informing students being assessed and implementing mandated assessment programs. The guidelines regarding externally mandated achievement tests (six pages) are complete and comprehensive.

The Nova Scotia Achievement Tests

Provincial testing programs in the province of Nova Scotia have a history which includes Provincial Exams, Nova Scotia Achievement Tests, Elementary Mathematics Program Assessment, and most recently the Elementary Language Arts Assessment. The most extensive of these are the Nova Scotia Achievement Tests which are given each year in October to students in grade nine and grade twelve. The NSATs include a series of tests in four subject areas: Language Arts, Mathematics, Social Studies and Science. The stated purpose of this testing program is to give some indication of program strengths and weaknesses (leading to program improvement) and to assist in the measurement of progress toward the goal of equal opportunity to quality education.

The Provincial Education Department has indicated that it hopes to improve the quality of this testing program by including in future reports (1) growth in achievement between the two levels tested, (2) coordination between testing and program implementation, (3) a writing sample in the Language Arts test, (4) further refinement of the expectations at each grade level, (5) more in depth reports and possibly (6) a "hands on" performance assessment in Mathematics or Science.

In April 1993 the Provincial Department of Education of the province of Nova Scotia published (for the first time) the results of the Nova Scotia Achievement Tests. (written in October of 1992) The "Background" written on the first page of that report states...

...the increased emphasis on testing has also created a different atmosphere - one in which business, industry and the public, as well as educators are interested in test results. In the spirit of accountability, the Department of Education believes Nova Scotians want to know and have a right to know what is expected of our students and how their achievement results compare with expectations... (p.1)

In the same report, the department of education identifies the following variables which have a statistically significant relationship to NSAT results; (1) average per capita income, (2) level of license held by teachers, and (3) average school size.

In November of 1992, the School Administrators Association of the Province of Nova Scotia requested that its membership report their views on standardized testing generally and the NSATs in particular. While he acknowledged the non-scientific nature of the results, Grant (1993) reported on the submissions received and his intent to "express the very real and broadly held concerns of school administrators from across the province." (p.1)

The concerns expressed regarding standardized testing (NSATs) included the following; (1) it is out of step with what is known about learning assessment and the goals of the public education system, (2) they assess only limited concepts of knowledge, skill, and intelligence, (3) students have no motivation to do their best work on NSATs, (4) it is too simplistic a process to measure the complexities of the educational process, (5) teacher's own assessments may be equally as valid, and (6) the uninformed use of test results may be dangerous to the public education system.

In documenting the responses to his report, Grant asks the question, "have we fairly determined the universe of educational outcomes which we should value...and can we reliably measure these outcomes once we have identified them..." (p.9) Among other objectives, this study attempts to report on the attitude of educators regarding these very questions.

Perceptions of Accountability

In this study it is assumed that it is important to know how educators feel about the various components of the accountability concept. It is assumed that perceptions and attitudes regarding accountability significantly influence the

potential for accountability trends to have a positive influence on public school education. Emerson said that "perception is fatal" and in terms of the future of "assessment for accountability", this may be exactly the case.

There is limited research available which documents teachers attitudes and perceptions of the elements of an accountability relationship but several studies have been done to do determine attitudes towards standardized testing.

In an interesting study Horber and Geisinger (1983) investigated the claim that "creative" people hold particularly negative attitudes toward multiple choice type tests. Their results did not provide support for this theory.

Green and Stager (1984) measured attitudes held by Wyoming teachers toward testing and test use. Generally attitudes towards classroom tests were favourable and attitudes towards standardized tests were unfavourable. Over ninety percent of the teachers in the study felt that standardized tests are not the best way to evaluate a teachers effectiveness and Forty percent felt that teachers do not understand standardized test results. In a later study, Green and Stager (1985) measured the attitudes of 555 teachers towards testing and the use of classroom, and standardized tests. They found that m es had

a more positive attitude toward classroom testing than females and (not surprisingly) that teachers who used tests most often had a more favourable attitude toward classroom testing. Attitudes toward standardized tests were described as negative or indifferent.

Soltz (1992) attempted to determine if the attitudes of teachers are a factor in the administration of externally mandated standardized tests. While he found that teacher's perceptions varied widely, they were not consistently related to their student's achievement and tests appeared to be administered in an objective fashion regardless of the teacher's attitude toward them. A study by Karmos (1984) indicated that results on some standardized tests are significantly related to student's perception of test importance and the use schools make of test results.

In order to determine perceptions of the purposes of standardized testing programs among testing directors, principals, supervisors and teachers, Marso and Pigge (1991) randomly selected ninety seven school districts in the U.S. They report that most educators, particularly elementary teachers and principals, perceive the most important purpose of standardized testing to be to support classroom instruction. Teachers and principals in secondary schools

differed from this view and ranked the guidance and counselling purpose of testing to be the most important.

Marso and Pigge (1992) also surveyed teachers in Ohio to determine perceptions of the " extent and effectiveness of their schools use of standardized test results." They found that teachers varied little between their extent and effectiveness ratings and generally perceived standardized test results to be used more for non-instructional rather than instructional purposes.

Green (1992) surveyed teachers in training and teachers in service and determined that preservice teachers had a less favourable attitude toward classroom testing (then teachers did) but held a more favourable attitude toward standardized testing.

Hall (1993) reports that a sample of two hundred and four Nova Scotia teachers indicates that two out of three feel that the Nova Scotia Achievement Tests are not an adequate indicator of student achievement and 84% indicated that the NSATs are not an adequate indicator of school effectiveness.

Summary

The literature review indicates that a great deal of attention has been devoted to the notion of accountability in education and the difficulties involved in assessing students for accountability purposes. Entire issues of leading educational journals have been devoted to the topic, books have been written and a great deal of research has been completed. Ten years after the release of "A Nation at Risk" assessment has undergone some evolutionary changes but many questions remain. The literature review included in this chapter touches on some of the salient features of a topic that is both voluminous and open ended and will continually be the subject of serious educational discourse. Educational assessment, as Herman (1992) suggests, continues " to be in the process of invention." (p.74)

Chapter III

Research Design and Methodology

Purpose of the Study:

The purpose of this study is to explore educator's perceptions of educational accountability, the role of standardized tests of achievement, and the association that exists among components of an accountability relation.

In particular, an investigation was carried out to determine the strength of the relationship between concepts such as (1) public involvement in education, (2) identification of educational outcomes, (3) measurability of educational outcomes, and (4) accountability for educational outcomes. Among the attitudes examined, are the predispositions of teachers toward standardized tests of achievement in general and Nova Scotia Achievement Tests in particular and the extent to which these attitudes are related to the above concepts and to the NSAT results.

Conceptual Framework:

For the purpose of this study, accountability (as defined in the introduction) was viewed as a concept that involved a number of sequential prerequisite concepts.

If schools and school systems are to accept the notion that they are accountable to the larger communities for specific educational outcomes, then they must believe that; (1) the public (school communities) have a legitimate role to play in the enterprise of education, (2) that educational outcomes are identifiable, (3) that the outcomes of education are measurable, and (4) that schools and school systems can be held to account for the measurement obtained.

Public Involvement

Identifiability of Outcomes

Measurability of Outcomes

Accountability for Measurement

Methodology:

Exploratory and descriptive research methods will be applied to this study in an attempt to describe teachers perceptions about, and attitudes toward, public involvement in education, identifying outcomes, measuring outcomes, and standardized achievement testing in the context of educational accountability.

In addition to the descriptive statistics, the major portion of the study is a comparative/correlational study which determines the strength of the relationships between teacher attitudes and dispositions toward achievement testing and accountability as well as other variables including test results, gender, subject taught, experience level, level taught, and license level. None of these variables are manipulated during the study.

Davis (1993) describes correlational studies; " Comparative and correlational studies use hypothesis directed observation to determine the differences between groups or the relationships holding among pre-identified but non-manipulated variables. The major features of comparative studies are that there are a large number of subjects, the purpose of the study is to determine the relationships holding among two or more

groups or variables, the variables measured are clearly specified, but the variables are not manipulated." (p.261)

While relationships are sought, there is no attempt to show causal relationships among the variables and no interventions will be recommended. If relationships are identified, further research may be recommended to determine cause.

Quantitative data was gathered by means of (1) examining school, district, and provincial records of the NSAT results, and (2) a questionnaire. Attitudes held by teachers and principals were measured using a Likert type attitudinal scale. Approximately 550 Junior and Senior High School teachers (grades 7-12) in Halifax County were questioned. In addition, forty four principals and vice principals completed the questionnaire. The questionnaire was distributed, completed and returned during the months of November and December 1993.

Research Instrumentation

The Questionnaire:

The first page of the questionnaire (appendix B) provided demographic information about the respondent. This information

included; (1) gender, (2) position, (3) level taught, (4) number of years teaching, (5) subject area, and (6) license level. Much of this demographic information (along with NSAT results) will be correlated with data gathered on the attitude (Likert type) scale in sections one and two.

Following the first page of information about the respondent, the questionnaire is divided into two sections. A Likert type scale assessed attitudes by asking respondents to indicate whether they strongly agree, agree, are undecided, disagree or strongly disagree with a series of statements.

There were seven areas in which attitudes were determined (1) public involvement in education (2) identifiability of educational outcomes (3) measurability of outcomes (4) accountability for outcomes (5) awareness of Nova Scotia Achievement Tests (6) utility of Nova Scotia Achievement Tests and (7) accountability for results of Nova Scotia Achievement Tests. Seven corresponding clusters of statements included both favourable and unfavourable statements about each attitude. Statements were randomly ordered to avoid a tendency to respond to similar items in similar ways.

SECTION ONE:

Section one provided data pertaining to attitudes toward the following:

(A) Public involvement in education:

(1) How do teachers and administrators feel about public involvement in public school education?

(2) Does community involvement impact on the quality of schools?

(B) Identifying outcomes:

(1) Should outcomes be identified, are they identified now and are they sufficiently clear?

(2) How important is it to identify outcomes?

(3) Should all teaching include a demonstration of learning?

(C) Measurability:

(1) Can we determine the effectiveness of schools and teachers by measuring outcomes?

(2) Are identified educational outcomes measurable?

(3) Should educators concern themselves with measuring outcomes?

(D) Accountability for outcomes:

(1) To what extent should school systems, schools and teachers be accountable for the performance of their students?

(2) Do teachers feel accountable for what they do?

(3) Can achievement testing be used as an accountability tool?

SECTION TWO:

Section two will gather data pertaining to attitudes toward the Nova Scotia Achievement Tests (NSAT's) including:

(E) level of awareness and understanding:

(1) Are teachers, administrators and parents aware of the purpose of NSAT's?

(2) Are teachers, administrators and parents aware of the results of NSAT's?

(3) Do teachers and administrators understand the purpose and results of NSAT's?

(F) level of utility and acceptability:

- (1) Are NSAT results of any value to teachers and administrators and are they concerned about them?
- (2) Are NSAT results a valid measure of educational outcomes.
- (3) Is the time spent on NSATs warranted?

(G) Accountability for results:

- (1) Do teachers and administrators feel accountable for NSAT results?
- (2) Should NSAT results be made public and should schools be compared?

Item/Attitude Correspondence on Pilot Questionnaire:

Attitudes toward...	As measured by ...
A. Community involvement	Items 1,5,9,12,13,17,22
B. Identification	Items 2,6,10,14,23,27,39,40
C. Measurability	Items 4,7,15,18,19,21,24,26 29,31
D. Accountability	Items 3,8,11,16,20,25,28,30 32,33,34,35,36,37,38
E. Awareness	Items 41,44,47,50,53,56 59,62,65,74

F. Utility	Items 42,45,48,51,54,57,60 66,68,69,72,73,75,76 78,80
G. Accountability	Items 43,46,49,52,55,58 61,64,67,70,71,77

The Pilot Study:

The pilot was administered on November 3, 1993 to a preliminary group of eighteen respondents. This entire pilot group consisted of principals and supervisors from the Musquodoboit/Sheet Harbour subsystem in Halifax County in Nova Scotia. The group required about sixteen minutes to complete the pilot questionnaire (appendix C). All eighteen respondents reacted to every statement on the scale. Following completion, the questionnaire was discussed with the group who gave valuable feedback on the wording of certain items, ambiguous items, repetitive items, grammatical considerations etc.

In order to identify the most appropriate items an item analysis was done by calculating the correlation coefficient for each item in a cluster against the mean for the cluster. The correlation coefficient (the Pearson R value) for each item in the original cluster (pilot) as well as for each item in the final cluster were calculated.

In addition to "R" values, frequency counts for each item as well as means and standard deviations were calculated. Based on these statistics, as well as feedback from the pilot group, items were discarded until the five most appropriate items for each cluster remained. The items that remained were examined closely to be certain that they would measure what was intended before the pilot was administered. As there were seven clusters, the final questionnaire contained thirty five items extracted from eighty items on the pilot scale.

The items that remained (as numbered on the pilot scale) after this process was applied are as follows:

Attitudes toward...	As measured by ...
A. Community involvement	Items - 1,5,12,13,22
B. Identification of outcomes	Items - 6,10,14,23,39
C. Measurability of outcomes	Items - 7,15,21,26,31
D. Accountability	Items - 3,8,33,37,38
E. Awareness (NSAT)	Items - 44,47,53,62,65
F. Utility (NSAT)	Items - 66,68,72,73,76
G. Accountability (NSAT)	Items - 43,55,58,67,70

**Item/Attitude Correspondence as numbered on Final
Questionnaire:**

Section I

A. Community involvement	Items - 1,3,5,6,9
B. Identification of outcomes	Items - 4,7,10,16,18
C. Measurability of outcomes	Items - 8,11,12,17,19
D. Accountability	Items - 2,13,14,15,20

Section II

E. Awareness (NSAT)	Items - 5,7,10,13,15
F. Utility (NSAT)	Items - 1,2,3,6,9
G. Accountability (NSAT)	Items - 4,8,11,12,14

The Sample

The sample consisted of 606 respondents from 23 different schools. The following graphs provide a profile of the sample.

Figure 3.1

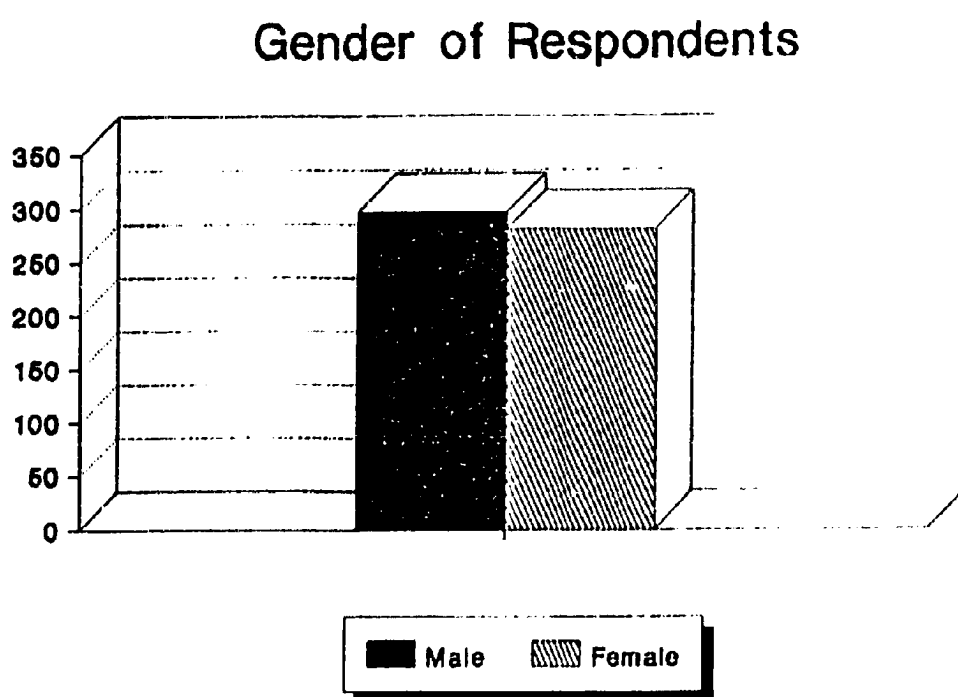


Figure 3.2

Position Held by Respondents

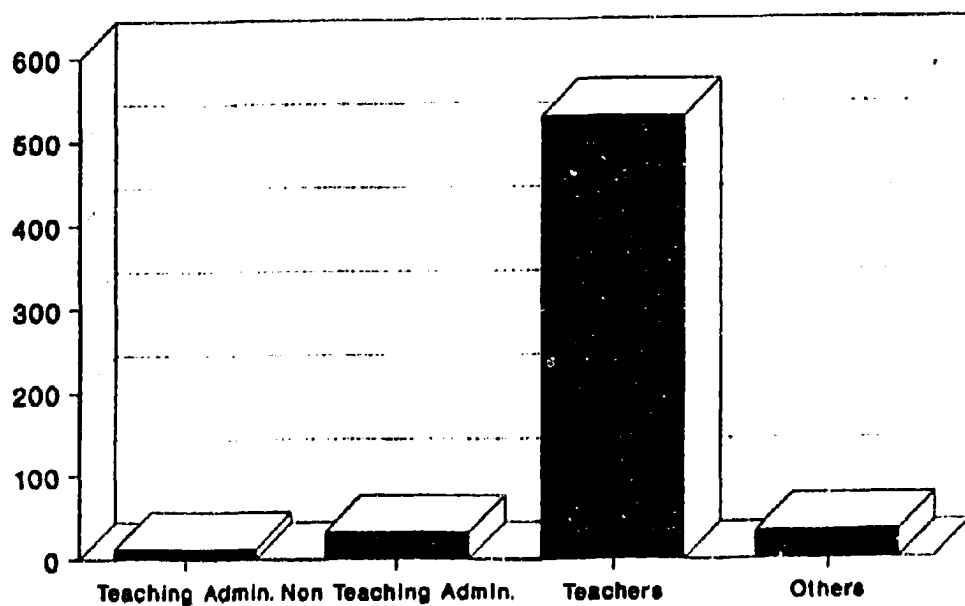


Figure 3.3

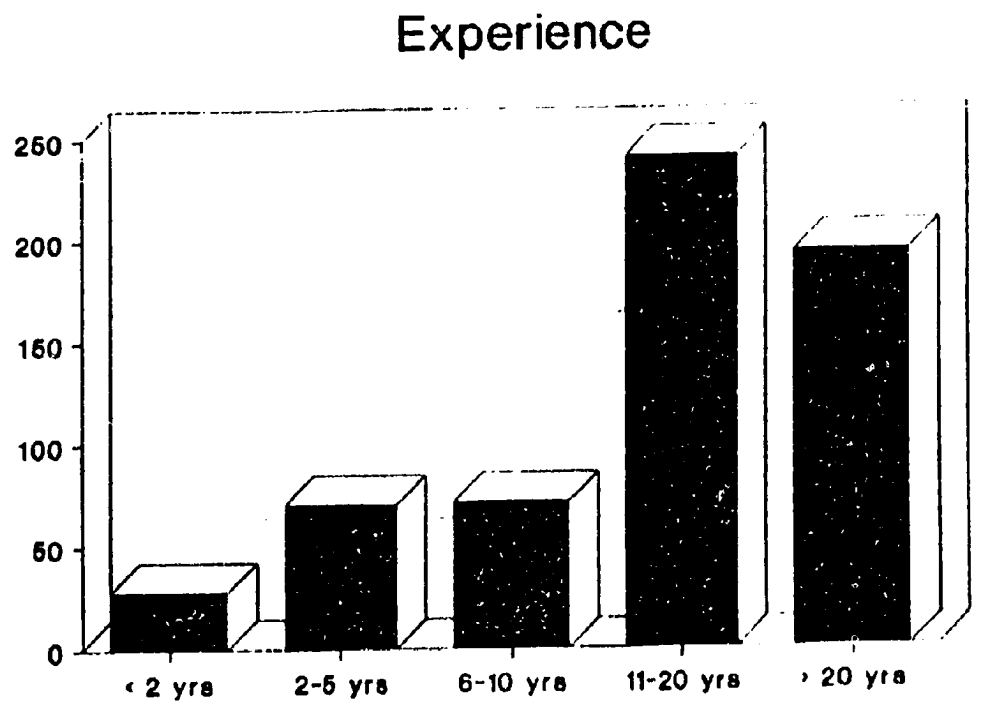


Figure 3.4

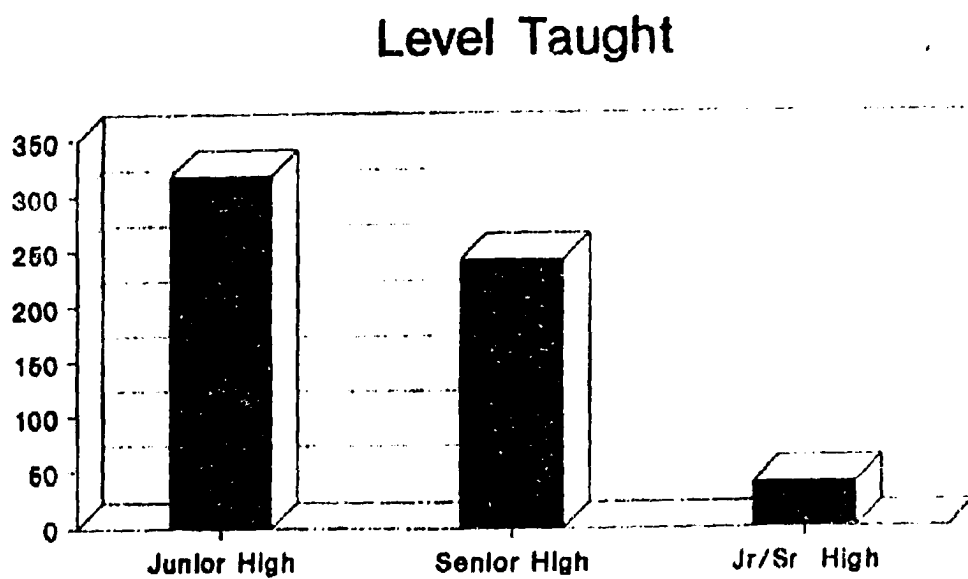


Figure 3.5

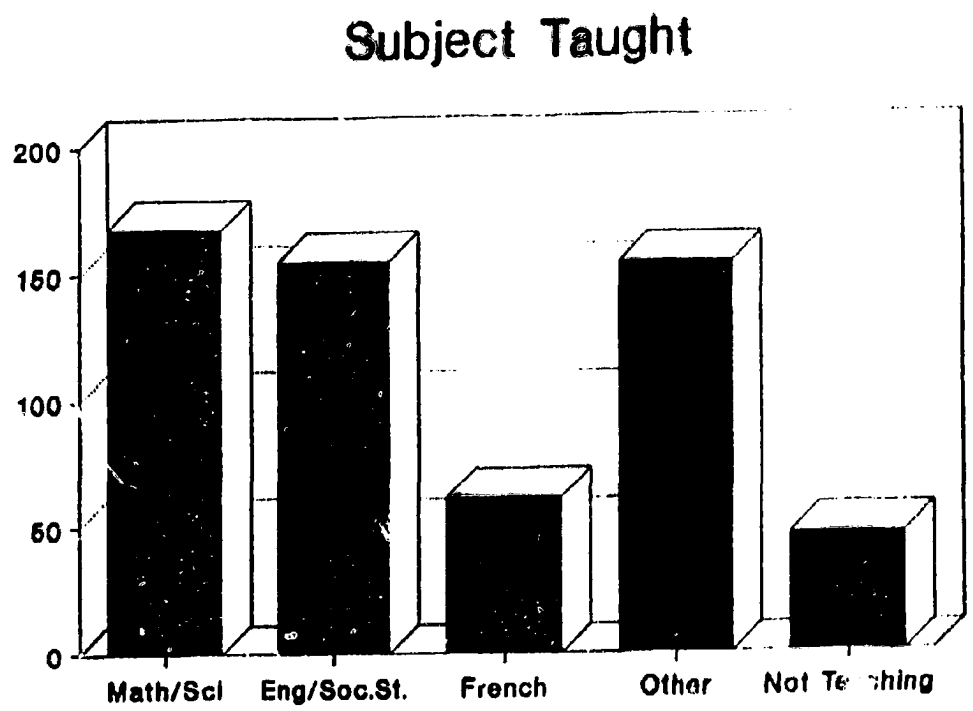
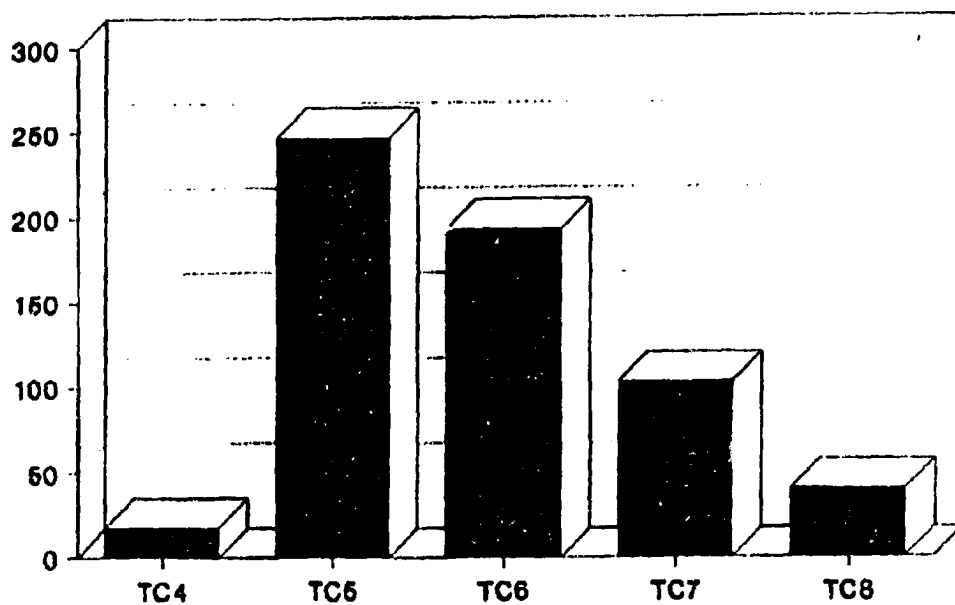


Figure 3.6

Teaching License

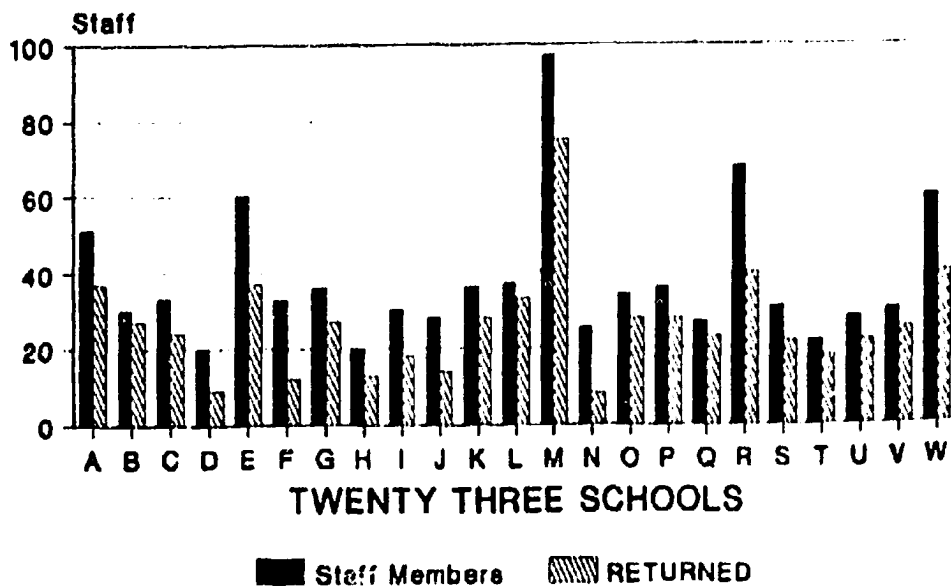


Procedures:

Six hundred and six junior high school and senior high school teachers and school based administrators in one Nova Scotia School District were included in this study which was carried out during the months of November and December 1993. The instrument used to conduct the study (described below) was delivered to twenty three schools including fifteen junior high schools and eight high schools. In most cases it was completed at one sitting (usually at a staff meeting) but in other cases it was handed out to teachers to be returned. As expected, the return rate was much higher for those teachers who completed the questionnaire at one sitting. The chart below provides an overview of the number, type and relative size of the schools participating, as well as the return rate.

Figure 3.7

Return Rate by School



The questionnaires required approximately ten minutes to complete and were picked up at the schools by the researcher. Overall 606 out of 870 or 70 per cent (70%) of the questionnaires were returned.

Descriptive and comparative statistics were calculated using an available statistical package (Statsview) on a home microcomputer.

Ethical Safeguards:

As some of the data gathered is of a sensitive nature, every precaution was taken to insure confidentiality and anonymity. Teachers, schools and areas were identified by number only with no possibility of tracing completed questionnaires to their specific source. Appropriate supervisory personnel at the school district under study have been contacted and permission has been granted to proceed with the survey. Results of the study will be presented by school, sub-system or other aggregate, never by individual respondent.

Permission and Approvals:

Permission to conduct the research was requested and obtained from the Chief Executive Officer as well as the Superintendent of Research and Planning of the selected Nova Scotia School District. (Appendix A) A summary of the study will be given to the selected school district upon completion.

Chapter 1V

Presentation and Discussion of Findings

This chapter contains a description of the responses to seven different clusters of statements which measure attitudes toward aspects of; (1) community involvement in education, (2) identifying educational outcomes, (3) measuring educational outcomes, (4) educational accountability, (5) awareness of Nova Scotia Achievement Tests, (6) utility of Nova Scotia Achievement Tests and (7) accountability for the results of Nova Scotia Achievement Tests.

In addition to the descriptive statistics (frequency counts, means and standard deviations) used to delineate the responses, comparative statistics are used to describe relationships between the category variables (gender, experience etc.) and the indices for the various clusters mentioned above. Comparative statistics include correlation coefficients (Pearson R), analysis of variance (ANOVA) and tests of significance.

Responses to each of the seven clusters are examined separately and the indices calculated for each cluster are compared to the category variables, to each other and to the results of the Nova Scotia Achievement Tests for each of the

twenty three schools involved in the sample.

The statements below appear exactly as they did in the questionnaire including the abbreviations "NSATs" for the "Nova Scotia Achievement Tests" and "HCBDSB" for the "Halifax County Bedford District School Board".

It is important to note that respondents were asked to "react" to the statements. They were instructed that they were not required to "know the answer" but only their reaction was sought. Respondents were instructed that if they felt the statement was unclear or they did not understand the statement, they should choose "undecided" as their response. For this reason, isolated statements for which there was no response (there were very few) are included in the "undecided" category. It was impossible to determine if respondents who choose "undecided" understood the statement perfectly and were undecided about their response, or if they did not understand the statement, found it vague or ambiguous or simply refused to react to it for other reasons. Also, it is acknowledged that some of the statements are compound, some are open to various interpretations and that some of the terms used in the statements are not clearly defined. It is possible that some respondents are reacting to an unintended interpretation of the statement or an unintended definition of words within the statement. These limitations do restrict the interpretation of

the responses. What follows is a description of the responses as they occurred.

This chapter is structured in the following way:

The responses to each of the statements in seven different clusters are examined, followed by the calculation of an index to summarize the responses to each cluster. A table is presented to indicate the extent to which each statement in the cluster is related (Pearson R) to the calculated index.

Then an analysis of variance, and in some cases a simple regression, were completed to determine if relationships existed between the calculated indices and the various category variables included in the study.

A table is presented to indicate the index for each category along with levels of significance. If significant differences are found between any of the categories, a second table is shown which shows the results of the tests of significance that were completed.

Following these analyses, the various indices are examined to determine their relationship to each other and are presented in a correlation matrix.

Finally, school scores on the various indices are examined to determine if any association exists between the various indices and the results of the Nova Scotia Achievement Tests for each school.

The effect sizes for any significant differences are included in the discussion of findings at the end of the chapter.

The statements are clustered under the following headings:

A. Attitude toward Community Involvement	IACI
B. Attitude toward Identifying Outcomes	IAIO
C. Attitude toward Measurability of Outcomes	IAMO
D. Attitude toward Accountability	IAA
E. Perception of Awareness of NSATs	IPA
F. Perception of Utility of NSATs	IPU
G. Perception of Accountability for results of NSAT	IPAC

A. Community Involvement

Table 4.1

Responses to Community Involvement Cluster "A"

(N=606)		SA	A	U	D	SD
1	" Community involvement in public school education is entirely appropriate."	37%	48%	8%	6%	1%
2	" The public has a genuine interest in, and a right to know, what schools are doing and how successful they are."	30%	56%	9%	5%	.5%
3	" Community involvement has no significant impact on the quality of schools."	1%	5%	10%	60%	24%
4	" It is appropriate that the public should audit the performance of public schools. "	5%	43%	28%	19%	5%
5	" Public interest in education is generally limited to times when schools are " under attack"	24%	51%	9%	15%	1%

Table 4.1 contains a summary of the responses to five statements measuring attitudes toward community involvement in education. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

Overwhelmingly, respondents felt that community involvement in education was appropriate. Eighty five percent either agreed or strongly agreed with statement one. There was very little resistance to this suggestion, only four of 606 respondents strongly disagreed with the statement.

Statement Two

While this was a compound statement, only about 9% of the respondents were undecided about it. 36% agreed (or strongly agreed) that the public interest was genuine and that they have a "right" to know how successful schools are. This is consistent with the response to statement one.

Statement Three

Eighty four percent of the respondents disagreed (or strongly disagreed) with the suggestion that community involvement does not affect the quality of schools. This response indicates that, not only did the respondents feel that community involvement was appropriate, but also that it has a real and significant impact on quality. It is not certain that

respondents felt that the impact would be positive but based on responses to statements one and two it is reasonable to conclude that a positive impact was implied. These first three statements indicate very strong support for developing community school partnerships.

Statement Four

The purpose of this statement was to determine if the respondent felt that public involvement should be extended to include an auditing function. While a large percentage of the respondents were undecided about this statement (28%), twice as many respondents agreed (48%), as disagreed (24%) with the notion that the public should audit schools. The large percentage of undecided may be due to the fact that there are various interpretations of the word "audit".

Statement Five

The purpose of this statement was to determine if respondents felt that community involvement was limited to difficult times. Three quarters of the respondents agreed (or strongly agreed) with this statement, indicating that while they felt that public involvement in education was appropriate, it may be limited to periods when schools are being censured. This must be balanced with the responses from statement two where most respondents indicated that the public interest in education was genuine. It is to be expected that interest would be heightened during times of duress.

Index of Attitudes toward Community Involvement:

Responses to these five statements on community involvement were converted to scores with appropriate weighting on favourable and unfavourable items. The mean of these scores was calculated to determine an index of attitudes toward community involvement (IACI) for each respondent. The higher the IACI, the more positive the attitude toward community involvement in schools. To determine the strength of the relationship between each statement and the IACI, the correlation coefficient (Pearson R) for each of the five statements was calculated.

Table 4.2

Correlation of Statements in Cluster "A" to IACI

Statement	Correlation to IACI
One	.598
Two	.578
Three	.51
Four	.493
Five	.600

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of attitudes toward community involvement (IACI) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

**Relationships of Category Variables to Attitudes Toward
Community Involvement:**

Table 4.3
Gender and IACI

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.552	.495	.029
Female	281	3.525	.488	.029

Table 4.4
Category Differences in Gender and IACI

Comparison	Mean Diff.	Fish. PLSD	F Test
Male vs. Female	0.028	0.08	0.456

No significant difference in attitude toward community involvement was found between male and female respondents.
($p=.4977$)

Table 4.5
Position and IACI

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.498	.468	.02
Teach/Admin	12	3.667	.492	.142
NT/ Admin.	32	3.913	.597	.105
Other	33	3.818	.483	.084

Table 4.6
Category Differences in Position and IACI

Comparison	Mean Diff.	Fish. PLSD	F Test
Tchr vs T/Ad	-.169	.273	.489
Tchr vs NT Admin	-.414	.17 *	7.607 *
Tchr vs Other	-.32	.168 *	4.672 *
T/Ad vs NT Admin	-.246	.317	.774
NT Admin vs Other	.152	.315	.297

* Significant at 95%

Non teaching administrators (most of principals and vice principals in the sample) and "others" (which includes mostly guidance counsellors, resource teachers etc.) have a significantly more favourable attitude toward community involvement than the teachers in the sample.

Table 4.7
Level Taught and IACI

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.599	.469	.026
Sr. High	242	3.461	.497	.032
JH/SH	40	3.58	.526	.083

Table 4.8
Category Differences in Level Taught and IACI

Comparison	Mean Diff.	Fish. PLSD	F Test
JRHI vs SRHI	.138	.081 *	2.754 *
JRHI vs JH/SR	.019	.16	.013
SRHI vs JH/SH	-.119	.163	.513

* Significant at 95%

Respondents in Junior High schools have a significantly more positive attitude toward community involvement than did respondents in senior high school.

Table 4.9
Experience and IACI

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.65	.42	.079
2-5 years	71	3.49	.437	.052
6-10 years	72	3.511	.436	.051
11-20 years	241	3.552	.479	.031
> 20 years	194	3.538	.546	.039

There were no significant differences in attitudes toward community involvement based on years of experience.

($p=.6374$, $F=.635$). Simple regression yields $R = .001$

Table 4.10
Subject Taught and IACI

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.495	.501	.039
Eng/Soc.St.	155	3.508	.454	.036
French	62	3.523	.447	.057
Other	154	3.529	.466	.038
Not Teaching	47	3.783	.54	.079

Table 4.11
Category Differences in Subject Taught and IACI

Comparison	Mean Diff.	Fish. PLSD	F Test
MA/SC vs N/T	-.288	.155 *	3.332 *
EN/SS vs N/T	-.275	.156 *	2.979 *
French vs N/T	-.26	.181 *	1.986
Other vs N/T	-.254	.156 *	2.554 *

* Significant at 95%

Only non teaching respondents were significantly different among the various subject groups. They held a significantly more positive attitude toward community involvement.

Table 4.12
Teaching License and IACI

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	3.567	.567	.134
TC5	247	3.517	.493	.031
TC6	193	3.533	.481	.035
TC7	103	3.586	.453	.045
TC8	40	3.585	.579	.092

No significant differences in attitudes toward community involvement were found related to teaching license.
($p=.7515$, $F=.479$). Simple regression yields $R = .046$

B. Identifying Educational Outcomes

Table 4.13

Responses to Identifying Outcomes Cluster "B"

(N=606)		SA	A	U	D	SD
1	" Identifying the outcomes of education should be at least as important as the inputs and processes of education."	29%	57%	11%	3%	.5%
2	" Teachers should focus on the process of education and not the results."	6.6%	24%	17%	44%	9%
3	" All teaching should include a culminating demonstration of learning."	7%	53%	24%	14%	2%
4	" Every attempt should be made to identify the specific outcomes of instruction in each of the core areas. "	9%	69%	15%	6%	1%

Table 4.13 contains a summary of the responses to four statements measuring attitudes toward various aspects of the identification of educational outcomes. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

The response to this statement indicates strong support for the importance of identifying the outcomes of education. Eighty five percent of the respondents felt that identifying educational outcomes was at least as important as identifying the processes and inputs of education, and only three percent of the respondents disagreed with this suggestion.

Statement Two

More than half of the respondents to this statement feel that the results of teaching deserve at least as much (and possibly more) attention as do the processes. About one third feel that the process of education deserves more attention.

Statement Three

While there were a large number of undecided, sixty percent of the respondents agreed that teaching should conclude with a demonstration of learning. In this statement the large number of undecided may be due to a lack of a clear understanding

about the meaning of the phrase "culminating demonstration". This phrase is borrowed from the proponents of "outcome based education" who insist that the "culminating demonstration" is necessary.

Statement Four

More than three quarters of the respondents felt it was necessary to identify the specific educational outcomes in the core areas. Since the statement says that "every attempt should be made", this may imply that respondents feel that these outcomes are not now specifically identified.

Index of Attitudes Toward Identification of Outcomes (IAIO):

Responses to four statements relating to the identification of outcomes were converted to scores with appropriate weighting on favourable and unfavourable items. The mean of these scores was calculated to determine an index of attitudes toward identification of outcomes (IAIO) for each respondent. The higher IAIO, the more positive the attitude toward the importance of identifying (and clarifying) educational outcomes. The correlation coefficient (Pearson R) for each of the four items against the IAIO was calculated.

Table 4.14
Correlation of Statements in Cluster "B" to IAIO

Statement	Correlation to IAIO
One	.616
Two	.625
Three	.587
Four	.626

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of attitudes toward identification of outcomes (IAIO) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

**Relationship of Category Variables to Attitudes Toward
Identifying Outcomes:**

Table 4.15
Gender and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.679	.52	.03
Female	281	3.66	.537	.032

No significant difference in attitudes toward identifying educational outcomes was found between male and female respondents. ($p=.6631$, $F=.19$)

Table 4.16
Position and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.647	.532	.023
Teach/Admin	12	3.667	.268	.077
NT/ Admin.	32	3.812	.421	.074
Other	33	3.75	.606	.105

There were no significant differences in attitudes toward identifying outcomes based on position held. ($p=.2689$, $F=1.314$)

Table 4.17
Level Taught and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.641	.507	.028
Sr. High	242	3.711	.519	.033
JH/SH	40	3.537	.639	.101

No significant differences in attitudes toward identifying outcomes were found to be related to level taught.

($p=.2999$, $F= 1.223$)

Table 4.18
Experience and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.714	.547	.103
2-5 years	71	3.63	.581	.069
6-10 years	72	3.566	.462	.054
11-20 years	241	3.64	.542	.035
> 20 years	194	3.728	.504	.036

There were no significant differences in attitudes toward identifying outcomes related to experience. ($p=.17$, $F= 1.6$)
Simple regression yields $R = .055$.

Table 4.19
Subject Taught and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.668	.534	.041
Eng/Soc.St.	155	3.59	.587	.047
French	62	3.73	.49	.062
Other	154	3.67	.477	.038

No significant differences in attitudes toward identifying outcomes were found related to subject taught.

($p=.3885$, $F=1.035$)

Table 4.20
Teaching License and IAIO

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	3.708	.464	.109
TC5	247	3.672	.517	.033
TC6	193	3.633	.536	.039
TC7	103	3.68	.57	.056
TC8	40	3.669	.492	.078

No significant differences in attitudes toward identifying outcomes were found related to license held. ($p=.9232$, $F=.227$)

Simple regression yields $R = .006$

C. Measuring Educational Outcomes

Table 4.21
Responses to Measurability Cluster "C"

(N=606)		SA	A	U	D	SD
1	" It is possible to devise a student assessment scheme that will provide significant information about the effectiveness of a teacher."	3%	28%	37%	23%	9%
2	" The cognitive outcomes of education can and should be measured."	12%	63%	20%	5%	3%
3	" It is not possible to quantitatively measure the effectiveness of schools."	2%	23%	28%	42%	4%
4	" It is possible to devise a student assessment scheme that will provide significant information about the effectiveness of a school. "	3%	43%	33%	18%	3%
5	" Most of the important outcomes of education are not measurable."	9%	33%	22%	34%	3%

Table 4.21 contains a summary of the responses to five statements measuring attitudes toward various aspects of measurability of educational outcomes. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

Respondents are not at all sure if teacher effectiveness can be measured through student assessment. This statement generated a even distribution of responses with a large number (almost forty percent) undecided about this issue. This statement requires more detail and specificity to provide more discriminating responses.

Statement Two

While one in five was undecided, about three quarters of the respondents felt that education outcomes are not only capable of being measured but should be measured. This is consistent with the statement in the previous section where about three quarters of the respondents suggested that specific outcomes should be identified. Only two respondents strongly disagreed with this statement.

Statement Three

While a large number (28%) are undecided about quantitatively measuring school effectiveness, at least forty five percent feel that a quantitative measure is possible and about twenty five percent feel that a quantitative measure is not possible.

Statement Four

This statement is very similar to statement three. As with the previous statement, a large number (one third) were undecided about measuring school effectiveness through student assessment, but far more (46%) agreed that it could be done than disagreed (21%).

Statement Five

There appears to be a fairly even distribution of responses to this statement. The "important" outcomes in this statement are not restricted to cognitive outcomes and this statement is open to broad interpretation. This ambiguity may have contributed to the range of responses.

Index of Attitudes Toward Measurability of Outcomes:

Responses to these five statements on measurability were converted to scores with appropriate weighting on favourable and unfavourable items. The mean of these scores was calculated to determine an index of attitudes toward measurability of outcomes (IAMO) for each respondent. The higher IAMO, the more positive the attitude toward the ability to, and the perceived importance of, measuring educational outcomes. The correlation coefficient (Pearson R) for each of the four items against the IAMO was calculated.

Table 4.22
Correlation of Statements in Cluster "C" to IAMO

Statement	Correlation to IAMO
One	.638
Two	.559
Three	.657
Four	.645
Five	.615

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the IAMO and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

Relationship of Category Variables to Attitudes Toward Measurability of Outcomes:

Table 4.23

Gender and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.255	.568	.033
Female	281	3.198	.583	.035

Table 4.24

Category Differences in Gender and IAMO

Comparison	Mean Diff.	Fish. PLSD	F Test
Male vs. Female	.057	.094	1.411

No significant difference in attitude toward measurability of outcomes was found between male and female respondents.

($p=.2353$, $F=1.411$)

Table 4.25
Position and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.199	.561	.024
Teach/Admin	12	3.133	.695	.201
NT/ Admin.	32	3.5	.693	.122
Other	33	3.418	.535	.093

Table 4.26
Category Differences in Position and IAMO

Comparison	Mean Diff.	Fish. PLSD	F Test
Tchr vs T/Ad	.066	.327	.052
Tchr vs NT Admin	-.301	.204 *	2.807 *
Tchr vs Other	-.219	.201 *	1.533

* Signifioant at 95%

Non teaching administrators have a significantly more favourable attitude toward measurability of outcomes then do the teachers in the sample.

Table 4.27
Level Taught and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.213	.576	.032
Sr. High	242	3.262	.563	.036
JH/SH	40	3.175	.57	.09

No significant difference in attitudes toward measurability of outcomes were found to be related to level taught.

($p=.2441$, $F= 1.367$)

Table 4.28
Experience and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.343	.588	.111
2-5 years	71	3.105	.53	.063
6-10 years	72	3.128	.546	.064
11-20 years	241	3.216	.572	.037
> 20 years	194	3.303	.591	.042

Table 4.29
Category Differences in Experience and IAMO

Comparison	Mean Diff.	Fish. PLSD	F Test
2-5 yrs vs >20 yrs	-.198	.156 *	1.564
6-10 yrs vs >20 yrs	-.175	.155 *	1.237

* Significant at 95%

Respondents with over twenty years experience have a significantly more positive attitude toward measurability of outcomes than those with two to five years experience.

Table 4.30
Subject Taught and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.221	.5	.039
Eng/Soc.St.	155	3.161	.617	.05
French	62	3.262	.516	.066
Other	154	3.208	.589	.047

No significant differences in attitudes toward measuring outcomes were found related to subject taught. ($p=.186$, $F=1.5$)

Table 4.31
Teaching License and IAMO

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	3.278	.509	.12
TC5	247	3.219	.557	.035
TC6	193	3.178	.573	.041
TC7	103	3.276	.612	.06
TC8	40	3.34	.587	.093

No significant differences in attitudes toward measuring educational outcomes were found related to teaching license ($p=.4174$, $F=.981$). Simple regression yields $R = .044$

D. Accountability

Table 4.32
Responses to Accountability Cluster "D"

(N=606)		SA	A	U	D	SD
1	" Teachers themselves are the most significant contributors to the quality of instruction in their classrooms."	29%	52%	6%	12%	2%
2	" Achievement test results can be used to compare the quality of instruction in schools with similar characteristics."	3%	34%	23%	30%	11%
3	" Recent interest in educational accountability is largely political and bureaucratic."	19%	46%	14%	19%	2%
4	" Current trends calling for more specific data on student achievement will ultimately damage the educational process. "	4%	18%	29%	45%	4%

Table 4.32 contains a summary of the responses to four statements measuring attitudes toward various aspects of accountability. The table shows the percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with the statements.

Statement One

As they have identified themselves as the most significant contributors, the response to this statement might indicate that teachers feel that they could be called to account for the quality of instruction in their classrooms. Eighty one percent agreed that they were the most significant contributors to that quality.

Statement Two

Although teachers feel responsible for the quality of instruction in their classrooms, they are uncertain about that quality being measured by achievement testing. Responses were widely dispersed.

Statement Three

Responses to this statement indicate some degree of cynicism regarding recent interest in educational accountability.

Statement Four

Again, a large number of undecided (almost thirty percent) but among the decided, a strong majority (by a margin of two to one) feel that the educational process will not be damaged by demands for more specific data on student achievement.

Index of Attitudes Toward Accountability:

As with previous clusters, responses to these five statements related to accountability were converted to scores. The mean of these scores determined an index of attitudes toward accountability (IAA) for each respondent. The higher IAA, the more positive the attitude toward various aspects of accountability. The correlation coefficient (Pearson R) for each of the four items against the IAA was calculated.

Table 4.33
Correlation of Statements in Cluster "D" to IAA

Statement	Correlation to IAA
One	.403
Two	.597
Three	.626
Four	.617

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of attitudes toward accountability (IAA) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

Relationship of Category Variables to Attitudes Toward Accountability:

Table 4.34
Gender and IAA

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.166	.562	.033
Female	281	3.065	.578	.034

Table 4.35
Category Differences in Gender and IAA

Comparison	Mean Diff.	Fish. PLSD	F Test
Male vs. Female	.101	.093 *	4.519 *

* Significant at 95%

Male respondents have a significantly more positive attitude toward the statements dealing with accountability than did female respondents.

Table 4.36
Position and IAA

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.09	.547	.024
Teach/Admin	12	3.208	.858	.248
NT/ Admin.	32	3.438	.632	.112
Other	33	3.288	.638	.111

Table 4.37
Category Differences in Position and IAA

Comparison	Mean Diff.	Fish. PLSD	F Test
Tchr vs T/Ad	-.118	.324	.172
Tchr vs NT Admin	-.348	.202 *	3.815 *
Tchr vs Other	-.198	.199	1.274
T/Ad vs NT Admin	-.299	.375	.48
NT Admin vs Other	.15	.275	.381

* Significant at 95%

Non teaching administrators have a significantly more favourable attitude toward accountability than did the teachers in the sample.

Table 4.38
Level Taught and IAA

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.122	.571	.032
Sr. High	242	3.124	.553	.036
JH/SH	40	3.2	.589	.093

No significant difference in attitudes toward accountability was found to be related to level taught.

Table 4.39
Experience and IAA

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.188	.703	.133
2-5 years	71	3.085	.503	.06
6-10 years	72	3.031	.646	.076
11-20 years	241	3.078	.537	.035
> 20 years	194	3.214	.573	.041

Table 4.40
Category Differences in Experience and IAA

Comparison	Mean Diff.	Fish. PLSD	F Test
6-10 yrs vs >20 yrs	-.183	.154 *	1.363
11-20 yrs vs >20 yrs	-.136	.107 *	1.549

* Significant at 95%

Respondents with more than twenty years experience had a more positive attitude toward accountability than did respondents in the six to twenty year range.

Table 4.41
Subject Taught and IAA

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.094	.52	.04
Eng/Soc.St.	155	3.11	.578	.046
French	62	3.125	.476	.061
Other	154	3.114	.614	.049

No significant differences in attitudes toward accountability were found related to subject taught.

($p=.9079$, $F=.253$)

Table 4.42
Teaching License and IAA

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	3.056	.61	.144
TC5	247	3.037	.583	.037
TC6	193	3.166	.534	.038
TC7	103	3.184	.558	.055
TC8	40	3.263	.628	.099

Table 4.43
Category Differences in Teaching License and IAA

Comparison	Mean Diff.	Fish. PLSD	F Test
TC5 vs TC6	-.128	.107*	1.387
TC5 vs TC7	-.147	.131*	1.221
TC5 vs TC8	-.225	.19*	1.355

* Significant at 95%

Generally, attitudes toward accountability became more positive with increase in teaching license. The significant differences were found when level five licenses were compared with the three higher license levels. ($p=.0327$)

E. Awareness of Nova Scotia Achievement Tests

Table 4.44

Responses to NSAT Awareness Cluster "E"

	(N=606)	SA	A	U	D	SD
1	" Most parents of grade nine and grade twelve students are not aware of their child's NSAT scores."	13%	50%	23%	12%	2%
2	" Parents want to know about their child's achievement (in school) relative to other students in the province."	7%	54%	23%	14%	2%
3	" Teachers in HCBDSB are generally aware of the purpose of the NSATs."	3%	42%	23%	29%	3%
4	" Teachers in HCBDSB are generally aware of the results of the NSATs."	2%	45%	21%	28%	4%
5	" Teachers do not understand enough about the NSATs to make any judgments about their usefulness."	5%	34%	26%	32%	3%

Table 4.44 contains a summary of the responses to five statements measuring perceptions of awareness of the Nova Scotia Achievement Tests. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

Generally, it appears to be the perception of teachers that parents are not aware of NSAT results. A large number were not sure about the level of parents awareness.

Statement Two

While many are undecided, teachers generally feel that parents wish to compare their child's achievement in school to the achievement of others in the province. Only about fifteen percent disagree with this notion.

Statement Three

There appears to be some uncertainty among respondents regarding the purpose of NSATs. There was a fairly even distribution of responses, but slightly more respondents perceive an awareness of the purpose of NSATs.

Statement Four

Less than one half of the respondents indicated that they felt that teachers were aware of the results of NSATs.

Statement Five

An even distribution of responses to this statement indicates a range of opinion regarding levels of understanding of NSATs.

Index of Perceptions of Awareness of NSATs:

As with other cases, responses to these five statements on awareness were converted to scores with appropriate weighting on favourable and unfavourable items. The mean of these scores was calculated to determine an index of perceptions of awareness (IPA) for each respondent. A higher IPA indicates a perception of greater general awareness of purpose and results of NSATs. The correlation coefficient (Pearson R) for each of the four items against the IPA was calculated.

Table 4.45

Correlation of Statements in Cluster "E" to IPA

Statement	Correlation to IPA
One	.505
Two	.407
Three	.678
Four	.673
Five	.577

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of perceptions of awareness (IPA) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

Relationship of Category Variables to Perceptions of Awareness of NSATs:

Table 4.46
Gender and IPA

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.01	.524	.03
Female	281	3.022	.555	.033

No significant difference in attitude toward perceptions of awareness of NSATs was found between male and female respondents. ($p=.79$, $F=.071$)

Table 4.47
Position and IPA

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.014	.53	.023
Teach/Admin	12	3.00	.426	.123
NT/ Admin.	32	3.025	.603	.107
Other	33	3.055	.662	.115

Differences in opinion of awareness of NSATs were not significantly related to position held. ($p=.9785$, $F=.065$)

Table 4.48
Level Taught and IPA

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.062	.537	.03
Sr. High	242	2.965	.538	.035
JH/SH	40	2.98	.524	.083

Table 4.49
Category Differences in Level Taught and IPA

Comparison	Mean Diff.	Fish. PLSD	F Test
Jr.High vs Sr.High	.096	.09 *	1.104

* Significant at 95%

Respondents at the junior high school level perceived greater awareness of NSATs than those at the senior high level.

Table 4.50
Experience and IPA

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.036	.493	.093
2-5 years	71	2.935	.478	.057
6-10 years	72	2.861	.484	.057
11-20 years	241	3.083	.546	.035
> 20 years	194	3.021	.565	.041

Table 4.51
Category Differences in Experience and IPA

Comparison	Mean Diff.	Fish. PLSD	F Test
2-5 yrs vs 11-20	-.148	.142 *	1.044
6-10 yrs vs >20 yrs	-.16	.145 *	1.165
6-10 yrs vs 11-20	-.222	.141 *	2.38 *

* Significant at 95%

With some exceptions, perceptions of awareness generally increased with experience. The significant differences occur after ten years experience.

Table 4.52
Subject Taught and IPA

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.007	.495	.038
Eng/Soc.St.	155	3.022	.608	.049
French	62	3.0	.46	.058
Other	154	3.032	.512	.041

No significant differences in perceptions of awareness were found related to subject taught. ($p=.9923$, $F=.064$)

Table 4.53
Teaching License and IPA

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	2.967	.446	.105
TC5	247	2.978	.52	.033
TC6	193	3.039	.543	.039
TC7	103	3.058	.576	.057
TC8	40	3.095	.562	.089

No significant differences in perceptions of awareness were found related to license held by respondents. ($p=.523$, $F=.804$)

F. Utility of Nova Scotia Achievement Tests

Table 4.54

Responses to NSAT Utility Cluster "F"

(N=606)		SA	A	U	D	SD
1	" The NSATs are a valid measure of many (but not all) of the desired outcomes of education."	2%	38%	27%	27%	6%
2	" While multiple choice type tests are limited, they do provide useful information about student achievement."	5%	61%	17%	13%	3%
3	" The interruption (at grade 9 and grade 12) caused by administering NSATs is not warranted"	5%	19%	29%	42%	5%
4	" We should continuously attempt to improve the NSATs and gain greater understanding of their potential uses."	12%	55%	19%	11%	3%
5	" School authorities should not spend their time investigating discrepancies in NSAT scores among schools."	10%	36%	26%	26%	2%

Table 4.54 contains a summary of the responses to five statements measuring perceptions of the utility of the Nova Scotia Achievement Tests. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

While more than a quarter of the sample was undecided about NSATs as a measure of desired outcomes, forty percent agreed with the statement while thirty five percent disagreed. There appears to be no consensus of opinion on this issue.

Statement Two

Generally (two thirds) respondents agreed that the NSATs do provide limited, but useful information. This statement is very similar to the previous statement. It may indicate greater support for multiple choice type testing because the term "limited" is used.

Statement Three

While many are undecided (29%), almost half of the respondents feel that the interruption caused by NSATs is warranted. This is a stronger statement for support for NSAT testing than statement two where only about forty percent felt that the NSATs were a valid measure.

Statement Four

To provide for greater understanding and improvement, a wide majority of respondents (two thirds) felt that NSATs warrant continued attention .

Statement Five

There was no strong consensus about the question of administrators spending their time investigating discrepancies in NSAT scores. Slightly more respondents felt that they should not.

Index of Perceptions of Utility of NSATs:

Responses to the five statements on utility of NSATs were converted to scores with reverse weighting where appropriate. The mean of these scores was calculated to determine an index of perceptions of utility (IPU) for each respondent. A higher IPU indicates a perception of greater utility of NSATs. The correlation coefficient (Pearson R) for each of the four statements against the IPU was calculated.

Table 4.55

Correlation of Statements in Cluster "F" to IPU

Statement	Correlation to IPU
One	.700
Two	.602
Three	.728
Four	.685
Five	.672

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of perceptions of utility (IPU) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

Relationship of Category Variables to Perceptions of the Utility of NSATs:

Table 4.56
Gender and IPU

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.243	.663	.038
Female	281	3.216	.645	.038

No significant difference in perceptions of utility of NSATs was found between male and female respondents. ($p=.6144$)

Table 4.57
Position and IPU

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.231	.645	.028
Teach/Admin	12	3.017	.663	.191
NT/ Admin.	32	3.456	.616	.109
Other	33	3.17	.781	.136

No significant differences in perception of utility of the NSATs were found related to position. ($p=.147$, $F=1.794$)

Table 4.58
Level Taught and IPU

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.294	.637	.036
Sr. High	242	3.183	.65	.042
JH/SH	40	3.135	.701	.111

Table 4.59
Category Differences in Level Taught and IPU

Comparison	Mean Diff.	Fish. PLSD	F Test
Jr. High vs Sr. High	.112	.109 *	1.1015

* Significant at 95%

Respondents at the junior high level perceived greater utility for the NSATs than respondents at the senior high.

Table 4.60
Experience and IPU

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.236	.76	.144
2-5 years	71	3.183	.659	.078
6-10 years	72	3.197	.582	.069
11-20 years	241	3.208	.641	.041
> 20 years	194	3.303	.673	.048

No significant differences in perceptions of utility of NSATs were found to be related to experience. ($p=.5249$, $F=.801$)
Simple regression yields $R= .052$

Table 4.61
Subject Taught and IPU

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.289	.586	.045
Eng/Soc.St.	155	3.195	.691	.056
French	62	3.2	.547	.069
Other	154	3.26	.652	.053

No significant differences in perceptions of utility were found related to subject taught. ($p=.4786$, $F=.875$)

Table 4.62
Teaching License and IPU

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	2.944	.556	.131
TC5	247	3.248	.599	.038
TC6	193	3.256	.656	.047
TC7	103	3.196	.753	.074
TC8	40	3.24	.715	.113

No significant differences in perceptions of NSATs were found to be related to teaching license held. ($p=.3781$, $F=1.055$)
Simple regression yields $R = .009$

G. Accountability for Nova Scotia Achievement Test Results

Table 4.63

Responses to NSAT Accountability Cluster "G"

(N=606)		SA	A	U	D	SD
1	" Schools and school systems should be held accountable for their performance on NSATs."	3%	21%	32%	34%	10%
2	" It may be useful to compare schools in Nova Scotia using NSAT results."	2%	38%	25%	27%	8%
3	" The quality of classroom teaching has a significant effect on NSAT results."	3%	35%	29%	27%	6%
4	" It is appropriate that NSAT results be made public for schools and school systems but not for individual students."	5%	27%	22%	36%	11%
5	" A serious attempt should be made to find out why some schools score consistently low on NSATs."	9%	58%	17%	13%	3%

Table 4.63 contains a summary of the responses to five statements measuring perceptions of accountability for the results of the Nova Scotia Achievement Tests. The table shows the (approximate) percentage of respondents who strongly agree (SA), agree (A), are undecided (U), disagree (D), or strongly disagree (SD) with each of the statements.

Statement One

Although responses were widely dispersed, most respondents disagreed with the idea that schools should be held accountable for NSAT results. The phrase "be held accountable" is open to broad interpretation.

Statement Two

While there were a large number of undecided, and a fairly even distribution of responses, more respondents agreed with this statement than disagreed indicating some level of support for the notion that schools could be compared using NSAT results.

Statement Three

Again, while responses were fairly evenly distributed, slightly more agreed than disagreed that the quality of teaching had an effect on NSAT scores. This statement was not intended to include "teaching to the test" as "quality

teaching" but some respondents may have interpreted it that way.

Statement Four

As with all statements in this cluster, responses were evenly distributed. Although this statement has two parts, it was assumed that no one would advocate individual student scores being made public, so it was only a question as to whether or not school scores should be published and open to public scrutiny. By only a slight margin respondents felt that they should not.

Statement Five

A strong majority (more than two thirds) of respondents feel that for schools that score consistently low on NSATs, serious investigation is warranted.

Index of Perceptions of Accountability for NSAT results:

The mean of the scores on the five statements dealing with accountability for NSAT results was calculated to determine an index of perceptions of accountability (IPAC) for each respondent. A higher IPAC indicates a perception of greater accountability for NSAT results. The correlation coefficient (Pearson R) for each of the four items against the IPAC was calculated.

Table 4.64**Correlation of Statements in Cluster "G" to IPAC**

Statement	Correlation to IPAC
One	.756
Two	.768
Three	.664
Four	.61
Five	.668

An analysis of variance (ANOVA) was done to determine the strength of the relationship between the index of perceptions of accountability (IPAC) and the following category variables: (1) gender, (2) position, (3) level taught, (4) experience, (5) subject taught and (6) teaching license.

Relationship of Category Variables to Perceptions of Accountability for NSAT Results:

Table 4.65
Gender and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
Male	297	3.071	.708	.041
Female	281	2.956	.671	.04

Table 4.66
Category Differences in Gender and IPAC

Comparison	Mean Diff.	Fish. PLSD	F Test
Male vs. Female	.116	.113 *	4.038 *

* Significant at 95%

Male respondents had a significantly more positive attitude toward accountability for the NSAT results than did females. (p=.045)

Table 4 67
Position and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
Teacher	528	3.025	.683	.03
Teach/Admin	12	2.95	.65	.188
NT/ Admin.	32	3.006	.759	.134
Other	33	2.994	.881	.153

No significant differences in perceptions of accountability for NSAT results were found to be attributable to position. ($p=.9767$, $F=.068$)

Table 4.68
Level Taught and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
Jr. High	318	3.05	.676	.038
Sr. High	242	2.98	.698	.045
JH/SH	40	3.07	.788	.125

No significant differences in perceptions of accountability for NSAT results were found among the three levels taught. ($p=.1752$, $F=1.59$)

Table 4.69
Experience and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
< 2 years	28	3.107	.737	.139
2-5 years	71	2.986	.633	.075
6-10 years	72	3.039	.684	.081
11-20 years	241	2.991	.685	.044
> 20 years	194	3.053	.733	.053

No significant differences in attitudes toward accountability for NSAT results were found to be related to experience.

($p=.8263$, $F= .375$) Regression yields $R = .005$

Table 4.70
Subject Taught and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
Math/Sci	168	3.03	.637	.049
Eng/Soc.St.	155	2.959	.711	.057
French	62	3.068	.663	.084
Other	154	3.071	.693	.056

No significant differences IPAC were found related to subject taught. ($p=.4056$, $F= 1.003$)

Table 4.71
Teaching License and IPAC

Group	Count	Mean	Std. Dev.	Std. Error
TC4	18	2.989	.747	.176
TC5	247	3.026	.641	.041
TC6	193	3.032	.69	.05
TC7	103	2.988	.752	.074
TC8	40	2.96	.889	.141

No significant differences in perceptions of accountability for NSAT results were found related to teaching license.

($p=.9631$, $F=.15$) Regression yields $R = .022$

Relationships Among Indices

Table 4.72

Relationships Among the Indices

	IACI	IAIO	IAMO	IAA	IPA	IPU	IPAC
IACI	1						
IAIO	.221	1					
IAMO	.327	.442	1				
IAA	.326	.345	.466	1			
IPA	.103	.096	.093	.082	1		
IPU	.226	.372	.379	.46	.195	1	
IPAC	.184	.29	.390	.403	.188	.62	1

Table 4.72 is a correlation matrix which shows how the seven indices that were calculated above are related to each other. With a sample size of 606, all of the R values in the table are significant at 95 per cent ($R > .08$) and as expected, all of the correlations are positive. The R values are included in the two tables below which display these relationships in order from strongest to weakest.

Table 4.73

Stronger Relationships Among Indices

Perception of accountability for NSATs	Perception of utility of the NSATs	.620
Attitude towards accountability	Attitude towards measurability	.466
Attitude toward accountability	Perception of utility of the NSATs	.460
Attitude toward identifying outcomes	Attitude toward measurability	.442
Perception of accountability for NSATs	Attitude toward accountability	.403
Perception of accountability for NSATs	Attitude toward measurability	.390
Perception of utility of the NSATs	Attitude toward measurability	.379
Perception of utility of the NSATs	Attitude toward identifying outcomes	.372
Attitude toward accountability	Attitude toward identifying outcomes	.345
Attitude toward community involvement	Attitude toward measurability	.327
Attitude toward community involvement	Attitude toward accountability	.326

Table 4.73 shows that the strongest relationships involve the concepts of measurability, utility, accountability and identifiability. It was expected that the extent to which the respondents felt positively or negatively about these concepts would rise or fall together. For example, it is reasonable to assume that those who feel that educational outcomes can and should be identified and measured, would also feel more positive about the utility of testing and accountability for test results.

Table 4.74
Weaker Relationships Among Indices

Perception of accountability for NSATs	Attitude toward identifying outcomes	.290
Attitude toward community involvement	Perceptions of utility of NSATs	.226
Attitude toward community involvement	Attitude toward identifying outcomes	.221
Perceptions of awareness of NSATs	Perceptions of utility of NSATs	.195
Perceptions of awareness of NSATs	Perceptions of accountability for NSATs	.188
Attitude toward community involvement	Perceptions of accountability for NSATs	.184
Perceptions of awareness of NSATs	Attitude toward community involvement	.103
Perceptions of awareness of NSATs	Attitude toward identifying outcomes	.096
Perceptions of awareness of NSATs	Attitude toward measurability	.093
Perception of awareness of NSATs	Attitude toward accountability	.082

Table 4.74 shows that the indices that were most weakly related to the others involved perceptions of awareness of NSATs and the notion of community involvement. This result was not unexpected. While it is reasonable to assume the attitude toward identifying, measuring and being accountable for outcomes are related, it is not at all clear that these concepts are related to perceptions of awareness or community involvement.

Relationship of Indices to Nova Scotia Achievement Test Results:

Table 4.75
Relationship of Indices to NSAT Results

ATTITUDE TOWARD	JH	SH
Community Involvement	.154	-.056
Identifying Outcomes	-.007	.063
Measurability of Outcomes	.046	.035
Accountability	.047	-.009
Perceptions of Awareness	.09	-.047
Perceptions of Utility	.116	.082
Perceptions of Accountability	.058	.182

Table 4.75 shows the correlation coefficient between the various indices (by respondent) and the mean NSAT score for each school. As NSATs are written at grade nine and grade twelve, separate columns are shown for junior and senior high.

Clearly, there is a very weak relationship between the attitudes and perceptions measured and actual performance on Nova Scotia Achievement Tests. This provides evidence against the "sour grapes theory" - that schools that do well on standardized testing have more positive attitudes towards testing and schools that do poorly on standardized testing feel more negatively about it. There is no significant differences among schools in any of the attitudes measured even though there are very strong differences in test scores.

Discussion of Findings

Community Involvement: There was very strong support for the notion that communities (the public) should be involved in public school education, that their interest was genuine and that their involvement makes a difference. Non teaching principals and vice-principals had a significantly more favourable attitude toward community involvement then did teachers (effect size .85) and Junior High School teachers had a more favourable attitude then did Senior High teachers

(effect size .28). Neither gender, license, experience or subject taught appeared to significantly influence attitudes toward community involvement.

Identifying Outcomes: A majority of respondents felt that identifying educational outcomes was important, that teachers should focus on results as well as process of education and that all teaching should include a demonstration of learning. Attitudes toward this concept appeared to be independent of gender, position, level, experience or license.

Measuring Outcomes: A strong majority of respondents felt that cognitive outcomes of education can and should be measured but were unsure if most of the "important" outcomes of education could be measured. They were undecided about whether or not school effectiveness could be determined through student assessment but generally agreed that it could provide significant information about school effectiveness. Respondents were less sure if teacher effectiveness could be measured through student assessment. The only category variables that appeared to affect attitudes toward measurability were position held and years of experience. Principals and Vice Principals and those respondents in the "other" category (mostly guidance counsellors) had a significantly more favourable attitude toward measurability

then did other groups (effect size .52) In addition, attitudes toward measurability generally became more positive as years of experience increased and those with over twenty years experience had significantly more favourable attitudes (effect size .3).

Accountability: A very strong majority of respondents felt that teachers were responsible for the quality of instruction in their classrooms but were unsure if the quality of instruction could be measured by achievement testing. They generally did not feel that demands for more specific data on student achievement would damage the educational process although they were sceptical that such demands were political and bureaucratic. Gender, position, experience, and license level were significantly related to attitudes towards accountability. Males had a more favourable attitude than did females (effect size .18). Non teaching administrators had a more favourable attitude than did teachers in the study (effect size .61). Respondents with more than 20 years experience and respondents with teaching licenses greater than level five also indicated a significantly more favourable attitude toward the accountability statements (effect size .3).

Awareness of NSATs: Less than half of the respondents perceived that teachers were aware of the purpose and results of the NSATs (although more agreed than disagreed, large numbers were undecided). A strong majority of respondents felt that parents would like to compare their child's achievement to others in the province but that they (parents) were generally not aware of the NSAT results. There was no consensus about the perception of the level of understanding (regarding NSATs) among the teachers. Respondents at the junior high level perceived greater awareness of the NSATs than respondents at the senior level (effect size .18) and teachers with more than ten years experience perceived greater awareness than those with less than ten years (effect size .3).

Utility of NSATs: While respondents were somewhat undecided about whether the NSATs provided a valid measure of some of the desired outcomes of education, most agreed that the interruption needed to administer them was warranted and that they do provide useful information about student achievement. Most thought that continued efforts should be made to improve and understand the NSATs. The only significant differences in any of the groups regarding perceptions of utility of the NSATs occurred in level taught, where junior high teachers perceived greater utility than senior high teachers (effect size .17).

Accountability for Results of NSATs: Although a (slight) majority felt that schools should not be held accountable for their performance on NSATs, more agreed (than disagreed) that it may be useful to compare schools using NSAT results (but that these results should not be made public). Only slightly more agreed than disagreed that the quality of classroom teaching has a significant effect on NSAT results. A strong majority agreed that efforts should be made to determine why some schools score consistently low. Among the respondents, males have a significantly more positive perception of accountability for NSAT results than do females (effect size .167). Other category variables did not appear to influence this perception.

The widely held view that English/Social Studies teachers are unfavourably disposed toward testing and Math/Science teachers are more favourably disposed is not supported by this study. No significant difference in attitudes towards testing by subject taught were detected.

Chapter V

Conclusions and Recommendations

The purpose of this study was to explore perceptions of education l accountability, the role of standardized testing, and the association that exists among components of an accountability relation.

In this study, the concept of accountability was defined as the responsibility of schools and school systems to demonstrate to the larger community (the public) that their teachers are teaching effectively and their students are learning appropriately. Accountability defined in these simple terms requires that those who embrace it assume that educational outcomes are identifiable and measurable and that the school community (the public) is entitled to hold schools and school systems accountable for the measurement obtained. How teachers and administrators feel about these assumptions is the essence of this study.

An instrument was developed that measured perceptions and attitudes toward identifying and measuring educational outcomes and the notion of accountability. In addition to examining these attitudes in a generic sense, perceptions of

awareness and utility of the Nova Scotia Achievement Test were measured and related to the notion of accountability.

Summary of Findings

Results of the study can be categorized into four areas; (1) a general description of the attitudes and perceptions measured, (2) a determination of differences in attitudes and perceptions among various demographic variables, (3) a description of the relationships among the attitudes measured, and (4) a description of the relationship between the attitudes measured and NSAT results.

General Description of Attitudes and Perceptions

A major portion of the study involved the generation of descriptive statistics that characterize attitudes and perceptions about public involvement in education (entitlement), identifying and measuring outcomes, and responsibility for these outcomes. Accountability embraces the notion that school communities (public) are entitled to some account of how schools are performing. It is important to know how educators feel about community participation in the process of education.

The data gathered indicated that generally, respondents feel very positively about the idea of community involvement in public school education. This would indicate that teachers and administrators in this sample have not retreated into a "closed shop" mentality where community members (lay people) are excluded from the "profession". While they stopped short of fully endorsing an "auditing" function for the public, there is reason to believe that specific information regarding the nature and purpose of a public "audit" may lead to acceptance of this idea. The most important finding in this section is that, for the most part, it appears that public school educators would welcome community/school linkages and perceive them as beneficial.

In this study it was assumed that accountability (as defined) requires that the outcomes of education are able to be identified and measured. Before determining if educators perceive that outcomes are measurable, it is necessary to determine if they are seen as identifiable. Results indicate that a wide majority of respondents feel that the specific outcomes of education can, and should, be identified. Generally, results from this section support the proposal that the results of education are very important indeed and that every attempt should be made to have them clearly designated. If there was no support for the notion that outcomes are

identifiable then it would be inappropriate to speak of measurability and accountability.

Just as most respondents believe that the outcomes of education can be identified, a strong majority feel that the (cognitive) outcomes of education can (and should) be measured. While some feel that many of the important outcomes (possibly in the affective domain) are not measurable, there is general acceptance that the cognitive results of teaching can be assessed. Responses generally support the idea that the effectiveness of schools can be (at least partially) determined by assessing students. While there is no clear indication that assessment schemes currently in place effectively measure the effectiveness of schools, respondents (for the most part) believe that such an assessment scheme is feasible. This is an important result for educators who subscribe to any model of accountability.

Even though they categorize recent interest in educational accountability as political and bureaucratic, teachers and administrators do not (generally) feel that current trends calling for more specific data on student achievement will damage the educational process. Teachers directly acknowledge their personal responsibility (accountability) for the quality of instruction in their classroom but only a slight majority

feel that the quality of instruction in the classroom affects standardized test results. The general feeling is that, while teachers believe in identifying and measuring outcomes, the number of variables (beyond their control) which affect student performance causes them to be reluctant about responsibility for the measurement obtained.

Perceptions of Nova Scotia Achievement Tests

Following an exploration of attitudes toward the generic concepts mentioned above, descriptive statistics were gathered to determine perceptions of awareness, utility and accountability for one particular measure of student performance, the Nova Scotia Achievement Tests.

While the statistics gathered indicate no strong consensus and many undecided, teachers generally feel that their colleagues are aware of the purpose and results of the NSATs. They perceive that parents want to know about their child's performance relative to others but feel that they (parents) are generally unaware of NSAT scores for their own children. Following a determination of levels of awareness, an attempt was made to determine if teachers and administrators felt that the NSATs were useful. Results indicate that despite the

reservations of a significant number who were undecided on the issue, generally it was felt that these tests do provide limited, but useful information about student achievement. Significantly more teachers feel that the interruption caused by this testing is warranted than feel that it is not. This was a somewhat surprising result as many educators perceive that teachers would like to see these tests eliminated altogether. Again, probably due to the myriad of factors involved in student achievement, teachers did not generally feel that schools or school systems should be held accountable for the results of NSATs or that the results should be made public. These results are only partially consistent with Grant (1993) who determined that administrators in Nova Scotia had widespread concerns over NSATs and Hall (1993) who concluded that NSATs are perceived not to be an adequate measure of student achievement or school effectiveness.

Relationship of Demographic Variables to Attitudes

Demographic variables included in the study included gender, position, level taught, experience, subject taught, and teaching license. The researcher was interested in whether or not these variables influenced the attitudes and perceptions under investigation. The following relationships were

determined; (1) males held a significantly more favourable attitude toward accountability than do females, (2) principals and other non-teaching administrators hold a more favourable attitude toward community involvement, measurability and accountability, and levels of awareness of NSATs, than do teachers and (3) teachers at the Junior High level hold a more favourable attitude toward community involvement, and utility and awareness of NSATs, than do teachers at the Senior High level. (4) More experience is related to more positive attitudes toward measurability, accountability and awareness of NSATs and (5) teachers with licenses above level five hold a more favourable attitude toward accountability than those below level five.

Otherwise, attitudes toward the various concepts was not significantly influenced by the demographic variables listed. The limited influence of the demographic variables is one of the most surprising results of the study. The widely held view is that English/Social Studies teachers are unfavourably disposed toward testing and Math/Science teachers are more favourably disposed is not supported by the study which measures no significant difference in attitudes towards testing by subject taught. In addition, another popular theory is not consistent with the data gathered. The "sour grapes theory"... that schools that do well on standardized testing

have positive attitudes towards testing generally and schools that do poorly on standardized testing feel negatively about it is not supported. There is no significant difference in attitudes measured among schools even though there are very strong differences in test scores.

Relationships Among the Attitudes Measured

The strongest relationships among the attitudes measured involve the concepts of measurability, utility, accountability and identifiability. It was expected that the extent to which the respondents felt positively or negatively about these concepts would rise or fall together. For example, it is reasonable to assume that those who feel that educational outcomes can and should be identified and measured, would also feel more positive about the utility of testing and accountability for test results. The indices that were most weakly related to the others involved perceptions of awareness of NSATs and the notion of community involvement. This result was not unexpected. While it is reasonable to assume the attitude toward identifying, measuring and being accountable for outcomes are related, it is not at all clear that these concepts are related to perceptions of awareness or community involvement. It is surprising to note that none of the demographic variables are strongly related to NSAT results.

Recommendations for Future Study

More than six hundred teachers and administrators were included in this study. While this large number did improve the quality of the measurement, it may not be necessary to include such a large sample in future studies of this nature. Correlational studies do not require large samples to provide reasonable results. Because sampling error varies inversely as the square root of sample size a sample of 606 is only 1.2 times as accurate as a sample of 400.

The most serious limitation of the study is that the instrument used to measure the attitudes under investigation was developed by the researcher and therefore the validity of the instrument could not be verified by other studies. This limitation does restrict the interpretation of the results. It would be interesting for future researchers to apply the same instrument to other groups such as elementary school teachers, parent groups, supervisory personnel, politicians and members of the business community. If the same instrument were used again, the cluster measuring levels of awareness of the NSATs should be left out because it had little or no relation to other clusters and because it is not an essential component of the accountability relation. In addition the instrument could

be further improved by refining the statements based on feedback from the large sample included in this study.

Recommendations for Practitioners

Practitioner support for the possibility of identifying and measuring specific educational outcomes is essential for the implementation of an accountability relation. This study not only indicates a general support for these prospects but also indicates that this support is not limited to particular groups within the study. It was not intended that this project endorse any particular form of assessment but it was hoped that it would provide testimony that an assessment scheme (for accountability purposes) was possible.

This project was intended to be a preliminary study which could provide some useful information for educators who are interested in exploring the notion of educational accountability through some form of student assessment. It may provide some encouragement in this regard. It provides evidence that educators are concerned about the results of education and have a sense of responsibility for learning outcomes.

REFERENCES

- Alberta Education (1993) Achieving Quality - Final Report of the Educational Quality Indicators Initiative. Alberta, Canada: Alberta Education.
- Ary, D., Jacobs, L.C. & Razavieh (1985) Introduction to Education Research (3rd ed.). New York: Holt, Rinehart and Winston
- Bacharach, S.B. (1991) Education Reform: Making Sense of It All. Boston: Allyn and Bacon.
- Balderson, J. (1991) Measuring Up The Canadian School Executive vol.11, no.8, pp. 3
- Berlak, H. (1985) Testing in a Democracy Educational Leadership, vol.43, no.2, pp. 16-20
- Berliner, D.C. (1993) Mythology and the American System of Education Phi Delta Kappan, vol.74, no.8, pp. 632
- Bovaird, J. (1991, April/May) To Test or Not To Test FWTAO Newsletter
- Boyer, E. (1983) High School: A Report on Secondary Education in America, New York: Harper Row
- Brandt, R. (1992) On Research on Teaching: A Conversation with Lee Schulman Educational Leadership, vol 49, no.7 pp.14-19
- Bridge, R.G., Judd, C.M., Moock, P.R. (1979) The Determinants of Educational Outcomes Cambridge: Ballinger Publishing Company

- Bryk, A.S. & Hermanson, K.L. (1993) Educational Indicator Systems: Observations on Their Structure, Interpretation and Use. In L. Darling-Hammond (Ed.) Review of Research in Education (pp.451-479). Washington, DC: American Educational Research Association.
- Campbell, W.G., Ballou, S.V., & Slade, C. (1982) Form and Style: Thesis, Reports, Term Papers Sixth Edition. Boston: Houghton Mifflin
- Cannell, J.J. (1988) Nationally Normed Elementary Achievement Testing in the United States: How All Fifty States Scored Above the National Average Educational Measurement: Issues and Practices 7-2, pp. 5-9
- Clarke, V.B. (1993) Waves of Education Reform The Canadian School Executive vol.13, no.6, pp. 7-11
- Coleman, J.S. et al. (1966) Equality of Educational Opportunity Washinton D.C. ERIC Document ED 012275
- Common, D.L. (1987) Evaluation and Quality in Public Education Education Canada Fall 1987
- Corporate Higher Education Forum (1992) To Be Our Best: Learning for the Future Canada: Montreal
- Council of Ministers of Education, Canada (1991, December) School Achievement Indicators Program. Background Paper No.1
- Darling-Hammond, L. (1991) The Implications of Testing Policy for Quality and Equality. Phi Delta Kappan. vol. 73, no.3, pp. 220-225
- Darling-Hammond, L. (1990) Achieving Our Goals: Superficial or Structural Reforms Phi Delta Kappan vol. 74, no.4, pp. 286-295

- Davis, B. (1993) Discovering Educational Research, Halifax: Saint Mary's University
- Denoyer, R.A. & White, M. (1990) Tests: Fallible Indicators of Educational Quality NASSP Bulletin, vol.74, no.523, pp. 49-52
- Eastman, N. (1991) Educational Evaluation: The Future Theory Into Practice vol. 30, no.1, pp. 74-79
- Eckstein, M.A. & Noah, H.J. (1992) Examinations: Comparative and International Studies, New York: Pergamon Press
- Economic Council of Canada (1992) A lot to Learn, Education and Training in Canada, Canada Communication Group - Publishing, Ottawa Canada
- Edmond, D.R. (1992) The Development and Use of Educational Indicator Systems. Education Canada, vol. 32, no.4, pp. 8-13
- Edmond, D.R. (1993) Development and Using Educational Indicator Systems The Canadian School Executive vol.12, no.6 pp.3-9
- Eisner, E.W. (1991) What Really Counts in Schools Educational Leadership vol.48, no.5, pp.10-17
- Elam, S.M., Rose, L.C., & Gallup, A.M. (1993) The 25th Annual Gallup/Phi Delta Kappa Poll of Public Attitude Toward Public School Phi Delta Kappan, vol. 75, no. 2, pp.137-152

- Foster, J.D. (1991) The Role of Accountability in Kentucky's Education Reform Act of 1990 Educational Leadership vol.48, no.5, pp.34-36
- Freedman, J. (1993) Failing Grades: Canadian Schooling in a Global Economy. Society for Advancing Educational Research, Canada: Full Court Press Inc.
- Fullen, M.G. (1991) The New Meaning of Educational Change (2nd. Ed. New York: Teachers College Press
- Glickman, C.D. (1990) Open Accountability for the '90s: Between the Pillars. Educational Leadership vol.47, no.7, pp. 38-41
- Glasnapp, D.R. & Poggio, J.P. (1991) Mandated Competency Testing Programs: Evaluating State and Local Initiatives Theory Into Practice vol.30, no.1, pp.61-67
- Goodlad, J. (1984) A Place Called School: Prospects for the Future, New York: McGraw Hill
- Grant, J. (1993) A Summary of Submissions of NSSAA on Standardized Testing. Halifax, Nova Scotia: Nova Scotia Teachers Union
- Green, K.E. (1992) Differing Opinions on Testing Between Pre-service and Inservice Teachers Journal of Education Research, vol.86, no.1, pp.37-42
- Green, K.E. & Stager, S.F. (1984) Wyoming Teachers' Use of Tests and Attitudes Towards Classroom and Standardized Tests. ERIC document ED 252575
- Green, K.E. & Stager, S.F. (1985) Teachers' Attitudes Towards Testing. Paper presented at the Annual Meeting of the National Council on Measurement in Education. ERIC document ED 253588

- Haertel, E. (1989) Student Achievement Tests as Tools of Educational Policy: Practices and Consequences. Test Policy and Test Performance: Education Language and Culture. Boston: Kluwer Academic Publishers
- Hall, P. (1993) Destreaming in Nova Scotia: Where are we Now? The Canadian School Executive vol.13, no.6, pp. 22-25
- Hanushek, E.A. (1981) Throwing Money at Schools Journal of Policy Analysis and Management vol.1, pp. 19-41
- Harris, K.H. & Longstreet, W.S. (1990) Alternative Testing and the National Agenda for Control The Cleaning House vol.64, no.2, pp. 90-93
- Herman, J.L. (1992) What Research Tells Us About Good Assessment Educational Leadership vol.48, no.8, pp.74-78
- Hoffman, B. (1962) The Tyranny of Testing New York: Crowell Collier Press
- Horber, D.T. & Geisinger, K.F. (1983) The Relationship of Attitudes Toward Multiple Choice Tests and Convergent Production, Divergent Production and Risk Taking. Paper presented at the annual meeting of AERA. ERIC document ED229435.
- Houts, P. (1977) The Myth of Measurability New York: Hartt
- Kaagan, S. (1990) State Education Indicators: A Mandate for Action NASSP Bulletin vol.74, no. 527, pp.54-65

Kaagan, S. & Smith, M.S. (1985) Indicators of Educational Quality. Educational Leadership vol 43, no.2, pp.21-24

Kagan, S. & Coley, R.J. (1989) State Education Indicators: Measured Strides, Missing Steps Centre for Policy Research in Education, Rutgers University and Policy Information Centre

Karmos, A.H. & Karmos, J.S. (1984) Attitudes Towards Standardized Tests and Their Relation to Achievement Test Performance. Measurement and Evaluation in Counseling and Development, vol.17, no.2, pp.56-66

Katzer, K., Cook, K.H., Crouch, W.W. (1991) Evaluating Information: A Guide for Users of Social Science Research. New York: McGraw Hill

LeBlanc, R.D. (1994) Accountability: Hindsight and Foresight The Canadian School Executive vol.13, no.7, pp.21-24

Lessinger, L.M. (1973) Accountability: Present Forces and Future Concerns Creating Appraisal and Accountability Systems, New Directions for Education. Washington: Jossey-Bass

Lieberman, A. (1991) Accountability As a Reform Strategy Phi Delta Kappan, vol.73, no.3, pp. 219-220

Linn, R.L. (1989) Educational Measurement (3rd Edition), New York: MacMillan

Linn, R.L. (1993) Educational Assessment: Expanded Expectations and Challenges Educational Evaluation and Policy Analysis, vol 15, no.1, pp. 1-16

- Madaus, G.F. (1985) Public Policy and the Testing Profession: You've Never Had It So Good Educational Measurement: Issues and Practices, vol.4, no.4, pp. 5-11
- Madaus, G.F. (1991) The Effects of Important Tests On Students: Implications for a National Examination or System of Examinations Phi Delta Kappan vol.73, no.3, pp. 226-231
- Madaus, G.F. & Tan, A.G.A (1993) The Growth of Assessment: Challenges and Achievements of American Education. In G.Cawelti (Ed.) The 1993 ASCD Yearbook, (pp. 53-79)
- Maguire, T.O. (1985) Is Canada a Nation at Risk? A report to the Canadian Educational Researchers Association ERIC Document ED269846
- Maharaj, H. (1991) Performance Testing: Move in the Right Direction Manitoba Teacher vol.70, no.1, pp.16-17
- Marso, R.N., & Pigge, F.L. (1992) Classroom Teachers Perception of the Extent and Effectiveness of their Schools Uses of Standardized Test Results ERIC Document ED342802
- Marso, R.N. & Pigge, F.L. (1991) Testing Directors' Principals' and Teacher Perceptions of the Actual Purposes of standardized Testing ERIC Document ED333041
- Martuza, V.R. (1977) Applying Norm Referenced and Criterion Referenced Measurement in Education Toronto: Allyn and Bacon

- McConaghy, T. (1991) National Testing, Canadian Style Phi Delta Kappan, vol.73, no.2, pp.172-173
- McCormick, R. (1982) Calling Education to Account. Halley Court Oxford England: Heinemann Educational Books
- McEwan, N. & Chow, H. (1990) Issues in Implementing Indicator Systems ERIC Document ED323262
- McEwan, N. & Zatko, G. (1989, June) Educational Quality Indicators: Framework for Action. Paper presented at the annual meeting of the Canadian Educational Researchers Association.
- Mehrens, W.A., & Kominski, J. (1989) Methods for Improving standardized Test Scores: Fruitful Fruitless or Fraudulant? Educational Measurement: Issues and Practices vol.8, no.1, pp. 14-22
- Mehrens, W.A., & Lehmann, I.J. (1980) Standardized Tests in Education Third Edition, New York: Holt Rinehart and Winston
- Mehrens, W.A., & Lehmann, I.J. (1984) Measurement and Evaluation in Education and Psychology Third Edition, New York: Holt Rinehart and Winston
- Messick, S. (1989) Validity Educational Measurement New York: MacMillan
- Moss, P.A. (1992) Shifting Concepts of Validity in Educational Measurement: Implications for Performance Assessment. Review of Educational Research, vol.62, no.3, pp. 229-258
- National Commission on Excellence in Education (1983) A Nation at Risk Washington D.C. U.S. Government Office

National Commission on Testing and Public Policy (NCTPP) (1990) From Gatekeeper to Gateway: Transforming Testing in America Chestnut Hill Mass. : NCTPP Boston College

Nova Scotia Dept. of Education (1993) Nova Scotia Achievement Tests, Report of 1992 Results. Halifax Nova Scotia.

Nova Scotia Dept. of Education (1992) Guide to Nova Scotia Achievement Tests, Level 12

Odden, A. (1990) Educational Indicators in the United States: The Need for Analysis Educational Researcher vol.19, no.5 pp.26

Packard, R., & Dereshiwsky, M.I. (1991) Results of Perception Assessment Scale. Arizona Career Ladder Pilot Test Project ERIC Document ED334148

Passow, A.H. (1990) How It Happened, Wave by Wave Education Reform, Making Sense of it All Allyn and Bacon: Boston

Pipho, C. (1986) States Move Closer to Reality Phi Delta Kappan vol.68, no.4

Popham, W.J. (1978) Criterion Referenced Measurement New Jersey: Prentice Hall

Popham, W.J. (1988) Educational Evaluation (2nd edition). Englewood Cliffs New Jersey: Prentice Hall

Popham, W.J. (1991) Appropriateness of Teachers Test-Preparation Practices Educational Measurement: Issues and Practices vol.10, no.4, pp. 12-15

Popham, W.J. (1993) Circumventing the High Costs of Authentic Assessment. Phi Delta Kappan, vol.74, no.6, pp.470-473

Principles for Fair Student Assessment Practices for Education in Canada (1993). Edmonton Alberta: Joint Advisory Committee

Ravitch, D. (1985) The Schools We Deserve New York: Basic Books

Report of Auditor General of Nova Scotia (1990), Halifax: Government of Nova Scotia

Report of Commission on Public Education Finance (1981), Halifax: Government of Nova Scotia

Report of Select Committee on Education (1992), Halifax: Government of Nova Scotia

Roberson, E.W. (1970) Educational Accountability Through Evaluation Educational Technology Publications: Englewood Cliffs, New Jersey

Rogers, W.T. (1991) Establishing Measurement Standards for Canada's Schools The Canadian School Executive vol.11, no.3

Rosenthal, R., & Jacobson, L. (1968) Pygmalion in the Classroom New York: Holt Rinehart and Winston

Rudman, H.C. (1987) The Future of Testing Is Now Educational Measurement: Issues and Practices vol.6, pp.5-11

Sizer, T.R. (1985) Common Sense Educational Leadership vol.42, no.6, pp.21-22

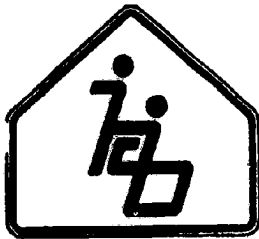
Sizer, T.R. (1984) Horace's Compromise: The Dilemma of the American High School, Boston: Houghton Mifflin

- Smith, M. (1988) Educational Indicators Phi Delta Kappan vol.69, no.7, pp.487
- Sockett, H. (1980) Accountability: Purpose and Meaning Calling Education to Account. Oxford England: Heinemann Educational Books
- Soltz, D.F. (1992) Does It Really Matter What Teachers Think of Tests? Paper presented at the annual meeting of American Educational Research Association. ERIC document ED344901
- Spady, W.G. & Marshall, K.J. (1991) Beyond Traditional Outcome Based Education. Educational Leadership, vol 49, no.2, pp.67-72
- Stiggins, R.J. (1991) Assessment Literacy Phi Delta Kappan vol.72, no.7, pp.534-539
- Strenio, A.J. (1981) The Testing Trap New York: Rawson Wade Publishers Inc.
- Sullivan, T.S. (1987) Variables Related to High School Achievement Ph.D. dissertation, University of Alberta
- Summers, A.A., & Wolfe, B.L. (1975) Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools Business Review Federal Reserve Bank of Philadelphia
- Timpane, M. (1977) Some Political Aspects of Accountability Mandates The Changing Politics of Education Edited by Moser, E.K. & Wagoner, J.L. McCutchan: Berkeley
- UNESCO (1990) World Report on Education

- United States Dept of Education (1988) Creating Responsible and Responsive Accountability Systems Washington, D.C. Office of Educational Research and Improvement
- United States Dept of Education (1991) America 2000: An Education Strategy: Sourcebook, Washington DC: US Dept of Education
- Wagner, R.B. (1989) Accountability in Education: A Philosophical Inquiry Routledge: New York
- Walberg, H.J. (1990) OECD: Indicators of Educational Productivity Educational Researcher vol.19, no.5
- Wiggins, G. (1991) Standards Not Standardization: Evoking Quality Student Work Educational Leadership vol.48, no.5, pp.18-25
- Wise, A.E. (1979) Legislated Learning: The Bureaucratization of the American Classroom Berkley: University of California Press.
- Worthen, B.R. (1993) Critical Issues That Will Determine the Future of Alternate Assessment Phi Delta Kappan vol.74, no.6, pp.444-456
- Worthen, B.R. & Sanders, J.R. (1991) The Changing Face of Educational Evaluation Theory Into Practice v. 30-1
- Worthen, B.R. & Spandel V. Putting the Standardized Test Debate in Perspective Educational Leadership vol.48 no. 5 pp.65-69
- Wright, R.L.D. (1976) Understanding Statistics New York: Harcourt Brace Jovanovich, Inc.

APPENDICES

APPENDIX A
PERMISSIONS TO PROCEED



Halifax County - Bedford District School Board

267 Cobequid Road, • P. O. Box 1000, • Lt. Sackville, N.S. • B4C 3Z5

MEMORANDUM

To: Junior and Senior High School Principals
From: Lloyd Gillis
Date: November 5, 1993
Re: Research Project - Bill Kilfoil

Our colleague, Bill Kilfoil, is on a one year study leave at Saint Mary's University. His thesis will explore educational accountability and the role of standardized achievement testing. We are interested in this research as it will be of benefit to our system.

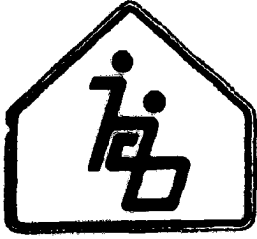
Bill will deliver questionnaires for all teachers and administrators by mid-November. The questionnaire will require approximately ten minutes for completion. He would like to have the questionnaire distributed, completed and collected during one sitting, perhaps at a staff meeting.

Thank you for your cooperation in assisting with this project.

L. E. Gillis
C. E. O.

LBG:mk

"committed to lifelong learning"



Halifax County - Bedford District School Board

267 Cobequid Road, • P. O. Box 1000, • Lr. Sackville, N.S. • B4C 3Z5

October 5, 1993

Mr. Bill Kilfoil
72 Sunnyholme Drive
Oyster Pond, N. S.
BOJ 1WO

Dear Bill:

I believe that your study will be of benefit to the system and I offer my support. As you are aware, we request that you obtain permission from the principal before conducting the survey in the school.

I would be pleased to meet with you and offer my assistance. Please call Marg and arrange a convenient time.

Yours truly,

Dr. J. F. Morrison
Superintendent of Research, Evaluation & Planning

JFM:mk

"committed to lifelong learning"

DEAR PRINCIPALS:

1. First of all, thank you very much for agreeing to take a few minutes with your staff to complete this questionnaire. It is much appreciated.
2. If possible, would you please distribute AND collect these questionnaires at a staff meeting. It will require less than **TEN MINUTES** to complete the questionnaire.
3. Would you please call me (889-3659) as soon as the questionnaires are completed. I will make arrangements to come to your school and pick up the completed questionnaires as soon as I hear from you.
4. I would like to have the completed questionnaires as soon as possible but anytime **BEFORE THE CHRISTMAS BREAK** would be fine.
5. The questionnaire asks people to indicate the extent to which they agree or disagree with a series of statements. I believe the instructions on the questionnaire are clear, but please remind your staff that I **am seeking their reaction to these statements**, they do not have to "know" the answer.
6. I would like everyone on staff (including principals and Vice principals) to complete the questionnaire if possible.
7. Please call me if you have any questions. Thanks again!

Bill Kilfoil
72 Sunnyholme Drive
Oyster Pond N.S.
B0J 1W0

889-3659

APPENDIX B
QUESTIONNAIRE

DATE: _____ SCHOOL NUMBER _____

INFORMATION ABOUT THE RESPONDENT

	PLEASE PLACE A CHECK MARK IN THE APPROPRIATE COLUMN	A	B	C	D	E
1	Gender (A) Male (B) Female			X	X	X
2	Position (A) Teacher (B) Teaching Principal or V.P. (C) Non teaching Principal or V.P. (D) Supervisor or Superintendent (E) Other					
3	Working mostly at which level (A) Elementary P-6 (B) Junior High 7-9 (C) Senior High 10-12 (D) Elementary and Junior High (E) Junior and Senior High					
4	Number of years in education (A) less then 2 years (B) 2 to 5 years (C) 6 to 10 years (D) 11 to 20 years (E) more then 20 years					
5	Subject(s) taught most often (A) Math and/or Science (B) English and/or Social Studies (C) French (D) Any other subject(s) (E) Not teaching at all					
6	Teaching License (A) TC4 or less (B) TC5 (C) TC6 (D) TC7 (E) TC8					

Thank you for your cooperation in completing this survey. The entire process will take about 10 minutes. In the following pages you should select the response that represents your reaction to each statement, you do not need to know the answer. Please react to every statement, if you are not clear about the statement, choose undecided. ALL responses require placing a check mark in the appropriate column.

Section ONE

	Indicate your reaction to the following statements by marking the appropriate box. Please react to every statement. Choose "undecided" if you are not sure or have no opinion.	S T R O N G L Y	A G R E E	U N D E C I D E D	D I S A G R E E	S D I S O N G R E E
1	Community involvement in public school education is entirely appropriate.					
2	Teachers have always been held accountable for what they do.					
3	The public has a genuine interest in, and has a right to know, what schools are doing and how successful they are.					
4	Identifying the outcomes of education should be at least as important as the inputs and processes of education.					
5	Public interest in education is generally limited to times when schools are "under attack".					
6	Community involvement has no significant impact on the quality of schools.					
7	Teachers should focus on the process of education not the results.					
8	It is possible to devise a student assessment scheme that will provide significant information about the effectiveness of a TEACHER.					
9	It is appropriate that the public should audit the performance of public schools.					
10	For educators, the desired outcomes of education ARE sufficiently clear.					

		S A	A	U	D	S D
11	The cognitive outcomes of education can and should be measured.					
12	It is NOT possible to quantitatively measure the effectiveness of schools.					
13	Recent interest in educational accountability is largely political and bureaucratic.					
14	Current trends calling for more specific data on student achievement will ultimately damage the educational process.					
15	Achievement test results can be used to compare the quality of instruction in schools with similar characteristics.					
16	All teaching should include a culminating demonstration of learning.					
17	It is possible to devise a student assessment scheme that will provide significant information about the effectiveness of a SCHOOL.					
18	Every attempt should be made to identify the specific outcomes of instruction in each of the core areas.					
19	Most of the important outcomes in education are NOT measurable.					
20	Teachers themselves are the most significant contributors to the quality of instruction in their classrooms.					

Section TWO

This next set of 15 statements refers specifically to the NOVA SCOTIA ACHIEVEMENT TESTS referred to as the NSAT's. The HALIFAX COUNTY/BEDFORD DISTRICT SCHOOL BOARD is referred to as HCBDSB.

	Indicate your reaction to the following statements by marking the appropriate box. Please react to every statement, choose "undecided" if you are not sure or have no opinion.	S T R O N G L Y	A G R E E	A G R E E	U N D E C I D E D	D I S A G R E E	S D I R S O A N G R E E
1	The interruption (at grade 9 and grade 12) caused by administering NSAT's is NOT warranted.						
2	While multiple choice type tests are limited they do provide useful information about student achievement.						
3	School authorities should not spend their time investigating discrepancies in NSAT scores among schools.						
4	A serious attempt should be made to find out why some schools score consistently low on NSAT's						
5	Most parents of grade 9 and grade 12 students are NOT aware of their child's NSAT scores.						
6	We should continuously attempt to improve the NSAT's and gain greater understanding of potential uses.						
7	Teachers in the HCBDSB are generally aware of the purpose of the NSAT's.						
8	The quality of classroom teaching has a significant effect on NSAT scores.						
9	The NSAT's are a valid measure of many (but not all) of the desired outcomes of education.						

		S A	A	U	D	S D
10	Parents want to know about their child's achievement (in school) relative to other students in the province.					
11	It may be useful to compare SCHOOLS in Nova Scotia using NSAT results.					
12	Schools and school systems should be held accountable for their performance on NSAT's.					
13	Teachers do not understand enough about the NSAT's to make any judgements about their usefulness.					
14	It is appropriate that NSAT results be made public for schools and school systems but not for individual students.					
15	Teachers in HCBDSB are generally aware of the results of the NSAT's.					

APPENDIX C
PILOT QUESTIONNAIRE

TODAYS DATE IS: _____ YOUR SCHOOL NUMBER IS _____

Thank you for your cooperation in completing this survey. The entire process will take about 10 minutes. In sections one and two you should select the response that represents your reaction to each statement, you do not have to know the answer. Please consider your responses carefully. ALL responses require placing an "X" in the appropriate column.

INFORMATION ABOUT THE RESPONDENT MALE () or FEMALE ()

	PLEASE PLACE AN 'X' IN THE APPROPRIATE COLUMN	A	B	C	D	E
1	Position (A) Teacher (B) Teaching Principal or V.P. (C) Non teaching Principal or V.P. (D) Supervisor or Superintendent (E) Other					
2	Level taught (if teaching) (A) Elementary P-6 (B) Junior High 7-9 (C) Senior High 10-12 (D) Elementary and Junior High (E) Junior and Senior High					
3	Total Number of years teaching: (A) less then 2 years (B) 2 to 5 years (C) 6 to 10 years (D) 11 to 20 years (E) more then 20 years					
4	Subject(s) taught most often (A) Math and/or Science (B) English and/or Social Studies (C) French (D) Health/Phy.Ed. (E) Others including specialists					
5	Subsystem (A) Western (B) Bedford (C) Eastern Suburban (D) Sackville (E) Musquodoboit/Sheet Harbour					
6	Teaching License (A) TC4 or less (B) TC5 (C) TC6 (D) TC7 (E) TC8					

	Indicate your reaction to the following statements by placing an "X" in the appropriate box. Please respond to every statement. Choose "undecided" if you are not sure or have no opinion.	S T R O N G L Y	A G R E E	U N D E C I D E D	D I S A G R E E	S D I S O N G R E Y
1	Community involvement in public school education is entirely appropriate.					
2	It is reasonable for educators to identify the desired outcomes of education.					
3	Teachers have always been held accountable for what they do.					
4	It is possible to devise an assessment scheme (administered to students) that will provide significant information about the effectiveness of a SCHOOL SYSTEM.					
5	The public has a genuine interest in, and has a right to know, what schools are doing and how successful they are.					
6	Every attempt should be made to identify the specific outcomes of instruction in each of the core areas.					
7	Most of the important outcomes in education are NOT measurable.					
8	Teachers themselves are the most significant contributors to the quality of instruction in their classrooms.					

		S A	A	U	D	S D
9	Community involvement with schools is often disruptive and unsettling.					
10	Identifying the outcomes of education should be at least as important as the inputs and processes of education.					
11	Generally teachers are apprehensive about the concept of accountability.					
12	Public interest in education is generally limited to times when schools are "under attack".					
13	Community involvement has no significant impact on the quality of schools.					
14	Teachers should focus on the process of education not the results.					
15	It is possible to devise an assessment scheme (administered to students) that will provide significant information about the effectiveness of a SCHOOL.					
16	Generally teachers have no problem with "accountability" and would welcome valid indicators of academic progress.					
17	Schools would improve significantly if there were more parent and community involvement in education.					
18	A properly designed assessment system will motivate students and teachers to work harder to enhance learning.					

		S A	A	U	D	S D
19	It is NOT essential that teachers be concerned with measuring the outcomes of education.					
20	Administrators, (principals supervisors etc.) have always been held accountable for what they do.					
21	It is possible to devise an assessment scheme (administered to students) that will provide significant information about the effectiveness of a TEACHER.					
22	It is appropriate that the public should audit the performance of public schools.					
23	For educators, the desired outcomes of education ARE sufficiently clear.					
24	Appropriate instruments to measure educational outcomes are already in place in our district.					
25	It is NOT appropriate to hold educators accountable by measuring what students learn.					
26	The cognitive outcomes of education can and should be measured.					
27	The public is NOT clear what schools are trying to accomplish.					
28	Taxpayers should feel that school systems are accountable to them for the quality of schools in their district.					
29	Considerable time and attention should be given to developing alternate methods (not currently in use) of measuring educational outcomes.					

		S A	A	U	D	S D
30	The quality of instruction in the classroom directly affects the amount of learning that takes place.					
31	It is NOT possible to quantitatively measure the effectiveness of schools.					
32	Teachers are NOT responsible to their employers for the quality of instruction in their classrooms.					
33	Recent interest in educational accountability is largely political and bureaucratic.					
34	Quality teaching is at least as important as any of the other factors that affect student learning.					
35	It is NOT possible to use the results of standardized achievement tests to compare the quality of schools.					
36	A determination about the quality of SCHOOL SYSTEMS can be made by comparing achievement test results.					
37	Current trends calling for more specific data on student achievement will ultimately damage the educational process.					
38	Achievement test results can be used to compare the quality of instruction in schools with similar characteristics.					
39	All teaching should include a culminating demonstration of learning.					
40	We should not focus on measuring the outcomes of education but rather on the resources and processes of education.					

This next set of 20 statements refers specifically to the NOVA SCOTIA ACHIEVEMENT TESTS written each year by all Nova Scotia students in grade nine and grade twelve. The NOVA SCOTIA ACHIEVEMENT TESTS are referred to as the NSAT's. The HALIFAX COUNTY/BEDFORD DISTRICT SCHOOL BOARD is referred to as HCBDSB.

	Indicate your reaction to the following statements by placing an "X" in the appropriate box. Please respond to every statement, choose "undecided" if you are not sure or have no opinion.	S T R O N G L Y	A G R E E	U N D E C I D E D	D I S A G R E E	S D I S O N G R E E
41	Teachers in HCBDSB are aware that the NSAT's are administered each year to grade 9 and 12 students.					
42	Elementary school teachers (P-6) are NOT concerned with NSAT results.					
43	It is appropriate that NSAT results should be made public for schools and school systems but not for individual students.					
44	Teachers in the HCBDSB are generally aware of the purpose of the NSAT's.					
45	Generally, teachers are favourably disposed toward the administration of NSAT's.					
46	It is appropriate that SCHOOLS be compared using NSAT results.					
47	Teachers in HCBDSB are generally aware of the results of the NSAT's.					
48	Junior High Teachers (7-9) are NOT concerned about NSAT results.					

		S A	A	U	D	S D
49	It is appropriate that school SYSTEMS be compared using NSAT results.					
50	Teachers understand the statistical methods used to report NSAT results.					
51	Senior high teachers (10-12) are NOT concerned about NSAT results.					
52	It is appropriate that individual students be compared using NSAT results.					
53	Teachers do NOT understand enough about the NSAT's to make any judgements about their usefulness.					
54	Teachers sometimes use NSAT results to influence their methods of instruction.					
55	Schools and school systems should be held accountable for their performance on NSAT's.					
56	It is important that teachers understand the statistical methods used to report NSAT results.					
57	Teachers sometimes use NSAT results to modify curriculum.					
58	It may be useful to compare SCHOOLS in Nova Scotia using NSAT results.					
59	Most teachers would advise parents not to be concerned with NSAT results.					
60	Teachers NEVER use NSAT results at all.					

		S A	A	U	D	S D
61	It may be useful to compare SCHOOL SYSTEMS in Nova Scotia using NSAT results.					
62	Most parents of grade nine and grade twelve students are NOT aware of their child's NSAT scores.					
63	It is NOT appropriate for Universities and other post-secondary institutions to consider NSAT results when admitting students.					
64	Provincial education authorities have an obligation to monitor student achievement in Nova Scotia.					
65	Parents want to know about their child's achievement in school relative to other students in the province.					
66	The NSAT's are a valid measure of many (but not all) of the desired outcomes of education.					
67	The quality of classroom teaching has a significant effect on NSAT scores.					
68	We should continuously attempt to improve the NSAT's and gain greater understanding about their potential uses.					
69	Standardized achievement tests will be with us for many years to come.					
70	A serious attempt should be made to find out why some schools score consistently low on NSAT.					

		S A	A	U	D	S D
71	NSAT results are determined almost entirely by factors outside the schools.					
72	School authorities should NOT spent their time investigating discrepancies in NSAT scores among schools and school systems.					
73	While multiple choice type tests are limited they can provide useful information about student achievement.					
74	Teachers are familiar with the concepts of authentic and performance assessment.					
75	Student achievement should be measured but NOT by using multiple choice type tests.					
76	The interruption (at grade nine and twelve) caused by administering NSAT's is NOT warranted.					
77	To a large extent, schools and schools systems are responsible for the achievement of students in their care.					
78	The NSAT results do provide useful information to schools					
79	Comparing one student's achievement to another (on NSAT's) is INAPPROPRIATE.					
80	NSAT results are really quite meaningless.					

APPENDIX D
ITEM CORRELATION MATRIX

Correlation Matrix Clusters A, B, C, and D

Correlation Matrix for Variables: X1 ... X18

	s1.1	s1.3	s1.5	s1.6	s1.9	s1.4	s1.7	s1.16
s1.1	1							
s1.3	.228	1						
s1.5	.063	.119	1					
s1.6	.208	.159	.042	1				
s1.9	.206	.197	.083	.086	1			
s1.4	.164	.324	-.054	.153	.196	1		
s1.7	.015	.028	.012	.126	.033	.185	1	
s1.16	.029	.053	-.046	.051	.155	.208	-.021	1
s1.18	.066	.142	.021	.086	.148	.25	.126	.318
s1.8	.131	.16	.094	.025	.358	.17	.005	.173
s1.11	.105	.174	-.055	.013	.215	.239	.118	.241
s1.12	.079	.103	.028	.127	.235	.217	.176	.092
s1.17	.035	.163	.076	.025	.237	.214	.027	.268
s1.19	.079	.065	.09	.085	.126	.184	.195	.124
s1.13	.101	.144	.31	.07	.181	.085	.157	.035
s1.14	.086	.133	1	.06	.283	.218	.243	.129

Note: 1 case deleted with missing values

Correlation Matrix for Variables: X1 ... X18

	s1.1	s1.3	s1.5	s1.6	s1.9	s1.4	s1.7	s1.16
s1.15	.055	.107	.053	-.009	.156	.177	.029	.184
s1.20	-.005	.116	.05	.054	-.05	.037	-.038	.127

Correlation Matrix Clusters A, B, C, and D

Correlation Matrix for Variables: X₁ ... X₁₈

	s1.18	s1.8	s1.11	s1.12	s1.17	s1.19	s1.13	s1.14
s1.18	1							
s1.8	.134	1						
s1.11	.296	.19	1					
s1.12	.21	.247	.277	1				
s1.17	.284	.38	.225	.25	1			
s1.19	.226	.13	.227	.273	.174	1		
s1.13	.064	.159	.098	.177	.164	.16	1	
s1.14	.196	.151	.229	.282	.16	.354	.311	1

Correlation Matrix for Variables: X₁ ... X₁₈

	s1.18	s1.8	s1.11	s1.12	s1.17	s1.19	s1.13	s1.14
s1.15	.185	.211	.206	.191	.389	.089	.095	.187
s1.20	.14	.078	.075	.072	.078	-.032	-.026	-.05

Correlation Matrix Clusters E, F and G

Correlation Matrix for Variables: X₁ ... X₁₅

	s2.5	s2.7	s2.10	s2.13	s2.15	s2.1	s2.2	s2.3
s2.5	1							
s2.7	.119	1						
s2.10	.121	.107	1					
s2.13	.062	.321	-.068	1				
s2.15	.173	.36	.073	.267	1			
s2.1	.124	.132	.128	.045	.066	1		
s2.2	.025	.062	.091	-.005	.002	.319	1	
s2.3	.076	.057	.189	-.001	.099	.403	.193	1
s2.6	-.023	.157	.229	-.078	.06	.359	.288	.319
s2.9	.031	.188	.174	-.033	.112	.364	.304	.311
s2.4	.009	.066	.266	-.118	.042	.267	.241	.459
s2.8	.033	.154	.232	.003	.024	.265	.22	.228
s2.11	.052	.184	.25	-.053	.122	.354	.272	.382
s2.12	-.008	.091	.213	-.096	.062	.265	.199	.297
s2.14	.071	.071	.215	-.053	.071	.091	.089	.188

Correlation Matrix for Variables: X₁ ... X₁₅

	s2.6	s2.9	s2.4	s2.8	s2.11	s2.12	s2.14
s2.6	1						
s2.9	.382	1					
s2.4	.434	.305	1				
s2.8	.348	.425	.325	1			
s2.11	.387	.424	.449	.38	1		
s2.12	.375	.407	.365	.4	.552	1	
s2.14	.18	.209	.233	.217	.28	.312	1

APPENDIX E
NOVA SCOTIA ACHIEVEMENT TEST RESULTS

[illegible]

[illegible]