Educating All Our Children:

Departmental Directive on Student Assessment, Evaluation and Reporting

2010

Table of Contents:

Directive

1.	Purpose	5
2.	Authorities and References	5
3.	Rationale	6
4.	Definitions	6
5.	Scope	7
6.	Principles	
7.	Assessment	
	7.1 Classroom based assessments	7
	7.2 Systemic Assessments	8
8.	Reporting	9
	8.1 School-based Reporting	
	8.2 System-based Reporting	9
9.	Access to Information/Protection of Privacy (ATIPP)	
10.	Review of the SAER Directive	

Departmental Directive: Student Assessment, Evaluation & Reporting

1. Purpose

It is important for student progress to be consistently evaluated and reported in relation to NWT curricular outcomes. Information derived from the fair assessment and evaluation of students provides valuable information on student success in relation to curriculum expectations and identifies areas of strength and challenges at the student, school, district education authority, divisional education authority and territory-wide level.

- 1.1 Student progress will be evaluated in relation to NWT curricular outcomes.
- 1.2 Information from the assessment and evaluation of students shall be used to inform the student and his/her parents/guardians of the student's progress on NWT curricular outcomes.
- 1.3 Information from the assessment and evaluation of students shall be used to assist teachers and administrators in the planning and modification of educational programs and curricula for individual students and groups of students.
- 1.4 Information from the assessment and evaluation of students will be used to assist in determining the effectiveness of educational programs at the levels of the school, education authority and Department.
- 1.5 In the development and use of assessment and evaluation tools there will be a respect for and the integration of culture-based learning.

2. Authorities and References

The NWT Education Act establishes roles, responsibilities, standards and educational outcomes for student assessment, evaluation and reporting. Arising from the Act are regulations and directives that provide clarification and further direction on specific aspects of the Act. In addition to NWT legislation, regulations and directives there are documents that are widely accepted across Canada as principles and ethical guidelines for the fair assessment, evaluation and reporting of student progress.

- NWT Education Act (1996)
- NWT Access to Information and Protection of Privacy Act
- Departmental Directive on Inclusive Schooling 2006
- Departmental Directive on the Management of Information in the Student Record and Other Records Pertaining to Students
- Student Records Regulations
- Principles for Fair Student Assessment Practices for Education in Canada
- Directive on Teacher Growth and Evaluation in the Northwest Territories
- Rethinking Classroom Assessment with Purpose in Mind; Western and Northern Canadian Protocol, 2005
- NWT Approved Curricula
- Education Act Appeals Regulations

3. Rationale

Assessment is an integral component of all teaching and learning processes. Teacher year plans, unit or course plans, and daily lesson plans identify specific learning outcomes, and outline how student achievement will be measured in relation to those outcomes. What teachers assess, and how they assess it, is based on what teachers teach and how students learn. On an on-going basis, classroom teachers are in the best position to determine what needs to be assessed, to select methods for assessing, and to develop classroom assessment instruments. From time to time, teachers will also be required to administer external assessments that will validate what they already know about their students, provide additional information about the achievements and learning needs of students, and/or answer questions about quality and equity of the education system.

4. Definitions

It is beneficial if all schools in the NWT have a common understanding of terms used in assessment and evaluation. A more comprehensive glossary of terms is provided in the procedures for student assessment, evaluation and reporting.

assessment

The process of collecting information on student achievement and performance that includes a variety of assessment tasks designed to monitor and improve student learning. Assessments may include classroom observations, performance testing, informal reading inventories, writing folders/samples, portfolios, teacher-developed tests, standardized tests, checklists/rating scales, oral/written examinations and assignments.

evaluation The process of making judgments and decisions, based on interpretation

of the evidence gathered through assessment, as to what extent students have attained the prescribed outcomes and standards in the course curriculum and identifying what knowledge and skills still need to be learned. Evaluation also involves the making of decisions about the quality, value or worth of a response for the purpose of providing

descriptive feedback (formative) and marks (summative).

reporting The process of summarizing and clearly communicating student progress

on curricular outcomes to various client groups, including students, parents, administrators and the Department. Reporting also includes the recording of student data into the student record for each student and

submitting data requested by the Department.

5. Scope

This Directive applies to all education bodies, kindergarten through grade 12 in the Northwest Territories. This includes, but is not limited to, the Department of Education, Culture and Employment, Divisional Education Councils, District Education Authorities and personnel associated therewith. Information in the Directive includes mandatory requirements as well as recommendations that apply to Education Authorities.

6. Assessment

Classroom based assessments are for the purpose of evaluating the progress of individual students and systemic assessments are a means of evaluating the performance of education systems. Each of these evaluations relies on the collection of data gathered through the assessment of students. However, the data are collected for different purposes, the methods of assessing and evaluating are different, and the results are reported to different clientele.

A. Classroom based assessments

Throughout the school year, teachers are to use a variety of classroom assessments to collect information on each student. There are a variety of methods for doing classroom based assessment; all with the intent of allowing the student to demonstrate what they know and can do.

Teachers and school administrators are to ensure classroom based assessments will:

- I. Align with the NWT approved curriculum and classroom instruction.
- Be used to inform instruction.
- III. Include a variety and provide a balance of formative and summative assessments.

- IV. Provide a variety of opportunities and methods for students to demonstrate their progress toward the achievement of curricular outcomes.
- V. Be used as a means to diagnose individual student strengths and needs and, if necessary, to develop and implement a plan for intervention.
- VI. Have a clear purpose, including criteria for evaluation that is communicated to students in advance.

Superintendents and School authorities will ensure that

VII. Teachers are evaluated on their classroom assessment practices in accordance with the Directive for *Teacher Growth and Evaluation in the Northwest Territories* unless the jurisdiction uses an alternative teacher evaluation practice that has been approved by the Minister.

B. Systemic Assessments

The Department has a responsibility for establishing the standards for educational achievement at the territorial and local levels. As such, the Department will

- I. Establish curricular outcomes and NWT high school graduation requirements.
- II. Administer territory-wide assessments to:
 - establish and maintain standards;
 - monitor student achievement; and
 - assist in improving the quality of education.
- III. Establish rules and guidelines for the implementation of systemic assessments and the dissemination of the results.
- IV. Collect data from schools on student progress and factors affecting that progress.
- V. Wherever possible, provide school authorities summaries and councils with summaries or copies of the information collected.
- VI. Inform district education authorities as to the purpose of data collected from schools and how it will be used.

Superintendents and school authorities will

- VII. Ensure systemic assessments are conducted according to the requirements prescribed for the administration of the assessment.
- VIII. Ensure systemic assessments are conducted in a fair and appropriate manner. This includes arranging for accommodations and excusing a student if s/he is not capable of responding to the assessment in its original or approved accommodated form.
- IX. Ensure systemic assessments are only used for the purposes for which they were intended.

7. Reporting

Teachers, administrators and other professionals are accountable for regularly reporting progress to students, parents, the public, educators and other education decision-makers.

A. School-based Reporting

- I. Schools will use a variety of methods to communicate student progress and achievement to students and parents/guardians throughout the school year.
- II. Student report cards are to fulfill requirements specified by the Department and education authority as set out in the procedures for student assessment, evaluation and reporting.
- III. Parents/guardians are to be informed as to the education program of the student and the grade level of the curriculum the student is working on.
- IV. Parents are to be provided current information on their child's attendance and advised when it may be a concern.
- V. School boards and authorities are to encourage the use of 3-way conferencing and student portfolios as a means to communicate and enhance understanding of student achievement.
- VI. For each student in grades one through nine, teachers and schools are required to report to the Department the education program the student was following and the grade level of curriculum the student was working on for the majority of the school year. This is to be reported annually and in the subject areas of Language Arts and Mathematics.
- VII. Schools are to ensure that student records, paper and electronic, are kept current and fulfill required recording for student progress and attendance in accordance with Student Records Regulations and the Departmental Directive on the Management of Information in the Student Record and Other Records Pertaining to Students.

B. System-based Reporting

- Schools and education authorities will ensure information on attendance, student progress, assessment and evaluation requested by the Department is provided on time, accurate, complete and submitted in the format requested.
- II. Wherever possible, the Department will supply the education authorities summaries and findings of any information it has collected and/or assessments it has required to be implemented at the school, education authority and territory-wide level.
- III. Education authorities will be provided a copy of any Departmental press release on student attendance and achievement in advance of release and given sufficient time to prepare their own release and responses.

- IV. Education authorities are not to release to the public any systemic assessments results before the Department has released territory-wide data.
- V. Information from the assessment and evaluation of students is NOT to be released or discussed in a forum where there is a potential the information may lead to the identification of individual students and teachers. Information from student assessment and evaluation must meet the provisions laid out in the NWT Access to Information and Protection of Privacy Act (ATIPP).

8. Principles

Student assessment and reporting practices shall be consistent with the *Departmental Directive on Inclusive Schooling, 2006* and abide by the principles provided in *The Principles for Fair Student Assessment Practices for Education in Canada* provided as Appendix 1 of this Directive and highlighted below.

A. Classroom Assessments

I. Development and Choosing Methods for Assessment

Assessment methods should be appropriate for and compatible with the purpose and context of the assessment. Assessment methods should be based on the goals and objectives of instruction. The territorial curricula identify what schools should teach, and what students should learn, in terms of learning outcomes. These outcomes are learned within the context of the culture, heritage and language of the community.

A wide variety of sources and methods, suitable to the backgrounds and prior experiences of the students, should be used to obtain comprehensive information about student growth. These should include information from the students themselves, their parents and their peers, and informal or systematic observation of student behaviors in the classroom, class and homework assignments, tests, examinations and a wide variety of student products.

II. Collecting Assessment Information

Students should be provided with sufficient opportunity to demonstrate growth in knowledge, skills and attitudes in all five domains: physical, emotional, social, intellectual and spiritual. Students come to school from different backgrounds, they develop individual learning styles, and learn at different rates. Assessment should be a positive component of the learning experience, which enhances each

student's self-esteem and allows the student to demonstrate, in a variety of ways, what he or she knows and can do.

III. Judging and Scoring Student Performance

Procedures for judging or scoring student performance should be appropriate for the assessment method being used and be consistently applied and monitored. To increase consistency and validity, scoring procedures should be developed in conjunction with the assessment instrument. Informing students prior to the assessment about the scoring procedures to be followed should help ensure that both students and their teachers hold similar expectations.

IV. Summarizing and Interpreting Results

Procedures for summarizing and interpreting assessment results should yield accurate and informative representations of a student's performance in relation to the goals and objectives of instruction for the reporting period. Interpretation of information gathered for a reporting period for a student is a complex, and at times, controversial issue. Results should be interpreted against a pre-determined content standard and in relation to a student's background, prior learning experiences and future goals. Where appropriate, results should be interpreted in collaboration with students, parents/guardians, support team, and students' peers, to develop and provide accurate, meaningful and appropriate information about students' growth.

V. Reporting Assessment Findings

Assessment reports should be clear, accurate, and of practical value to the audiences for whom they are intended. Parents/guardians and students need to know what specific knowledge and skills the student is expected to acquire; what level of mastery is considered satisfactory and what level is excellent. There needs to be on-going communication between parents/guardians, students and teachers about what the student is able to do, what areas require further attention, and about ways of supporting the student in his or her learning.

B. Assessments Produced External to the Classroom

I. Developing and Selecting Methods for Assessment

Users should select assessment methods that have been developed to be as fair as possible for students who have different backgrounds or special needs. Users should select methods that are appropriate for the intended purposes and suitable for the students to be assessed. Assessments should not be used for purposes other than what they were intended for.

II. Collecting and Interpreting Assessment Information

Users should follow directions for proper administration of an assessment method and interpretation of assessment results. All individuals who administer the assessment, score the responses and interpret the results should have the necessary knowledge and skills to perform these tasks.

III. Informing Students Being Assessed

The students being assessed and, where applicable, their parents/guardians should be provided with complete information presented in an understandable way. Parents and students should be informed as to the process used to select the students participating in the assessment, format of the assessment, date and time for the assessment, conditions under which the results of the assessments will be released, and to whom the results will be released.

IV. Implement Mandated Assessment Programs

When the assessment is mandated, all persons involved in the assessment are to be informed as to the purposes of the assessment and the uses to be made of the results. Persons involved in the assessment include administrators, teachers, students and parents/guardians. All individuals involved in a systemic assessment are to follow the directions and fulfill the conditions specified for the implementation of the assessment. The use of systemic assessments, results reports and explanations of results are to be consistent with the purpose(s) of the assessment and the intended uses of the results.

Results of assessments should be interpreted in light of factors that may influence them. Important factors include characteristics of the students and the assessment method in relation to the learning outcomes being reported on.

Access to Information/Protection of Privacy (ATIPP) 9.

When students are at school, school staff act in the place of parents to protect students. Educators must ensure compliance with ATIPP concerning information on student assessments. Specifically, ATIPP legislation serves to protect student privacy through establishing controls over the collection, use, disclosure, protection, accuracy and retention of personal information.

Consideration must be given to issues such as:

- publication of assessment results;
- posting of student marks;
- display of graded student work;
- gathering and sharing of student exemplars;
- administration of assessments by outside agents.

10. **Review of the SAER Directive**

There will be a review of this directive five years after the date of implementation.

Jackson Lafferty

Minister of Education, Culture & Employment

Date

Principles for Fair Student Assessment Practices for Education in Canada

The Principles for Fair Student Assessment Practices for Education in Canada contains a set of principles and related guidelines generally accepted by professional organizations as indicative of fair assessment practice within the Canadian educational context. Assessments depend on professional judgment; the principles and related guidelines presented in this document identify the issues to consider in exercising this professional judgment and in striving for the fair and equitable assessment of all students.

Assessment is broadly defined in the *Principles* as the process of collecting and interpreting information that can be used (i) to inform students, and their parents/guardians where applicable, about the progress they are making toward attaining the knowledge, skills, attitudes, and behaviors to be learned or acquired, and (ii) to inform the various personnel who make educational decisions (instructional, diagnostic, placement, promotion, graduation, curriculum planning, program development, policy) about students. Principles and related guidelines are set out for both developers and users of assessments. Developers include people who construct assessment methods and people who set policies for particular assessment programs. Users include people who select and administer assessment methods, commission assessment development services, or make decisions on the basis of assessment results and findings. The roles may overlap, as when a teacher or instructor develops and administers an assessment instrument and then scores and interprets the students' responses, or when a ministry or department of education or local school system commissions the development and implementation of an assessment program and scoring services and makes decisions on the basis of the assessment results.

The Principles of Fair Student Assessment Practices for Education in Canada was developed by a Working Group guided by a Joint Advisory Committee. The Joint Advisory Committee included two representatives appointed by each of the following professional organizations: Canadian Education Association, Canadian School Boards Association, Canadian Association for School Administrators, Canadian Teachers' Federation, Canadian Guidance and Counselling Association, Canadian Association of School Psychologists, Canadian Council for Exceptional Children, Canadian Psychological Association, and Canadian Society for the Study of Education. In addition, the Joint Advisory Committee included a representative of the Provincial and Territorial Ministries and Departments of Education.

Financial support for the development and dissemination of the *Principles* was provided principally by the Walter and Duncan Gordon Charitable Foundation, with additional support provided by various Faculties, Institutes, and Colleges of Education and Provincial and Territorial Ministries and Departments of Education in Canada. This support is gratefully acknowledged.

The Joint Advisory Committee invites users to share their experiences in working with the *Principles* and to submit any suggestions that could be used to revise and improve the *Principles*. Comments and suggestions should be sent to the Joint Advisory Committee at the address shown below.

The Principles for Fair Student Assessment Practices for Education in Canada is not copyrighted. Reproduction and dissemination are encouraged. Please cite the Principles as follows:

Principles for Fair Student Assessment Practices for Education in Canada. (1993). Edmonton, Alberta: Joint Advisory Committee. (Mailing Address: Joint Advisory Committee, Centre for Research in Applied Measurement and Evaluation, 3-104 Education Building North, University of Alberta, Edmonton, Alberta, T6G 2G5).

The Principles for Fair Student Assessment Practices for Education in Canada is the product of a comprehensive effort to reach consensus on what constitutes sound principles to guide the fair assessment of students. The principles and their related guidelines should be considered neither exhaustive nor mandatory; however, organizations, institutions, and individual professionals who endorse them are committing themselves to endeavor to follow their intent and spirit so as to achieve fair and equitable assessments of students.

Organization and Use of the Principles

The principles and their related guidelines are organized in two parts. Part A is directed at assessments carried out by teachers at the elementary and secondary school levels. Part A is also applicable at the post-secondary level with some modifications, particularly with respect to whom assessment results are reported. Part B is directed at standardized assessments developed external to the classroom by commercial test publishers, provincial and territorial ministries and departments of education, and local school jurisdictions¹.

Five general principles of fair assessment practices are provided in each Part. Each principle is followed by a series of guidelines for practice. In the case of Part A where no prior sets of standards for fair practice exist, a brief comment accompanies each guideline to help clarify and illuminate the guideline and its application.

The Joint Advisory Committee recognizes that in the field of assessment some terms are defined or used differently by different groups of people. To maintain as much consistency in terminology as possible, an attempt has been made to employ generic terms in the *Principles*.

_

¹Boards, boroughs, counties, and school districts.

A. Classroom Assessments

Part A is directed toward the development and selection of assessment methods and their use in the classroom by teachers. Based on the conceptual framework provided in the *Standards for Teacher Competence in Educational Assessment of Students* (1990), it is organized around five interrelated themes:

- I. Developing and Choosing Methods for Assessment
- II. Collecting Assessment Information
- III. Judging and Scoring Student Performance
- IV. Summarizing and Interpreting Results
- V. Reporting Assessment Findings

The Joint Advisory Committee acknowledges that not all of the guidelines are equally applicable in all circumstances. However, consideration of the full set of principles and guidelines within Part A should help to achieve fairness and equity for the students to be assessed.

I. Developing and Choosing Methods for Assessment

Assessment methods should be appropriate for and compatible with the purpose and context of the assessment.

Assessment method is used here to refer to the various strategies and techniques that teachers might use to acquire assessment information. These strategies and techniques include, but are not limited to, observations, text- and curriculum-embedded questions and tests, paper-and-pencil tests, oral questioning, benchmarks or reference sets, interviews, peer-and self-assessments, standardized criterion- referenced and norm-referenced tests, performance assessments, writing samples, exhibitions, portfolio assessment, and project and product assessments. Several labels have been used to describe subsets of these alternatives, with the most common being "direct assessment," "authentic assessment," "performance assessment," and "alternative assessment." However, for the purpose of the *Principles*, the term assessment method has been used to encompass all the strategies and techniques that might be used to collect information from students about their progress toward attaining the knowledge, skills, attitudes, or behaviors to be learned.

1. Assessment methods should be developed or chosen so that inferences drawn about the lanowledge, skills, attitudes, and behaviors possessed by each student are valid and not open to misinterpretation.

Validity refers to the degree to which inferences drawn from assessments results are meaningful. Therefore, development or selection of assessment methods for collecting information should be clearly linked to the purposes for which inferences and decisions are to be made. For example, to monitor the progress of students as proofreaders and

editors of their own work, it is better to assign an actual writing task, to allow time and resources for editing (dictionaries, handbooks, etc.), and to observe students for evidence of proofreading and editing skill as they work than to use a test containing discrete items on usage and grammar that are relatively devoid of context.

2. Assessment methods should be clearly related to the goals and objectives of instruction, and be compatible with the instructional approaches used.

To enhance validity, assessment methods should be in harmony with the instructional objectives to which they are referenced. Planning an assessment design at the same time as planning instruction will help integrate the two in meaningful ways. Such joint planning provides an overall perspective on the knowledge, skills, attitudes, and behaviors to be learned and assessed, and the contexts in which they will be learned and assessed.

3. When developing or choosing assessment methods, consideration should be given to the consequences of the decisions to be made in light of the obtained information.

The outcomes of some assessments may be more critical than others. For example, misinterpretation of the level of performance on an end-of-unit test may result in incorrectly holding a student from proceeding to the next instructional unit in a continuous progress situation. In such "high-stake" situations, every effort should be made to ensure the assessment method will yield consistent and valid results. "Low stake" situations, such as determining if a student has correctly completed an in-class assignment, can be less stringent. Low stake assessments are often repeated during the course of a reporting period using a variety of methods. If the results are aggregated to form a summary comment or grade, the summary will have greater consistency and validity than its component elements.

4. More than one assessment method should be used to ensure comprehensive and consistent indications of student performance.

To obtain a more complete picture or profile of a student's knowledge, skills, attitudes, or behaviors, and to discern consistent patterns and trends, more than one assessment method should be used. Student knowledge might be assessed using completion items; process or reasoning skills might be assessed by observing performance on a relevant task; evaluation skills might be assessed by reflecting upon the discussion with a student about what materials to include in a portfolio. Self-assessment may help to clarify and add meaning to the assessment of a written communication, science project, piece of art work, or an attitude. Use of more than one method will also help minimize inconsistency brought about by different sources of measurement error (for example, poor performance because of an "off-day"; lack of agreement among items included in a test, rating scale, or questionnaire; lack of agreement among observers; instability across time).

5. Assessment methods should be suited to the backgrounds and prior experiences of students.

Assessment methods should be free from bias brought about by student factors extraneous to the purpose of the assessment. Possible factors to consider include culture, developmental stage, ethnicity, gender, socio-economic background, language, special interests, and special needs. Students' success in answering questions on a test or in an oral quiz, for example, should not be dependent upon prior cultural knowledge, such as understanding an allusion to a cultural tradition or value, unless such knowledge falls

within the content domain being assessed. All students should be given the same opportunity to display their strengths.

6. Content and language that would generally be viewed as sensitive, sexist, or offensive should be avoided.

The vocabulary and problem situation in each test item or performance task should not favour or discriminate against any group of students. Steps should be taken to ensure that stereotyping is not condoned. Language that might be offensive to particular groups of students should be avoided. A judicious use of different roles for males and females and for minorities and the careful use of language should contribute to more effective and, therefore, fairer assessments.

7. Assessment instruments translated into a second language or transferred from another context or location should be accompanied by evidence that inferences based on these instruments are valid for the intended purpose.

Translation of an assessment instrument from one language to another is a complex and demanding task. Similarly, the adoption or modification of an instrument developed in another country is often not simple and straightforward. Care must be taken to ensure that the results from translated and imported instruments are not misinterpreted or misleading.

II. Collecting Assessment Information

Students should be provided with a sufficient opportunity to demonstrate the knowledge, skills, attitudes, or behaviors being assessed.

Assessment information can be collected in a variety of ways (observations, oral questioning, interviews, oral and written reports, paper-and-pencil tests). The guidelines which follow are not all equally applicable to each of these procedures.

1. Students should be told why assessment information is being collected and how this information will be used.

Students who know the purpose of an assessment are in a position to respond in a manner that will provide information relevant to that purpose. For example, if students know that their participation in a group activity is to be used to assess cooperative skills, they can be encouraged to contribute to the activity. If students know that the purpose of an assessment is to diagnose strengths and weaknesses rather than to assign a grade, they can be encouraged to reveal weaknesses as well as strengths. If the students know that the purpose is to assign a grade, they are well advised to respond in a way that will maximize strength. This is especially true for assessment methods that allow students to make choices, such as with optional writing assignments or research projects.

2. An assessment procedure should be used under conditions suitable to its purpose and form.

Optimum conditions should be provided for obtaining data from and information about students so as to maximize the validity and consistency of the data and information collected. Common conditions include such things as proper light and ventilation, comfortable room temperature, and freedom from distraction (e.g., movement in and out of the room, noise). Adequate work-space, sufficient materials, and adequate time limits appropriate to the purpose and form of the assessment are also necessary. For example, if the intent is to assess student participation in a small group, adequate work space should be provided for each student group, with sufficient space between subgroups so that the groups do not interfere with or otherwise influence one another and so that the teacher has the same opportunity to observe and assess each student within each group.

3. In assessments involving observations, checklists, or rating scales, the number of characteristics to be assessed at one time should be small enough and concretely described so that the observations can be made accurately.

Student behaviors often change so rapidly that it may not be possible simultaneously to observe and record all the behavior components. In such instances, the number of components to be observed should be reduced and the components should be described as concretely as possible. One way to manage an observation is to divide the behavior into a series of components and assess each component in sequence. By limiting the number of components assessed at one time, the data and information become more focused, and time is not spent observing later behavior until prerequisite behaviors are achieved.

4. The directions provided to students should be clear, complete, and appropriate for the ability, age and grade level of the students.

Lack of understanding of the assessment task may prevent maximum performance or display of the behavior called for. In the case of timed assessments, for example, teachers should describe the time limits, explain how students might distribute their time among parts for those assessment instruments with parts, and describe how students should record their responses. For a portfolio assessment, teachers should describe the criteria to be used to select the materials to be included in a portfolio, who will select these materials, and, if more than one person will be involved in the selection process, how the judgments from the different people will be combined. Where appropriate, sample material and practice should be provided to further increase the likelihood that instructions will be understood.

5. In assessments involving selection items (e.g., true-false, multiple-choice), the directions should encourage students to answer all items without threat of penalty.

A correction formula is sometimes used to discourage "guessing" on selection items. The formula is intended to encourage students to omit items for which they do not know the answer rather than to "guess" the answer. Because research evidence indicates that the benefits expected from the correction are not realized, the use of the formula is discouraged. Students should be encouraged to use whatever partial knowledge they have when choosing their answers, and to answer all items.

6. When collecting assessment information, interactions with students should be appropriate and consistent.

Care must be taken when collecting assessment information to treat all students fairly. For example, when oral presentations by students are assessed, questioning and probes should be distributed among the students so that all students have the same opportunity to demonstrate their knowledge. While writing a paper-and-pencil test, a student may ask to have an ambiguous item clarified, and, if warranted, the item should be explained to the entire class.

7. Unanticipated circumstances that interfere with the collection of assessment information should be noted and recorded.

Events such as a fire drill, an unscheduled assembly, or insufficient materials may interfere in the way in which assessment information is collected. Such events should be recorded and subsequently considered when interpreting the information obtained.

8. A written policy should guide decisions about the use of alternate procedures for collecting assessment information from students with special needs and students whose proficiency in the language of instruction is inadequate for them to respond in the anticipated manner.

It may be necessary to develop alternative assessment procedures to ensure a consistent and valid assessment of those students who, because of special needs or inadequate language, are not able to respond to an assessment method (for example, oral instead of written format, individual instead of group administered, translation into first language, providing additional time). The use of alternate procedures should be guided by a written policy developed by teachers, administrators, and other jurisdictional personnel.

III. Judging and Scoring Student Performance

Procedures for judging or scoring student performance should be appropriate for the assessment method used and be consistently applied and monitored.

Judging and scoring refers to the process of determining the quality of a student's performance, the appropriateness of an attitude or behavior, or the correctness of an answer. Results derived from judging and scoring may be expressed as written or oral comments, ratings, categorizations, letters, numbers, or as some combination of these forms.

1. Before an assessment method is used, a procedure for scoring should be prepared to guide the process of judging the quality of a performance or product, the appropriateness of an attitude or behavior, or the correctness of an answer.

To increase consistency and validity, properly developed scoring procedures should be used. Different assessment methods require different forms of scoring. Scoring selection items (true-false, multiple-choice, matching) requires the identification of the correct or, in some instances, best answer. Guides for scoring essays might include factors such as the major points to be included in the "best answer" or models or exemplars corresponding to different levels of performance at different age levels and against which comparisons can be made. Procedures for judging other performances or products might include specification of the characteristics to be rated in performance terms and, to the extent possible, clear descriptions of the different levels of performance or quality of a product.

2. Before an assessment method is used, students should be told how their responses or the information they provide will be judged or scored.

Informing students prior to the use of an assessment method about the scoring procedures to be followed should help ensure that similar expectations are held by both students and their teachers.

3. Care should be taken to ensure that results are not influenced by factors that are not relevant to the purpose of the assessment.

Various types of errors occur in scoring, particularly when a degree of subjectivity is involved (e.g., marking essays, rating a performance, judging a debate). For example, if the intent of a written communication is to assess content alone, the scoring should not be influenced by stylistic factors such as vocabulary and sentence structure. Personal bias errors are indicated by a general tendency to rate all students in approximately the same way (e.g., too generously or too severely). Halo effects can occur when a rater's general impression of a student influences the rating of individual characteristics or when a previous rating influences a subsequent rating. Pooled results from two or more independent raters (teachers, other students) will generally produce a more consistent description of student performance than a result obtained from a single rater. In combining results, the personal biases of individual raters tend to cancel one another.

4. Comments formed as part of scoring should be based on the responses made by the students and presented in a way that students can understand and use them.

Comments, in oral and written form, are provided to encourage learning and to point out correctable errors or inconsistencies in performance. In addition, comments can be used to clarify a result. Such feedback should be based on evidence pertinent to the learning outcomes being assessed.

5. Any changes made during scoring should be based upon a demonstrated problem with the initial scoring procedure. The modified procedure should then be used to rescore all previously scored responses.

Anticipating the full range of student responses is a difficult task for several forms of assessment. There is always the danger that unanticipated responses or incidents that are relevant to the purposes of the assessment may be overlooked. Consequently, scoring should be continuously monitored for unanticipated responses and these responses should be taken into proper account.

6. An appeal process should be described to students at the beginning of each school year or course of instruction that they may use to appeal a result.

Situations may arise where a student believes a result incorrectly reflects his/her level of performance. A procedure by which students can appeal such a situation should be developed and made known to them. This procedure might include, for example, checking for addition or other recording errors or, perhaps, judging or scoring by a second qualified person.

IV. Summarizing and Interpreting Results

Procedures for summarizing and interpreting asssessment results should yield accurate and informative representations of a student's performance in relation to the goals and objectives of instruction for the reporting period.

Summarizing and interpreting results refers to the procedures used to combine assessment results in the form of summary comments and grades which indicate both a student's level of performance and the valuing of that performance.

1. Procedures for summarizing and interpreting results for a reporting period should be guided by a written policy.

Summary comments and grades, when interpreted, serve a variety of functions. They inform students of their progress. Parents, teachers, counsellors, and administrators use them to guide learning, determine promotion, identify students for special attention (e.g., honours, remediation), and to help students develop future plans. Comments and grades also provide a basis for reporting to other schools in the case of school transfer and, in the case of senior high school students, post-secondary institutions and prospective employers. They are more likely to serve their many functions and those functions are less likely to be confused if they are guided by a written rationale or policy sensitive to these different needs. This policy should be developed by teachers, school administrators, and other jurisdictional personnel in consultation with representatives of the audiences entitled to receive a report of summary comments and grades.

2. The way in which summary comments and grades are formulated and interpreted should be explained to students and their parents/guardians.

Students and their parents/guardians have the "right-to-know" how student performance is summarized and interpreted. With this information, they can make constructive use of the findings and fully review the assessment procedures followed.

It should be noted that some aspects of summarizing and interpreting are based upon a teacher's best judgment of what is good or appropriate. This judgment is derived from training and experience and may be difficult to describe specifically in advance. In such circumstances, examples might be used to show how summary comments and grades were formulated and interpreted.

3. The individual results used and the process followed in deriving summary comments and grades should be described in sufficient detail so that the meaning of a summary comment or grade is clear.

Summary comments and grades are best interpreted in the light of an adequate description of the results upon which they are based, the relative emphasis given to each result, and the process followed to combine the results. Many assessments conducted during a reporting period are of a formative nature. The intent of these assessments (e.g., informal observations, quizzes, text-and-curriculum embedded questions, oral questioning) is to inform decisions regarding daily learning, and to inform or otherwise refine the instructional sequence. Other assessments are of a summative nature. It is the summative

assessments that should be considered when formulating and interpreting summary comments and grades for the reporting period.

4. Combining disparate kinds of results into a single summary should be done cautiously. To the extent possible, achievement, effort, participation, and other behaviors should be graded separately.

A single comment or grade cannot adequately serve all functions. For example, letter grades used to summarize achievement are most meaningful when they represent only achievement. When they include other aspects of student performance such as effort, amount (as opposed to quality) of work completed, neatness, class participation, personal conduct, or punctuality, not only do they lose their meaningfulness as a measure of achievement, but they also suppress information concerning other important aspects of learning and invite inequities. Thus, to more adequately and fairly summarize the different aspects of student performance, letter grades for achievement might be complemented with alternate summary forms (e.g., checklists, written comments) suitable for summarizing results related to these other behaviors.

5. Summary comments and grades should be based on more than one assessment result so as to ensure adequate sampling of broadly defined learning outcomes.

More than one or two assessments are needed to adequately assess performance in multifacet areas such as Reading. Under-representation of such broadly defined constructs can be avoided by ensuring that the comments and grades used to summarize performance are based on multiple assessments, each referenced to a particular facet of the construct.

6. The results used to produce summary comments and grades should be combined in a way that ensures that each result receives its intended emphasis or weight.

When the results of a series of assessments are combined into a summary comment, care should be taken to ensure that the actual emphasis placed on the various results matches the intended emphasis for each student.

When numerical results are combined, attention should be paid to differences in the variability, or spread, of the different sets of results and appropriate account taken where such differences exist. If, for example, a grade is to be formed from a series of paper-and-pencil tests, and if each test is to count equally in the grade, then the variability of each set of scores must be the same.

7. The basis for interpretation should be carefully described and justified.

Interpretation of the information gathered for a reporting period for a student is a complex and, at times, controversial issue. Such information, whether written or numerical, will be of little interest or use if it is not interpreted against some pertinent and defensible idea of what is good and what is poor. The frame of reference used for interpretation should be in accord with the type of decision to be made. Typical frames of reference are performance in relation to pre-specified standards, performance in relation to peers, performance in relation to aptitude or expected growth, and performance in terms of the amount of improvement or amount learned. If, for example, decisions are to be made as to whether or not a student is ready to move to the next unit in an instructional sequence, interpretations based on pre-specified standards would be most relevant.

8. Interpretations of assessment results should take account of the backgrounds and learning experiences of the students.

Assessment results should be interpreted in relation to a student's personal and social context. Among the factors to consider are age, ability, gender, language, motivation, opportunity to learn, self-esteem, socio-economic background, special interests, special needs, and "test-taking" skills. Motivation to do school tasks, language capability, or home environment can influence learning of the concepts assessed, for example. Poor reading ability, poorly developed psycho-motor or manipulative skills, lack of test-taking skills, anxiety, and low self-esteem can lead to lower scores. Poor performance in an assessment may be attributable to a lack of opportunity to learn because required learning materials and supplies were not available, learning activities were not provided, or inadequate time was allowed for learning. When a student performs poorly, the possibility that one or more factors such as these might have interfered with a student's response or performance should be considered.

9. Assessment results that will be combined into summary comments and grades should be stored in a way that ensures their accuracy at the time they are summarized and interpreted.

Comments and grades and their interpretations, formulated from a series of related assessments, can be no better than the data and information upon which they are based. Systematic data control minimizes errors which would otherwise be introduced into a student's record or information base, and provides protection of confidentiality.

10. Interpretations of assessment results should be made with due regard for limitations in the assessment methods used, problems encountered in collecting the information and judging or scoring it, and limitations in the basis used for interpretation.

To be valid, interpretations must be based on results determined from assessment methods that are relevant and representative of the performance assessed. Administrative constraints, the presence of measurement error, and the limitations of the frames of reference used for interpretation also need to be accounted for.

V. Reporting Assessment Findings

Assessment reports should be clear, accurate, and of practical value to the audiences for whom they are intended.

1. The reporting system for a school or jurisdiction should be guided by a written policy. Elements to consider include such aspects as audiences, medium, format, content, level of detail, frequency, timing, and confidentiality.

The policy to guide the preparation of school reports (e.g., reports of separate assessments; reports for a reporting period) should be developed by teachers, school administrators, and other jurisdictional personnel in consultation with representatives of the audiences entitled to receive a report. Cooperative participation not only leads to more adequate and helpful reporting, but also increases the likelihood that the reports will be understood and used by those for whom they are intended.

Written and oral reports should contain a description of the goals and objectives of instruction to which the assessments are referenced.

The goals and objectives that guided instruction should serve as the basis for reporting. A report will be limited by a number of practical considerations, but the central focus should be on the instructional objectives and the types of performance that represent achievement of these objectives.

3. Reports should be complete in their descriptions of strengths and weaknesses of students, so that strengths can be built upon and problem areas addressed.

Reports can be incorrectly slanted towards "faults" in a student or toward giving unqualified praise. Both biases reduce the validity and utility of assessment. Accuracy in reporting strengths and weaknesses helps to reduce systematic error and is essential for stimulating and reinforcing improved performance. Reports should contain the information that will assist and guide students, their parents/guardians, and teachers to take relevant follow-up actions.

4. The reporting system should provide for conferences between teachers and parents/guardians. Whenever it is appropriate, students should participate in these conferences.

Conferences scheduled at regular intervals and, if necessary, upon request provide parents/guardians and, when appropriate, students with an opportunity to discuss assessment procedures, clarify and elaborate their understanding of the assessment results, summary comments and grades, and reports, and, where warranted, to work with teachers to develop relevant follow-up activities or action plans.

5. An appeal process should be described to students and their parents/guardians at the beginning of each school year or course of instruction that they may use to appeal a report.

Situations may arise where a student and his/her parents/guardian believe the summary comments and grades inaccurately reflect the level of performance of the student. A procedure by which they can appeal such a situation should be developed and made known to them (for example, in a school handbook or newsletter provided to students and their parents/guardians at the beginning of the school year).

6. Access to assessment information should be governed by a written policy that is consistent with applicable laws and with basic principles of fairness and human rights.

A written policy, developed by teachers, administrators, and other jurisdictional personnel, should be used to guide decisions regarding the release of student assessment information. Assessment information should be available to those people to whom it applies -- students and their parents/guardians, and to teachers and other educational personnel obligated by profession to use the information constructively on behalf of students. In addition, assessment information might be made available to others who justify their need for the information (e.g., post-secondary institutions, potential employers, researchers). Issues of informed consent should also be addressed in this policy.

7. Transfer of assessment information from one school to another should be guided by a written policy with stringent provisions to ensure the maintenance of confidentiality.

To make a student's transition from one school to another as smooth as possible, a clear policy should be prepared indicating the type of information to go with the student and the form in which it will be reported. Such a policy, developed by jurisdictional and ministry personnel, should ensure that the information transferred will be sent by and received by the appropriate person within the "sending" and "receiving" schools respectively.

B. Assessments Produced External to the Classroom

Part B applies to the development and use of standardized assessment methods used in student admissions, placement, certification, and educational diagnosis, and in curriculum and program evaluation. These methods are primarily developed by commercial test publishers, ministries and departments of education, and local school systems.

The principles and accompanying guidelines are organized in terms of four areas:

- I. Developing and Selecting Methods for Assessment
- II. Collecting and Interpreting Assessment Information
- III. Informing Students Being Assessed
- IV. Implementing Mandated Assessment Programs

The first three areas of Part B are adapted from the Code of Fair Testing Practices for Education (1988) developed in the United States. The principles and guidelines as modified in these three sections are intended to be consistent with the Guidelines for Educational and Psychological Testing (1986) developed in Canada. The fourth area has been added to contain guidelines particularly pertinent for mandated educational assessment and testing programs developed and conducted at the national, provincial, and local levels.

I. Developing and Selecting Methods for Assessment

Developers of assessment methods should strive to make them as fair as possible for use with students who have different backgrounds or special needs. Developers should provide the information users need to select methods appropriate to their assessment needs.

Users should select assessment methods that have been developed to be as fair as possible for students who have different backgrounds or special needs. Users should select methods that are appropriate for the intended purposes and suitable for the students to be assessed.

Developers should:

Users should:

- Define what the assessment method is intended to measure and how it is to be used. Describe the characteristics of the students with which the method may be used.
- 2. Warn users against common misuses of the assessment method.
- Describe the process by which the method was developed. Include a description of the theoretical basis, rationale for selection of content and procedures, and derivation of scores.
- 4. Provide evidence that the assessment method yields results that satisfy its intended purpose(s).
- 5. Investigate the performance of students with special needs and students from different backgrounds. Report evidence of the consistency and validity of the results produced by the assessment method for these groups.
- Provide potential users with representative samples or complete copies of questions or tasks, directions, answer sheets, score reports, guidelines for interpretation, and manuals.

- 1. Determine the purpose(s) for assessment and the characteristics of the students to be assessed. Then select an assessment method suited to that purpose and type of student.
- Avoid using assessment methods for purposes not specifically recommended by the developer unless evidence is obtained to support the intended use.
- 3.

 Review available assessment methods for relevance of content and appropriateness of scores with reference to the intended purpose(s) and characteristics of the students to be assessed.
- Read independent evaluations of the methods being considered. Look for evidence supporting the claims of developers with reference to the intended application of each method.
- 5. Ascertain whether the content of the assessment method and the norm group(s) or comparison group(s) are appropriate for the students to be assessed. For assessment methods developed in other regions or countries, look for evidence that the characteristics of the norm group(s) or comparison group(s) are comparable to the characteristics of the students to be assessed.
- 6. Examine specimen sets, samples or complete copies of assessment instruments, directions, answer sheets, score reports, guidelines for interpretation, and manuals and judge their appropriateness for the intended application.

- 7. Review printed assessment methods and related materials for content or language generally perceived to be sensitive, offensive, or misleading.
- 8. Describe the specialized skills and training needed to administer an assessment method correctly, and the specialized knowledge to make valid interpretations of scores.

- 9. Limit sales of restricted assessment materials to persons who possess the necessary qualifications.
- 10. Provide for periodic review and revision of content and norms, and, if applicable, passing or cut-off scores, and inform users.
- Provide evidence of the comparability of different forms of an instrument where the forms are intended to be interchangeable, such as parallel forms or the adaptation of an instrument for computer administration.
- Provide evidence that an assessment 12. method translated into a second language is valid for use with the second language. This information should be provided in the second language.
- Advertise an assessment method in a way 13. that states it can be used only for the purposes for which it was intended.

- 7. Review printed assessment methods and related materials for content or language that would offend or mislead the students to be assessed.
- 8. Ensure that all individuals who administer the assessment method, score the responses, and interpret the results have the necessary knowledge and skills to perform these tasks (e.g., learning assistance teachers, speech and language pathologists, counsellors, school psychologists, psychologists).
 - 9. Ensure access to restricted assessment materials is limited to persons with the necessary qualifications.
- 10. Obtain information about the appropriateness of content, the recency of norms, and, if applicable, the appropriateness of the cut-off scores for use with the students to be assessed.
- 11. Obtain information about the comparability of interchangeable forms, including computer adaptations.
- Obtain evidence about the validity of the use of an assessment method translated into a second language.
 - Verify advertising claims made for an assessment method.

II. Collecting and Interpreting Assessment Information

13.

Developers should provide information to help users administer an assessment method correctly and interpret assessment results accurately. Users should follow directions for proper administration of an assessment method and interpretation of assessment results.

Developers should:

Users should:

- Provide clear instructions for administering the assessment method and identify the qualifications that should be held by the people who should administer the method.
- Ensure that the assessment method is administered by qualified personnel or under the supervision of qualified personnel.
- 2. When feasible, make available appropriately modified forms of assessment methods for students with special needs or whose proficiency in the original language of administration is inadequate to respond in the anticipated manner.
- 2. When necessary and feasible, use appropriately modified forms of assessment methods with students who have special needs or whose proficiency in the original language of administration is inadequate to respond in the anticipated manner.

Ensure that instruments translated from one language to another are administered by persons who are proficient in the translated language.

- 3. Provide answer keys and describe procedures for scoring when scoring is to be done by the user.
- 3. Follow procedures for scoring as set out for the assessment method.
- 4. Provide score reports or procedures for generating score reports that describe assessment results clearly and accurately. Identify and explain possible misinterpretations of the scores yielded by the scoring system (grade equivalents, percentile ranks, standard scores) used.
- 4. Interpret scores taking into account the limitations of the scoring system used. Avoid misinterpreting scores on the basis of unjustified assumptions about the scoring system (grade-equivalents, percentile ranks, standard scores) used.

- 5. Provide evidence of the effects on assessment results of such factors as speed, test- taking strategies, and attempts by students to present themselves favourably in their responses.
- 6. Warn against using published norms with students who are not part of the population from which the norm or comparison sample was selected or when the prescribed assessment method has been modified in any way.

7. Describe how passing and cut-off scores, where used, were set and provide evidence regarding rates of misclassification.

8. Provide evidence to support the use of any computer scoring or computer generated interpretations. The documentation should include the rationale for such scoring and interpretations and their comparability with the results of scoring and interpretations made by qualified judges.

- 5. Interpret scores taking into account the effects of such factors as speed, test-taking strategies, and attempts by students to present themselves favourably in their responses.
- 6. Interpret scores taking account of major differences between the norm group(s) or comparison group(s) and the students being assessed. Also take account of discrepancies between recommended and actual procedures and differences in familiarity with the assessment method between the norm group(s) and the students being assessed.

Examine the need for local norms, and, if called for, develop these norms.

7. Explain how passing or cut-off scores were set and discuss the appropriateness of these scores in terms of rates of misclassification.

Examine the need for local passing or cut-off scores and, if called for, reset these scores.

- 8. Ensure that any computer administration and computer interpretations of assessment results are accurate and appropriate for the intended use. If necessary, ensure that relevant information not included in computer reports is also considered.
- 9. Observe jurisdictional policies regarding storage of and subsequent access to the results. Ensure that computer files are not accessible to unauthorized users.
- 10. Ensure that all copyright and user agreements are observed.

III. Informing Students Being Assessed

Direct communication with those being assessed may come from either the developer or the user of the assessment method. In either case, the students being assessed and, where applicable, their parents/guardians should be provided with complete information presented in an understandable way.

Developers or Users should:

- 1. Develop materials and procedures for informing the students being assessed about the content of the assessment, types of question formats used, and appropriate strategies, if any, for responding.
- 2. Obtain informed consent from students or, where applicable, their parents/guardians in the case of individual assessments to be used for identification or placement purposes.
- 3. Provide students or, where applicable, their parents/guardians with information to help them decide whether to participate in the assessment when participation is optional.
- 4. Provide information to students or, where applicable, their parents/guardians of alternate assessment methods where available and applicable.

Control of results may rest with either the developer or user of the assessment method. In either case, the following steps should be followed.

Developers or Users should:

- 1. Provide students or, where applicable, their parents/guardians with information as to their rights to copies of instruments and completed answer forms, to reassessment, to rescoring, or to cancellation of scores and other records.
- 2. Inform students or, where applicable, their parents/guardians of the length of time assessment results will be kept on file and of the circumstances under which the assessment results will be released and to whom.

3. Describe the procedures that students or, where applicable, their parents/guardians may follow to register concerns about the assessment and endeavor to have problems resolved.

IV. Implementing Mandated Assessment Programs²

Under some circumstances, the administration of an assessment method is required by law. In such cases, the following guidelines should be added to the applicable guidelines outlined in Sections I, II, and III of Part B.

Developers and Users should:

- 1. Inform all persons with a stake in the assessment (administrators, teachers, students, parents/guardians) of the purpose(s) of the assessment, the uses to be made of the results, and who has access to the results.
- 2. Design and describe procedures for developing or choosing the methods of assessment, selecting students where sampling is used, administering the assessment materials, and scoring and summarizing student responses.
- 3. Interpret results in light of factors that might influence them. Important factors to consider include characteristics of the students, opportunity to learn, and comprehensiveness and representativeness of the assessment method in terms of the learning outcomes to be reported on.
- 4. Specify procedures for reporting, storing, controlling access to, and destroying results.
- 5. Ensure reports and explanations of results are consistent with the purpose(s) of the assessment, the intended uses of the results and the planned access to the results.
- 6. Provide reports and explanations of results that can be readily understood by the intended audience(s). If necessary, employ multiple reports designed for different audiences.

References

Code of Fair Testing Practices for Education. (1988). Washington, D.C.: Joint Committee on Testing Practices.

² The Joint Advisory Committee wishes to point out it has not taken a position on the value of mandated assessment and testing programs. Rather, given the presence of these programs, the intent of the guidelines presented in Section IV, when combined with applicable guidelines in the first three sections of Part B, is to help ensure fairness and equity for the students being assessed.

Guidelines for Educational and Psychological Testing. (1986). Ottawa, Ont.: Canadian Psychological Association.

Standards for Teacher Competence in Educational Assessment of Students. (1990). Washington: D.C.: American Federation of Teachers, National Council on Measurement in Education, and National Educational Association.

The membership of the Working Group (WG) that developed the *Principles for Fair Student Assessment Practices for Education in Canada* and of the Joint Advisory Committee that oversaw the development was as follows:

Marvin Betts Michael Jackson Jean Pettifor Gary Broker Michel Laurier (WG) Sharon Robertson Clement Dassa (WG) Tom Maguire (WG) Don Saklofske Romulo Magsino Dick Dodds Marvin Simner Linda McAlpine Tom Dunn (WG) Marielle Simon (WG) Ross Traub (WG) **Bob Gilchrist** Allan McDonald Nicholas Head Stirling McDowell Sue Wagner Kim Wolff Douglas Hodgkinson Craig Melvin Kathy Oberle (WG) Barbara Holmes (WG) Todd Rogers (Chair, Frank Oliva Working Group and Committee) Joint Advisory