Principles for Fair Student Assessment Practices for Education in Canada

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 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

- 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
- 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 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* 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 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. It is also applicable at the postsecondary 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, departments of education and local school jurisdictions (boards, boroughs, counties and school districts).

Five general principles of fair assessment practices are provided in each part. Each principle is followed by a series of guidelines for practice. In 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*.

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* (American Federation of Teachers 1990), it is organized around five interrelated themes:

1. Developing and choosing methods for assessment
2. Collecting assessment information
3. Judging and scoring student performance
4. Summarizing and interpreting results
5. Reporting assessment findings

The Joint Advisory Committee acknowledges that not all 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.

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 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 inferences drawn about the knowledge, 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 assessment 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 and so on) 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.

Some assessment outcomes 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 on 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 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, socioeconomic background, language, special interests and special needs. Students' success in answering questions on a test or in an oral quiz, for example, should not depend on 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 insensitive, sexist or offensive should be avoided.

The vocabulary and problem situation in each test item or performance task should not favor 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, females and minorities and the careful use of language should contribute to more effective and 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, adopting or modifying 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.

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 various ways (observations, oral questioning, interviews, oral and written reports, paper-and-pencil tests). The guidelines that 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 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 students know 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 to maximize validity and consistency. Common conditions include such things as proper light and ventilation, comfortable room temperature and freedom from distraction (for example, 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 the groups do not interfere with or otherwise influence one another. This gives the teacher 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 their ability, age and grade level.

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 criteria to be used to select materials to be included, who will select these materials and, if more

than one person will be involved in the selection process, how judgments will be combined. Where appropriate, sample material and practice should be provided to increase the likelihood that instructions will be understood.

➢5. In assessments involving selection items (for example, 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 the benefits expected from the correction are not realized, 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 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 alternative 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). Use of alternative procedures should be guided by a written policy developed by teachers, administrators and other jurisdictional personnel.

Judging and Scoring Student Performance

Procedures for judging and 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 characteristics to be rated in performance terms and, to the extent possible, clear descriptions of 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 about scoring procedures to be followed prior to the use of an assessment method should help ensure similar expectations are held by both students and their teachers.

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

Various errors occur in scoring, particularly when a degree of subjectivity is involved (for example, 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 (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, personal biases of individual raters tend to cancel one another.

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

Comments, in oral and written form, are provided to encourage learning and to point out correctable errors or inconsistencies in performance. Comments can also 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 on 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 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 account.

➢6. A process students may use to appeal a result should be described to them at the beginning of each school year or course of instruction.

Situations may arise where a student believes a result incorrectly reflects his or 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 judging or scoring by a second qualified person.

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.

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 (honors, 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, postsecondary 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.

Some aspects of summarizing and interpreting are based on 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 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 on 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 (for example, 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 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 also suppress information concerning other important aspects of learning and invite inequities. Thus, to more adequately and fairly summarize different aspects of student performance, letter grades for achievement might be complemented with alternate summary forms (checklists, written comments) suitable for summarizing results related to other behaviors.

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

More than one or two assessments are needed to adequately assess performance in multifaceted areas such as reading. Underrepresentation 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 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 the actual emphasis placed on 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 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 prespecified 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 a student is ready to move to the next unit in an instructional sequence, interpretations based on prespecified 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, socioeconomic 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 psychomotor 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 the response or performance should be considered.

➤9. Assessment results to 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 on which they are based. Systematic data control minimizes errors which would otherwise be introduced into a

student's record or information base and protects 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 relevant and representative of the performance assessed. Administrative constraints, the presence of measurement error and limitations of the frames of reference used for interpretation also need to be accounted for.

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 (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 the reports will be understood and used by those for whom they are intended.

➤2. 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 their achievement.

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

Reports can be incorrectly slanted toward 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 reduce systematic error and is essential for stimulating and reinforcing improved performance. Reports should contain information to 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 appropriate, students should participate in these conferences.

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

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

Situations may arise where a student and the parents/guardians believe the summary comments and grades inaccurately reflect the student's level of performance. 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 consistent with applicable laws and 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 to whom it applies—students and their parents/guardians, 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 (postsecondary 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 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 the information transferred will be sent and received by the appropriate people 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, curriculum and program evaluation. These methods are primarily developed by commercial test publishers, ministries, departments of education and local school systems.

The principles and accompanying guidelines are organized in four areas:

1. Developing and selecting methods for assessment
2. Collecting and interpreting assessment information
3. Informing students being assessed
4. Implementing mandated assessment programs

The first three areas of Part B are adapted from the *Code of Fair Testing Practices for Education* (Joint Committee on Testing Practices 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* (Canadian Psychological Association 1986). 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.

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.

Developers' Responsibilities

➢1. 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.

➢ 3. 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 the assessment method yields results that satisfy its intended purpose.

➢ 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.

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 appropriate for the intended purposes and suitable for students to be assessed.

Users' Responsibilities

➢1. Determine the purpose for assessment and the characteristics of students to be assessed. Then select an assessment method suited to that purpose and type of student.

➢2. 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 and characteristics of students to be assessed.

➢4. Read independent evaluations of methods being considered. Look for evidence supporting 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 or comparison group 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 or comparison group are comparable to the characteristics of students to be assessed.

➢ 6. Provide potential users with representative samples or complete copies of questions or tasks, directions, answer sheets, score reports, guidelines for interpretation and manuals.	➢ 6. Examine specimen sets, samples or complete copies of assessment instruments, directions, answer sheets, score reports, guidelines for interpretation and manuals. Judge their appropriateness for the intended application.
➢7. Review printed assessment methods and related materials for content or language generally perceived to be insensitive, offensive or misleading.	➢ 7. Review printed assessment methods and related materials for content or language that would offend or mislead students to be assessed.
➢ 8. Describe the specialized skills and training needed to administer an assessment method correctly and the specialized knowledge to make valid interpretations of scores.	➢ 8. Ensure all individuals who administer the assessment method, score the responses and interpret the results have the necessary knowledge and skills to perform these tasks (learning assistance teachers, speech and language pathologists, counselors, school psychologists, psychologists).
➢ 9. Limit sales of restricted assessment materials to persons who possess the necessary qualifications.	➢9. Ensure access to restricted assessment materials is limited to people with the necessary qualifications.
➢10. Provide for periodic review and revision of content and norms and, if applicable, passing or cut-off scores. Inform users.	➢10. Obtain information about the appropriateness of content, the recency of norms, and, if applicable, the appropriateness of cut-off scores for use with students to be assessed.
➢11. 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.	➢11. Obtain information about the comparability of interchangeable forms, including computer adaptations.
➢12. Provide evidence that an assessment method translated into a second language is valid for use with that language. This information should be provided in the second language.	➢12. Obtain evidence about the validity of the use of an assessment method translated into a second language.
➢13. Advertise an assessment method in a way that states it can be used only for the purposes for which it was intended.	➢13. Verify advertising claims made for an assessment method.

Collecting and Interpreting Assessment Information

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' Responsibilities	***Users' Responsibilities***
➢1. Provide clear instructions for administering the assessment method and identify the qualifications people who should administer the method should have.	➢1. Ensure the assessment method is administered by qualified personnel or under their supervision.
➢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 instruments translated from one language to another are administered by people proficient in the translated language.

➢ 3. Provide answer keys and describe procedures for scoring when scoring is to be done by the user.

➢ 4. Provide score reports or procedures for generating score reports that describe assessment results clearly and accurately. Identify and explain possible misinterpretations of scores yielded by 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 favorably 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.

➢ 3. Follow procedures for scoring as set out for the assessment method.

➢ 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. Take into account the effects of such factors as speed, test-taking strategies and attempts by students to present themselves favorably in their responses.

➢ 6. Take into account major differences between the norm group(s) or comparison group(s) and the students being assessed. Consider discrepancies between recommended and actual procedures and differences in familiarity with the assessment method between norm group(s) and 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 any computer administration and computer interpretations of assessment results are accurate and appropriate for the intended use. If necessary, ensure relevant information not included in computer reports is also considered.

➢ 9. Observe jurisdictional policies regarding storage of and subsequent access to the results. Ensure computer files are not accessible to unauthorized users.

➢10. Ensure all copyright and user agreements are observed.

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' Responsibilities

➢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 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 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 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' Responsibilities

➤1. Provide students or 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 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 students or their parents/guardians may follow to register concerns about the assessment and try to have problems resolved.

Implementing Mandated Assessment Programs[1]

Under some circumstances, administration of an assessment method is required by law. In such cases, the following guidelines should be added to the applicable guidelines outlined in the first three sections of Part B.

Developers' and Users' Responsibilities

➤1. Inform all persons with a stake in the assessment (administrators, teachers, students, parents/guardians) of the purpose of the assessment, how results will be used 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 purposes of the assessment, the intended uses of results and planned access to results.

➤ 6. Provide reports and explanations of results that can be readily understood by the intended audiences. If necessary, employ multiple reports designed for different audiences.

Note

1. 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 this section, 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.

References

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

Canadian Psychological Association. *Guidelines for Educational and Psychological Testing*. Ottawa: Author, 1986.

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

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 T6G 2G5.)*