

D.H. Saklofske
University of Saskatchewan

H.L. Janzen
University of Alberta

School-Based Assessment Research in Canada

Abstract

School-based assessment practices in Canada share many similarities with the United States. The place of testing in the larger assessment process is still seen as very important and school psychologists are frequently requested to administer intelligence, personality, and achievement measures to school children. This has resulted in using tests developed in other countries and, at times, modifying or standardizing them for use here or developing Canadian tests. Teacher-made tests are most preferred by teachers for the assessment of student achievement. Emerging trends in education will require some changes in the way assessments are conducted and an increased recognition of the importance of cognitive processes being assessed by educational personnel.

Résumé

Les méthodes d'évaluation en milieu scolaire au Canada ont de nombreux points en commun avec celles des États-Unis. L'importance des tests dans le processus d'évaluation global est toujours perçue comme essentielle et les psychologues scolaires sont fréquemment tenus d'administrer aux écoliers des tests d'intelligence, de personnalité et de niveau. Cela explique qu'on se serve de tests mis au point dans d'autres pays et qu'on les modifie ou qu'on les normalise parfois pour le Canada ou que l'on conçoive des tests résolument canadiens. Les tests conçus par des professeurs ont la faveur des enseignants pour le contrôle des progrès des élèves. Les orientations futures de l'éducation nécessiteront d'apporter certains changements à la conduite des évaluations et de faire reconnaître au personnel scolaire l'importance des processus cognitifs qu'ils évaluent.

The task of describing school-based assessment practices in Canada is a rather complex one because of geographical size, social and cultural diversity, and the fact that educational matters are the jurisdiction of the provinces. Teachers are most involved in the ongoing measurement and evaluation of student learning through the administration of teacher-made tests (criterion-referenced and curriculum-based) and norm-referenced standardized achievement tests. The more specialized diagnostic assessments requested by classroom teachers may be performed by various specially trained professionals such as school psychologists, guidance counsellors, itinerant teachers, school social workers, and speech therapists. Dumont (1987) observed that the job descriptions for school psychologists obtained from "the various ministries of education in Canada always include assessment and diagnosis" (p. 106).

The purposes and methods of educational and psychological assessment in Canadian schools share a great many similarities to practices in the United States. So far, Canada has remained relatively free from the kinds of landmark court cases heard in the United States regarding the assessment and school program placement of minority children (e.g., *Larry P. v. Riles*, 1979; *Marshall v. Georgia*, 1984). Both countries endorse the right of access to education for all children. However, while the United States has enacted some comprehensive pieces of legislation that are intended to ensure appropriate educational services to all children with special needs (e.g., PL 94-142, PL 99-457), Canada has no similar national legislation of this kind that would have an impact on diagnostic and program services. *The Canadian Charter of Rights and Freedoms* (1982) certainly has some potential in this regard (Dickinson & MacKay, 1989; Fox, 1988; Kimmins, Hunter, & Mackay, 1985). As well, Canada has signed both the Universal Declaration of Human Rights and the Declaration of the Rights of the Child which support the right to appropriate education. Ray (1989) concluded that "Canadians may assert with confidence that their education deals more effectively with many rights than was customary, and it receives high marks in international reviews. . . and domestic opinion" (p. 155).

Assessment Methods

Canadian educational personnel employ multiple assessment approaches to aid in describing and understanding children in order to provide the best and most appropriate educational services. The "four pillars of assessment" discussed by Sattler (1988), including informal assessments, interviews, observations, and norm-referenced tests, form the basis for collecting the kind and amount of information necessary for defining and evaluating the psychological and educational needs of children. The techniques and tools that comprise these four assessment categories are further cast within a frame-

work that recognizes the necessity of measuring and evaluating a whole complex range of human factors and external conditions. In keeping with the theme of this journal issue, the following sections will briefly overview both nontest and test-based methods employed in the assessment of intelligence, personality, and achievement as they are generally practised in Canada.

Nontest Assessment

School personnel have a wide choice of tests that may be employed in the assessment of childrens' needs yet much of the valuable data leading to hypothesis generation or decision making are based on nontest methods. Indeed, such methods as interviews, task analysis, work samples, and observations frequently comprise not only the initial stages of the assessment process but also are utilized throughout it. While classroom teachers are continuously engaged in ongoing evaluations of students, and therefore routinely employ all of these methods, school psychologists usually become involved with a student once a referral/request is made for their professional services. The initial data collection should most likely utilize interviews, observations, and record reviews that will shed further light on the reason for referral and provide the kinds of background information that will help guide the assessment process.

Information focusing on home behaviours, the family system, social and developmental history, classroom or instructional factors, and student perceptions of the "problem" are necessary in formulating hypotheses relevant to determining the nature of the problem and what further actions may be required. For example, Thompson (1985) reported that home environment factors (e.g., parental behaviours, home literacy, educational ambition, socio-economic efficiency) are important determiners and predictors of educational performance and that this effect increases with age. Cummins (1982) has further shown that the length of residency and age of arrival for children born outside Canada can lower scores on certain cognitive and language-based tests during the first five years of enrollment in an English-language curriculum.

Interviews are important in the construction of a model of the student which is the central purpose of assessment. Although much has been written about interviews, there is some indication that we are still concerned about whether we can trust what we see, hear, and feel. It is still most common to read test-based journal articles written on student assessment by school psychologists. While the assessment reports generated by school-based specialists may be heavily weighted by test scores, most will also contain information gleaned from either structured or unstructured interviews with

the student. Although different interview techniques may be selected, the need to "fit" the interview to the client is particularly highlighted in cross-cultural assessment, and Massey (1988) has outlined a useful set of guidelines for conducting such interviews.

While assessment is sometimes viewed as being synonymous with testing, it is much more than the administration of tests. During these days of fear of litigation, diagnosticians rely heavily on the safety of tests and test scores. Furthermore, school systems frequently request test data for purposes of designation and funding for special education. However, blindly interpreting such data without the full use of background, interview and self-report information, and current observations is not professionally defensible. Systematic behavioural observations are required to support and supplement obtained test data, perhaps even enhance it. Observational recordings may be used to describe both test and nontest behaviours, support the validity of test scores, aid in explaining test score variance, and add important information necessary to formulate a particular diagnosis and program recommendation.

Observation permits the opportunity to gather data of relevance to understanding and describing the child in school. Observation has an advantage in monitoring behaviours that change rapidly, exhibit particular growth or performance curves, and may not be typical or frequent. Of particular importance in the assessment process is the observation of children in their everyday environments such as the classroom. It is difficult to create tests that are both robust and sensitive enough to yield information about the frequency, intensity, duration, and situational specificity of a wide range of psychological constructs and human behaviours. For example, careful observation and recording procedures may aid in determining whether a child's poor performance in school is related to inability, lack of prerequisite skills, unwillingness, anxiety, or inappropriate instructional techniques.

Observation, like an interview, is both a technique and a skill. Knowing how to systematically observe is as important as knowing what to observe. Various observation schedules in the form of checklists and rating scales that focus on physical variables, affect, activity level, speech and language, adaptive behaviours, and social skills are commercially available but may also be created by teachers and psychologists as required. Two Canadian studies may be cited in this latter regard. Whyte (1984) developed a student self-report inventory that differentiated between learning-disabled and regular achieving adolescents on the basis of visual-spatial development, language and communication skills, memory and sequencing, social and emotional adjustment, and attentional skills. Whyte commented that the use of these techniques may permit the identification of learning-disabled stu-

dents and “make the more expensive and time consuming psychological testing more effective and efficient” (p. 24). At the other end of the educational ladder, Simner (1987) constructed a five-item questionnaire for identifying preschool children who may be at risk for early school failure. *The Teacher School Readiness Inventory* (TSRI) requires teachers to rate their observations of student in-class distractibility/attention, verbal fluency, interest and participation, letter identification, and printing skills. Results from Simner’s study suggest that the TSRI had high interrater reliability, was significantly related to performance across the curriculum, and had a “hit-rate” of 86% in identifying children who were experiencing academic problems in the first year of school.

Educational and Psychological Testing

The history and growth of psychological and educational test use in Canadian schools is fairly similar to the American experience. However, the disproportionate population size between the two countries has resulted in it sometimes being more economical to “import” tests developed and standardized in the United States and other countries for use in Canadian school systems rather than attempt to create them here. This situation can sometimes present rather major problems in the assessment process especially when norm-referenced tests are employed or the product being measured is tied to specific and unique experiences as in the case of curriculum-based assessment (Janzen & Saklofske, 1990). Thus, it is not always such a straight forward matter of simply using well constructed American tests in Canada. Since tests and measurements are such essential tools in assessment, some American instruments that are brought into Canada may be renormed or modified following the accumulation of data from research and clinical use. As an alternative, a number of tests have been created in Canada, especially in those areas that involve school achievement that is tied to local and provincial curriculums.

Assessment of Intelligence

The measurement of children’s intelligence is considered to be an important component in a comprehensive and multilevel assessment program. Despite the ongoing controversy surrounding the use of intelligence tests and Mercer’s (1988) contention that the intelligence testing debate is “evidence of growing intellectual crisis in measurement psychology that may be the prelude to a scientific revolution in that field” (p. 1), these are the most frequently administered of all individual tests requested of school psychologists. Many school boards also routinely administer group intelligence tests. IQ scores may be compared with achievement data to determine the extent of

student underachievement and learning difficulty, often with the aid of "mathematically-based formulae" (McLeod, 1988). Intelligence is viewed as an important variable in an analysis of school achievement and the findings from Canadian studies (e.g., Hardman & Oldridge, 1985; Janzen *et al.*, 1989 a,b) of the predictive validity of tests such as the *Weschler Intelligence Scale for Children-Revised* (WISC-R) and the *Stanford-Binet Intelligence Scale: Fourth Edition* (SBIS) are in line with American results.

Wilson and Humphries (1986) provide a not uncommon description, based on events in Ontario, of the demand for intellectual assessment as a major data source used for the designation and placement of students in special classes; "there is widespread dependence in school boards on norm-based instruments of which the most frequently used is the WISC-R" (p. 3). This widespread and frequent use of the WISC-R in assessment has generated a fairly large amount of research, in part, because of the controversy that Canadian children may perform differently from the American standardization sample and that some parts of the test may be biased against Canadian children. Since Vernon (1977) recommended the substitution of certain Canadian items on the Information subtest, a number of studies have been conducted to determine the efficacy of such changes (e.g., Cyr & Atkinson, 1987; Marx, 1984; Peters, 1976) culminating most recently in Beal's (1988) assertion that there is no solid evidence for changing the item content on the WISC-R. While national norms have not been developed for the WISC-R in Canada, some local or provincial norms do exist (e.g., Holmes, 1981). It is generally agreed that the American norms for this test are appropriate for use here although Canadian children, on average, score slightly above the subtest and IQ mean scale scores. This is partly due to the fact that the test was standardized 15 years ago. The WISC-R will soon be replaced by a new version which is concurrently being validated in Canada (Beal, 1989).

The WISC-R has also been employed in validity and clinical studies of exceptional children including retarded, learning disabled, and gifted students (see Janzen & Saklofske, 1990). Some of this research parallels that carried out in the United States; for example, WISC-R profile and pattern analysis related to achievement and learning disabilities (Bellemare, Inglis, & Lawson, 1986; Schmidt, Kuryliw, Saklofske, & Yackulic, 1989; Schmidt & Saklofske, 1983; Walsh, Marx, & Sudmant, 1983). A major concern relates to the use of this and other intelligence tests with culturally different groups of children, especially with Canadian aboriginal children. There is a divergence of opinion about the continued use of the WISC-R in the assessment of these children yet a fairly consistent pattern of average performance scores but below-average verbal scores has been reported by several researchers (Common & Frost, 1988). As well, a recent study of aboriginal children

living in villages and cities in British Columbia found that experiential background, not ethnicity, was a major determinant of information processing ability and the development of reading skills (Williams, 1988). These kinds of findings have resulted in school psychologists developing guidelines for the assessment of Canadian aboriginal children, including the examiner's responsibilities (Goldstein, 1988).

It appears that Canadian school psychologists and educators will continue to rely on individual intelligence tests developed and standardized in the United States. While intelligence is a universal construct, its measurement and phenotypic expression can vary as a function of culture, experience, language, and so forth unless one chooses to measure more basic biological processes such as reaction time or evoked potentials. This is clearly realized in a multicultural country like Canada where research investigations of new tests with the potential for widespread use are usually initiated fairly quickly following their introduction. Such is the case, for example, with the *Stanford-Binet Intelligence Scale-Fourth Edition* (Janzen *et al.*, 1989 a,b; Kline, 1989), *Kaufman Assessment Battery for Children* (Gardner, 1986, 1988; Saklofske & Jedlicki, 1985), *Matrix Analogies Test-Short Form* (Saklofske & Murray, 1989), and *Draw A Person: A Quantitative Scoring System* (Saklofske & Braun, 1989).

It should be noted that several intelligence tests have been specifically developed for use in Canada although these are mainly group tests that are most often administered by classroom teachers as part of a school district or even province-wide testing program. *The Canadian Cognitive Abilities Test* (CCAT) is one of the most often used group tests of intelligence and was recently standardized along with the *Canadian Test of Basic Skills* (CTBS). The CCAT covers the K-12 grade range and yields measures of verbal, quantitative, and nonverbal ability. While not aimed specifically for the Canadian context, the work of Dr. J.P. Das of the University of Alberta (see Das & Naglieri, 1990) in developing tests for the assessment of attention, processing, and planning abilities must certainly be recognized.

Finally, it is interesting to note a contemporary study (Paul & MacLeave, 1989) that examined the perceptions of educators about intelligence and intelligence testing. The study was conducted within the Gander Bonavista-Connaigre Roman Catholic school district. Caution is necessary in generalizing to other areas of Canada. The reported findings included: (1) intelligence testing and the use made of these data were viewed more positively by educators who were high versus low "hereditarians" and by educators with university training beyond a Bachelor's degree in contrast to those only with an undergraduate degree, (2) secondary school teachers compared with ele-

mentary teachers supported the increased use of test scores, and (3) educators across all positions surveyed "appeared to be satisfied with current testing practices" (p. 380). These results and also the knowledge that intelligence tests are so often used in the psychoeducational assessment of children suggest that their place is assured for the foreseeable future.

Personality Assessment

The assessment of school children by psychologists and counsellors may include various personality measures depending on the nature of the referral and the decisions to be made. The personality characteristics of relevance to school personnel in understanding and providing for a child's personal and educational needs are the same as those of general interest in child and adolescent clinical psychology. Thus, as part of a diagnostic assessment, school psychologists may administer one or more comprehensive objective personality questionnaires (e.g., *Children's Personality Questionnaire*; *Junior Eysenck Personality Questionnaire*), projective techniques (e.g., *Children's Apperception Test*, *Kinetic School Drawings*, *Kahn Test of Symbol Arrangement*) or briefer instruments that measure very specific personality variables (e.g., *Children's Depression Scale*, *Nowicki-Stickland Locus of Control Scale for Children*).

A glimpse of Canadian educational and psychological journals (e.g., *The Alberta Journal of Educational Research*, *Canadian Journal of Behavioural Science*, *Canadian Journal of School Psychology*, *Canadian Journal of Special Education*) gives an indication of the investigations undertaken to examine personality factors in school children and how they may be related to school performance. Byrne's (1986) investigation of the self-concept of high school students replicated previous findings regarding the multidimensional, hierarchical structure and stability of self-concept as well as its relationship with academic achievement. Studies of learning disabled (LD) children have shown that they can be distinguished from non-LD students through an analysis of the graphic elements contained in the *Kinetic School Drawing Test* (Andrews & Janzen, 1988). Rogers and Saklofske (1985) found that LD students had lower self-concepts, a more external locus of control, and lower performance expectations than normally achieving elementary school children. Furthermore, measures of general and academic locus of control and academic self-concept contributed significantly in predicting the extent to which LD children were rated by teachers as experiencing academic success in their special education programs.

Turning to studies of attitudes, Nyberg and Clarke (1982) reported that the responses of grade 5 and 8 students on the *School Subjects Attitude Scales*

were related to teachers' perceptions of student attitudes, student choice of most and least liked subject, type of school program (e.g., academic), and cultural and sex differences. Following the suggestion that attitudinal and personality variables are important in the reading process, Summers (1980) has provided evidence for the validity of the *Estes Reading Attitude Scale* as a measure of school-oriented attitude toward reading in the intermediate grades. Furthermore, attitudes toward reading were not changed by the opportunity for sustained silent reading (Summers & McLelland, 1982).

In another area, self ratings of social behaviour by secondary school students were observed to be related to school achievement and gender (Loranger, Poirier, & Gauthier, 1983). As a final example, Rampaul, Singh, and Didyk (1984) reported significant positive correlations between self-concept, academic achievement, and teacher expectations among grade 3 and 4 Canadian aboriginal students in northern Manitoba. However, the authors observed "academic retardation with increasing age" among this sample of students (p. 213). These and other studies of personality in school children provide the school psychologist with a research base that will enhance the assessment process. Few school psychologists routinely administer selected personality tests to all referred children but rather will determine what to measure and what test(s) to employ based on the individual child's presenting needs. This may be suggested by the teacher's referral of a child for assessment; the child "appears nervous and anxious and doesn't participate in class," "has a low self-concept, difficult to motivate," or "has problems staying on task, is somewhat aggressive." Of course, school psychologists will continuously be "on the lookout" for personality factors that may support or alter their hypotheses about a child and guide data collection and diagnostic interpretation. Massey (1988, p. 29) and other Canadian researchers (e.g., Berry, 1975) have stated that attitudinal, motivational, and "other personality characteristics such as the amount of eye-contact, reaction to physical proximity, male-female interaction and lack of spontaneous speech also affect the interpretation of results" in the assessment of culturally different children.

It is very rare that personality-type questionnaires would be routinely administered to large numbers of children as part of a screening assessment program, as occurs with group intelligence and achievement tests, unless they are part of a more research-oriented project. An example of such a situation may be found in two studies of depression in Canadian elementary school children conducted by Paananen and Janzen (1986) and Saklofske, Janzen, and Paananen (1987). In order to determine the incidence and characteristics of depression among grade 4-6 children, over 600 children completed measures of self-esteem, locus of control, and depression; achievement and ability data were obtained from cumulative records. These results suggested, among

other things, that less than 10% of these children suffer from a depressive syndrome, depressed children have lower self-esteem and an external locus of control, and that the *Children's Depression Inventory* and *Children's Depression Scale* provided a similar index of depression.

The greatest number of personality measures used in Canada are most often American in origin and range from comprehensive, multifactor standardized tests to brief scales published in various journals. In recent years, there has been some encouragement by test developers to simultaneously standardize new instruments in both Canada and the United States. For example, a new battery called the *Behaviour Assessment System for Children* (BASC), authored by Kamphaus and Reynolds and to be published in 1990 by the American Guidance Service, is intended to provide a psychometrically sound approach to assessing the emotional and behavioural problems of children and adolescents. The authors advertised for Canadian sites so that appropriate norms may be developed. In this regard, Saklofske, Yackulic, and Kowalchuk from the University of Saskatchewan have recently completed the data collection phase for a provincial standardization study of the BASC with school children in grades 2-8 and a second study of conduct disordered boys with the expectation that this multisource measure (child self-report personality profile, parent rating scale, teacher rating scale, parent personality profile, classroom observation scale, developmental questionnaire) will prove especially useful in the assessment process.

In the event that personality tests from other countries are employed in the assessment of Canadian school children, it may be necessary to renorm or even modify the instrument. The *Junior Eysenck Personality Questionnaire*, developed in England, may be administered to Canadian children without scoring changes in the scales but the Canadian data show higher mean scale scores for the neuroticism and tough-mindedness scales, lower scores on social desirability, and almost no difference in extraversion (Eysenck & Saklofske, 1983). In contrast, the *Junior Impulsiveness Inventory*, also from England, did require some item changes when introduced in Canada (Saklofske & Eysenck, 1983).

Finally, it is important to note that various children's personality measures have been developed and standardized in Canada. Alberta has been the venue of some of this work and includes: the *Frost Self-Description Questionnaire: Extended Scales* (Frost, 1979) which measures various forms of anxiety and aggression; the *Canadian Self-Esteem Inventory for Children* (Battle, 1976, 1979); the *Culture-Free Self-Esteem Inventory for Children* (Battle, 1981); and the *Student's Perception of Ability Scale* (Boersma & Chapman, 1985) which is now being revised. This latter instrument follows

from research that has shown a strong relationship between self-perceptions of ability and school achievement.

Assessment of Achievement

The assessment of student learning and achievement can be initiated at different levels (classroom quiz to board and province-wide testing programs) and in various different ways (teacher-made tests of specific instructional content, standardized group achievement tests, diagnostic tests, homework assignments, project reports, etc.). Educational specialists are less involved than other staff in the everyday assessment of school achievement, but they are consulted by the teacher when additional information and possible program changes may be required for a student.

In this area of assessment, tests are much less portable from country to country and even across regions within a country like Canada. School performance and the learning outcomes of main importance for a given grade or subject area are most specific to the defined curriculum and objectives of school districts and provinces and may further vary to some extent across teachers and schools. Thus, very few of the comprehensive standardized achievement test batteries used in American schools are employed in Canada. Some briefer instruments and diagnostic tests prepared outside of Canada are useful as screening devices or for providing an analysis of particular learning problems. The *Wide Range Achievement Test* (WRAT) and the most recently revised version are sometimes used to obtain a general indication of reading, arithmetic, and spelling ability. The WRAT has not been normed in Canada and there is some difference of opinion as to the utility of this test for use here (Siegal, 1984; Snart, Dennis, & Brailsford, 1983). Another example of this kind is reflected in efforts to develop guidelines for the effective use of non-Canadian developed instruments such as the *Boder Test of Reading-Spelling* for screening decoding and encoding problems in French immersion students (Wiss & Burnett, 1988).

Several carefully standardized individual and group achievement tests have been specifically constructed for use in Canadian schools. The *Canadian Achievement Tests* (CAT) and the *Canadian Tests of Basic Skills* (CTBS) are two instruments frequently used by school systems as a broad achievement measure across all school grades. For example, the CTBS measures reading, language, work-study, and mathematics skills. Both tests were normed on large numbers of Canadian school students which permits the comparison of obtained test scores with local, provincial, and national results. The *British Columbia Quick Individual Educational Test*, yielding measures of spelling, word identification, passage comprehension, and arithmetic, is based on the

British Columbia curriculum and was normed in that province (Wormeli, 1983). McDonald (1985) concluded that this test is effective in the identification of students with learning difficulties. From a different perspective, the *Canadian Readiness Test* is a battery of seven subscales measuring four general areas of prereading behaviour and has been found to have predictive validity for reading and mathematics achievement (Collis, Ollila, & Yore, 1986; Evanechko, Ollila, Downing, & Braun, 1973). As well, a large number of achievement tests are constructed by provincial education departments and local school boards as a means of assessing learning outcomes based upon specific objectives.

Classroom teachers base their evaluations of student learning less often on standardized testing programs alone, but rather employ a variety of assessment methods and procedures. Teacher-made tests appear to be the most often used source of achievement data for student evaluation. Anderson (1989) reported that teacher-made tests were viewed by British Columbia science teachers as the most important information source with the least emphasis being given to standardized objective tests and other techniques such as oral tests, work contracts, and student self-reports. In an earlier study, Wahlstrom and Danley (1976) reported that Ontario elementary teachers prefer classroom observations of student performance but secondary teachers favour tests.

While classroom teachers are more inclined to use their own tests to measure student achievement, a number of Canadian educational researchers have expressed concern about the "gap" between test theory and research in the measurement and evaluation of student performance on the one hand, and the evaluation practices of teachers on the other (McLean, 1985). Wilson (1989) concluded that "it is still unclear whether teachers view evaluation activities in a coherent, holistic manner" (p. 143) and suggested two possible reasons for this situation:

The ready-made instruments may perform one part of the teacher's purpose for evaluation but not another. . . similarly, the professional's concern with reliability of scoring, item analysis, and systematic content converge may also seem of only esoteric interest, and even an imposition, to teachers practising the delicate balancing acts required by the different, sometimes contrary, forces operating on them already. (p. 143)

This separation of measurement theory and research from actual classroom evaluation practices has led to the recommendation that "applied edu-

cational measurement should center its attention on the classroom" (Anderson, 1989, p. 133). Evidence of a positive response to this plea is now being seen across Canada. To cite only a few examples of this new direction, McLean (1988) has described how functional language theories, which view achievement as performance, and as task and situation dependent, can provide the basis for effective teaching and achievement measurement. McLean has extended this "functional approach" to the areas of mathematics and science in the middle school. A recent issue of *Research Forum* (1989, No. 4), titled "Evaluation: Facing the Challenge", contained a number of articles that describe and examine the different kinds of measurement and evaluation practices employed by teachers including curriculum-based, criterion-referenced, naturalistic, and multidimensional evaluation procedures.

An emerging trend involves the identification of deficient cognitive functions using techniques such as Feuerstein's *Learning Potential Device* (Watts, 1985) and the analysis of both process and product in student learning, such as in spelling (Evans & Smith, 1989). The movement toward cognitive education programs will certainly require further modifications to the way teachers and psychologists view and measure student learning.

A most exciting four-year research project evaluating Feuerstein's *Instrumental Enrichment* (IE) and Mulcahy's *Strategies Program for Effective Learning and Thinking* (SPELT) has just been completed at the University of Alberta (Mulcahy, 1989). While the first draft of this report to the Alberta government is not yet available for circulation at this time, the author's have been given permission to cite this work. In brief, a sample of approximately 900 students experienced one of the IE, SPELT, or conventional instruction programs. Gifted, learning disabled, and normally achieving students in grades 4 and 7 were followed through to the completion of grades 6 and 9, respectively, permitting the comparison of a variety of achievement indices and cognitive strategies gained from the out-of-content versus in-content approaches. Generally, SPELT instruction was superior to IE and regular instructional methods, most notably in reading comprehension and related cognitive strategies. The learned cognitive strategies were maintained over time, together with improvements in achievement. The efficacy of the SPELT program was further observed in student gains in self-concept, locus of control, problem-solving strategies, metacognitive reading awareness, and comprehension monitoring skills. Learned reading and comprehension skills such as skimming, rereading, paraphrasing, inferring, and checking were successfully maintained. This study certainly provides evidence for the potential of such cognitive education programs but also points to the need for new and different assessment strategies.

Conclusions

There are many similarities in the school-based assessment practices of Canadian and American teachers, school psychologists, and other educational personnel. Although some school psychologists are still employed to administer tests (mainly individual intelligence tests), this narrow role appears to be on the verge of an even more widespread decline. Assessment is more than the administration of tests; it requires input from the various sources representing the "four pillars of assessment" (Sattler, 1988) but must be further viewed as a process including referral, screening, classification, planning, and program modification (Salvia & Ysseldyke, 1985).

While classroom teachers are continuously assessing student learning, the kinds of assessment information obtained by educational specialists frequently extends beyond data from a single test and is not limited to only a measure of, say, intelligence if it is to have utility in understanding and providing for a child's needs. There is also the recognition that the assessment must include information about the child and the child's situation or environment. This position is especially relevant in a multicultural mosaic such as Canada. While there is no equivalent to Mercer's (1979) *System of Multicultural Pluralistic Assessment* in Canada, recent descriptions of cross-cultural assessment have provided important guidelines for data collection, interpretation, and diagnosis (Massey, 1988; see also Samuda, Kong, Cummins, Pascual-Leone, & Lewis, 1989).

Many different tests may be employed to measure children's intelligence, personality, and school achievement. The need to examine the psychometric characteristics of instruments developed outside of Canada is recognized as is the importance of Canadian standardization of these norm-referenced tests. Achievement tests, such as the CTBS and curriculum-based achievement tests developed at the local and provincial levels, provide useful information but the fact that teachers show a preference for constructing their own measures needs to be addressed. The teacher requires assessment data that have instructional relevance, and educational measurement and evaluation specialists can certainly support the teacher's in-class assessment roles. Finally, emerging trends in cognitive education will require the development of new assessment techniques that will more effectively describe learning processes, strategies, and outcomes.

REFERENCES

- Anderson, J.O. (1989). Evaluation of student achievement: Teacher practices and educational measurement. *The Alberta Journal of Educational Research*, 35(2), 123-133.
- Andrews, J., & Janzen, H. (1988). Objectifying the Kinetic School Drawing (KSD): A comparison of LD and non-LD grade 5 students. *Canadian Journal of School Psychology*, 3, 27-51.
- Battle, J. (1976). Test-retest reliability of the Canadian Self-Esteem Inventory for Children. *Psychological Reports*, 38, 1343-1345.
- Battle, J. (1979). Self-esteem of students in regular and special classes. *Psychological Reports*, 44, 212-214.
- Battle, J. (1981). *Culture-free self-esteem inventory for children and adults*. Seattle: Special Child Publications.
- Beal, A. Lynne. (1988). Canadian content in the WISC-R: Bias or jingoism. *Canadian Journal of Behavioural Science*, 20(2), 154-166.
- Beal, A. Lynne. (1989). *Wechsler Intelligence Scale of Children III: A study to determine its validity for Canadian children*. A paper presented at the Canadian Association of School Psychologists' Conference, Edmonton, Alberta.
- Bellemare, F.G., Inglis, J., & Lawson, J.S. (1986). Learning disability indices derived from a principle component analysis of the WISC-R: A study of learning disability and normal boys. *Canadian Journal of Behavioural Sciences*, 18, 86-91.
- Berry, J.W. (1975). Psychological testing across cultures. In A. Wolfgang (Ed.) *Education of immigrant students*. Toronto: Ontario Institute for Studies in Education.
- Boersma, F.J., & Chapman, J.W. (1985). *Manual for the Student's Perception of Ability Scale*. Unpublished manuscript. University of Alberta, Edmonton, Alberta.
- Byrne, B.M. (1986). Self-concept/academic achievement relations: An investigation of dimensionality, stability, and causality. *Canadian Journal of Behavioural Science*, 18(2), 173-186.
- Collis, B., Ollila, L.O., & Yore, L.D. (1986). Predictive validity of the Canadian Readiness Test and selected measures of cognitive development for grade one reading and mathematics achievement. *The Alberta Journal of Educational Research*, 32(1), 2-11.
- Common, R.W., & Frost, L.G. (1988). The implications of the mismeasurement of Native students' intelligence through use of standardized intelligence tests. *Canadian Journal of Native Education*, 15, 9-30.
- Cummins, J. (1982). Tests, achievement, and bilingual students. *Focus, National Clearing House for Bilingual Education*, 1-7.
- Das, J.P., & Naglieri, D.H. (1990). Assessment of attention, simultaneous-successive coding and planning. In H.C. Haywood & D. Tsurie (Eds.), *Interactive assessment*. New York: Springer-Verlag.
- Cyr, J.J., & Atkinson, L. (1987). Test item bias in the WISC-R. *Canadian Journal of Behavioural Science*, 16, 30-35.

- Dickinson, G.M., & McKay, A.W. (1989). *Rights, freedoms and the education system in Canada*. Toronto: Edmond Montgomery.
- Dumont, F. (1987). School psychology in Canada: An evolving profession. *McGill Journal of Education*, 22(2), 101-116.
- Evanechko, P., Ollila, L., Downing, J., & Braun, C. (1973). An investigation of the reading readiness domain. *Research in the Teaching of English*, 7(1), 61-78.
- Evans, P., & Smith, A.L. (1989). Spelling errors: grade 1 to university-analysis of process and product. *The Alberta Journal of Educational Research*, 35(1), 20-33.
- Eysenck, S.B.G., & Saklofske, D.H. (1983). A comparison of responses of Canadian and English children on the Junior Eysenck Personality Questionnaire. *Canadian Journal of Behavioural Science*, 15(2), 121-130.
- Fox, Tricia A. (1988). Prisoner and students rights in regard to education: Are they implemented in Canada? *McGill Journal of Education*, 23(3), 265-276.
- Frost, B.P. (1979). *The Frost Self Description Questionnaire: Extended Scales*. Department of Educational Psychology, University of Calgary, Calgary, Alberta.
- Gardner, J. (1986). A comparison of the K-ABC and WISC-R for Cantonese, English and Punjabi speaking Canadians. *Canadian Journal of School Psychology*, 2, 45-54.
- Gardner, J. (1988). Factor structure of the K-ABC for Cantonese, English and Punjabi speaking Canadians. *Canadian Journal of School Psychology*, 4, 39-59.
- Goldstein, C.R. (1988). Cautions concerning assessment of Native Indian students. *Canadian Journal of School Psychology*, 4, 17-20.
- Hardman, S., & Oldridge, O. (1985). Predictive validity of WISC-R on the Woodcock-Johnson Achievement Battery, the B.C. Quick Individual Education Test and teacher rankings of student underachievement. *Canadian Journal of School Psychology*, 1, 31-38.
- Holmes, B. (1981). *Individually administered intelligence tests: An application of anchor test norming and equating procedures in British Columbia*. Unpublished doctoral dissertation. University of British Columbia.
- Janzen, H., Andrews, J., Fritzke, B., Clark, L., & Boersma, F. (1989a). *Relationship between the Stanford-Binet, 4th Edition and the WISC-R: A concurrent validity study*. Unpublished manuscript, Department of Educational Psychology, University of Alberta.
- Janzen, H., Andrews, J., Fritzke, B., Clark, L., & Boersma, F. (1989b). *The discriminant validity of the Stanford-Binet, 4th Edition*. Unpublished manuscript. Department of Educational Psychology, University of Alberta.
- Janzen, H.L., & Saklofske, D.H. (1990). Testing for diagnostic purposes. In L. Stewin, R. Short, & S. McCann (Eds.), *Psychology for the classroom*. Toronto: Copp Clark Pitman, (in press).
- Kamphaus, R.W., & Reynolds, C.R. (1990). *Behavior Assessment System for Children*. Circle Pines, MN: American Guidance Service, (in press).
- Kimmins, R., Hunter, W., & McKay, A.W. (1985). Educational legislation and litigation pertaining to the practice of school psychology in Canada. *Canadian Journal of School Psychology*, 1(1), 1-10.
- Kline, R.B. (1989). *The new Stanford-Binet is not a four factor test*. Paper presented at the 97th Annual Convention of the American Psychological Association, New Orleans, Louisiana.

- Loranger, M., Poirier, M., & Gauthier, D. (1983). Questionnaire d'auto-évaluation des conduites soules a l'école. *Canadian Journal of Behavioural Science, 15*(3), 280-300.
- Massey, Don S. (1988). Cross cultural psychological assessment: A clinical model of service delivery. *Canadian Journal of School Psychology, 4*, 21-37.
- Marx, R.W. (1984). Canadian content and the WISC-R Information Subtest. *Canadian Journal of Behavioural Science, 16*, 30-35.
- McDonald, D.D. (1985). Evaluation of the B.C. Quick Individual Educational Test. *Canadian Journal of School Psychology, 1*, 57-60.
- McLean, L.D. (1985). *The craft of student evaluation in Canada*. Toronto: Canadian Education Association.
- McLean, L.D. (1988). Achievement measures made relevant to pedagogy. *McGill Journal of Education, 23*(3), 243-252.
- McLeod, J. (1988). Individual differences from the perspective of computer-guided psycho-educational diagnosis. In D.H. Saklofske, & S.B.G. Eysenck (Eds.), *Individual differences in children and adolescents* (pp. 10-26). London: Hodder and Stoughton.
- Mercer, J.R. (1979). *System of Multicultural Pluralistic Assessment technical manual*. New York: Psychological Corporation.
- Mercer, J.R. (1988). Scientific revolutions and the psychomedical paradigm: The growing crisis in measurement psychology. *Canadian Journal of School Psychology, 4*, 1-16.
- Mulcahy, R.F. (1989). *Alberta Cognitive Education Program*. Department of Educational Psychology, University of Alberta, Edmonton.
- Nyberg, V.R., & Clarke, S.C.T. (1982). School Subjects Attitudes Scales. *The Alberta Journal of Educational Research, 28*(2), 175-287.
- Paananen, N., & Janzen, H. (1986). Incidence and characteristics of depression in elementary school children. *Canadian Journal of School Psychology, 2*(1), 7-19.
- Paul, J.F., & MacLeave, A.M. (1989). Educators' perceptions of intelligence and intelligence testing. *Canadian Psychology, 30*(2a), 380.
- Peters, H.D. (1976). The validity of the Wechsler Intelligence Scale for Children-Revised. *Canadian Journal of Behavioural Science, 8*, 414-417.
- Rampaul, W.E., Singh, M., & Didyk, J. (1984). The relationship between academic achievement, self-concept, creativity, and teacher expectations among Native children. *The Alberta Journal of Educational Research, 30*(3), 213-225.
- Ray, D. (1989). Human rights: Is Canadian schooling a privilege, a right or an obligation? *McGill Journal of Education, 23*(2), 143-155.
- Rogers, H., & Saklofske, D.H. (1985). Self-concepts, locus of control and performance expectations of learning disabled children. *Journal of Learning Disabilities, 18*(50), 273-278.
- Saklofske, D.H., & Braun, S. (1989). *A psychometric investigation of the Draw-A-Person: A Quantitative Scoring System utilizing a Canadian sample*. Paper presented at the Canadian Association of School Psychologists' Conference, Edmonton, Alberta.

- Saklofske, D.H., & Eysenck, S.B.G. (1983). Impulsiveness and adventuresomeness in Canadian children. *Psychological Reports, 52*, 147-152.
- Saklofske, D.H., Janzen, H.L., & Paananen, N. (1987). Predictors of childhood depression: Literature review and empirical findings. *Canadian Journal of Special Education, 3*(1), 1-14.
- Saklofske, D.H., & Jedlicki, G. (1985). *Concurrent and content validity of the K-ABC*. Paper presented at the 17th Annual Convention of the National Association of School Psychologists, Las Vegas.
- Saklofske, D.H., & Murray, W. (1989). *A normative and validity study of the Matrix Analogies Test-Short Form with Canadian children*. Paper presented at the Canadian Association of School Psychologists' Conference, Edmonton, Alberta.
- Salvia, J., & Ysseldyke, J.E. (1985). *Assessment in special and remedial education* (3rd ed.). Boston: Houghton Mifflin.
- Samuda, R.J., Kong, S.L., Cummins, J., Pascual-Leone, J., & Lewis, J. (1989). *Assessment and placement of minority students*. Toronto: Hogrefe and Huber.
- Sattler, J.M. (1988). *Assessment of children* (3rd ed.). San Diego: Jerome M. Sattler.
- Schmidt, H.P., Kuryliw, A.J., Saklofske, D.H., & Yackulic, R.A. (1989). Stability of WISC-R scores for a sample of learning disabled children. *Psychological Reports, 64*, 195-201.
- Schmidt, H.P., & Saklofske, D.H. (1983). Comparison of WISC-R patterns of children of average and exceptional ability. *Psychological Reports, 52*, 539-544.
- Siegal, L.S. (1984). On the adequacy of the WRAT: A reply to Snart, Dennis and Brailsford. *Canadian Psychology, 24*, 75-76.
- Simner, M. (1987). Predictive validity of the Teacher's School Readiness Inventory. *Canadian Journal of School Psychology, 3*(1), 21-32.
- Snart, F., Dennis, S., & Brailsford, A. (1983). Concerns regarding the Wide Range Achievement Test. *Canadian Psychology, 3*, 21-32.
- Summers, E.G. (1980). Validity of the Estes Reading Attitude Scale for intermediate grades. *The Alberta Journal of Educational Research, 26*(1), 36-43.
- Summers, E.G., & McLelland, J.V. (1982). A field-based evaluation of sustained silent reading in intermediate grades. *The Alberta Journal of Educational Research, 28*(2), 100-112.
- Thompson, W.W. (1985). Environmental effects on educational performance. *The Alberta Journal of Educational Research, 31*(1), 11-25.
- Vernon, P.E. (1977). Final report on modifications of WISC-R for Canadian use. *Canadian Psychological Association Bulletin, 5*, 5-7.
- Wahlstrom, M., & Danley, R.R. (1976). *Assessment of student achievement*. Toronto: Ontario Ministry of Education.
- Walsh, J., Marx, R.W., & Sudmant, W. (1983). WISC-R factor score profiles and academic achievement: A test of Kaufman's profile Analysis procedures. *Canadian Journal of Behavioural Science, 15*(3), 203-210.
- Watts, W.J. (1985). An error analysis on the Raven's using Feuerstein's Deficient Cognitive Functions. *The Alberta Journal of Educational Research, 31*(1), 41-53.
- Whyte, L.A. (1984). Characteristics of learning disabilities persisting into adolescence. *The Alberta Journal of Educational Research, 20*(1), 14-25.

- Williams, D.G. (1988). Developmental trends in information processing and reading: A cross-cultural comparison. *Canadian Journal of School Psychology, 4*, 39-47.
- Wilson, R.J. (1989). Evaluating student achievement in an Ontario high school. *The Alberta Journal of Educational Research, 35*(2), 134-144.
- Wilson, R.J., & Humphries, T. (1986). The impact of recent legislation on the role of Canadian school psychologists. *Canadian Journal of School Psychology, 2*(1), 1-6.
- Wiss, C.A., & Burnett, W. (1988). The application of the Boder Test of Reading-Spelling Patterns to the assessment of learning disabilities in French immersion poor readers. *Canadian Journal of School Psychology, 4*, 61-70.
- Wormeli, C.T. Jr. (1983). *Manual for the B.C. QUIET*. University of British Columbia Education Clinic. Vancouver, BC.

Donald H. Saklofske received his Ph.D. from the University of Calgary and is currently a professor in the Educational Psychology Department at the University of Saskatchewan. He has published papers on intelligence, personality and assessment and recently co-edited a book on individual differences.

Henry L. Janzen received his Ph.D. from the University of Calgary. He is a professor in the Department of Educational Psychology at the University of Alberta, Director of the Psychological Testing Center at the Faculty of Education, Clinical Services, and a consulting school psychologist. Dr. Janzen is active in many psychological associations, has published extensively, and is co-host of the radio program *That's Living* at CJCA 930.

Donald H. Saklofske a obtenu son doctorat à l'Université de Calgary et il enseigne actuellement au département de psychologie éducationnelle à l'Université de Saskatchewan. Il a publié des articles sur l'intelligence, le caractère et l'évaluation et a récemment coédité un livre sur les différences individuelles.

Henry L. Janzen a obtenu son Ph.D. à l'Université de Calgary. Il est professeur au département de psychologie éducationnelle de l'Université d'Alberta, directeur du centre de tests psychologiques des services cliniques de la faculté des sciences de l'éducation et psychologue scolaire-conseil. Il est membre de nombreuses associations de psychologie, a publié de nombreux ouvrages et est le coanimateur de l'émission *That's Living* diffusée sur les ondes de CJCA 930.

