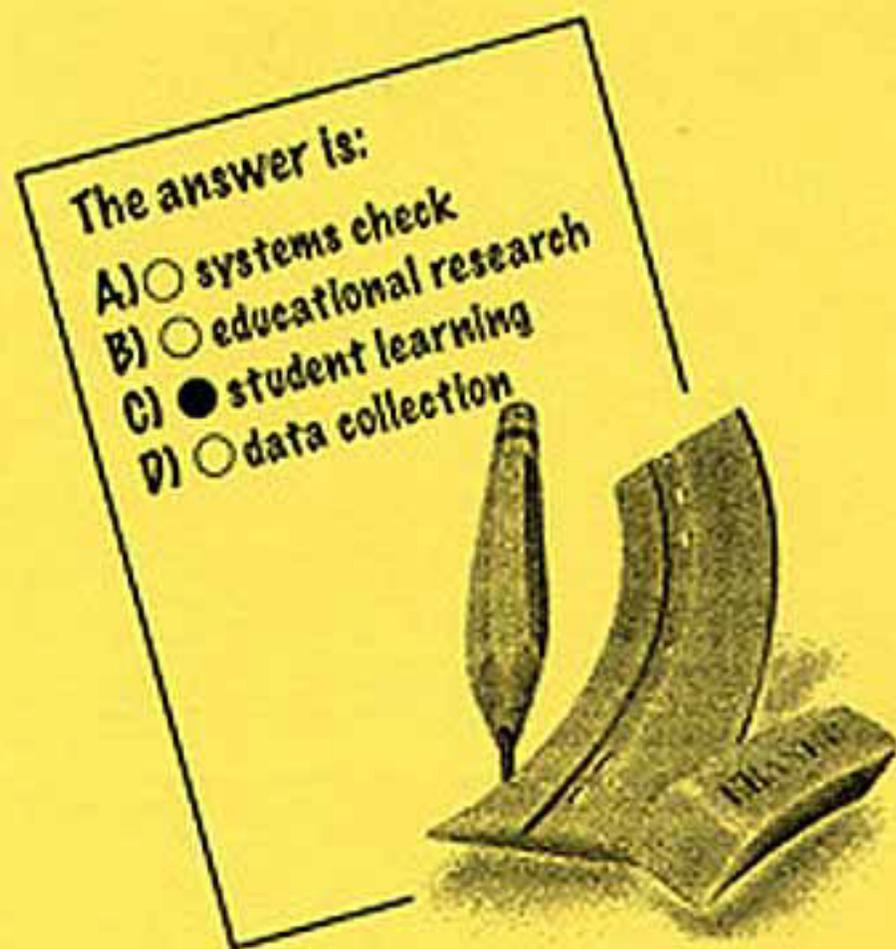


NBTA

Position Paper

on

External Assessment



May 2007

Position Paper on External Assessments

After the discontinuation of high stakes testing of New Brunswick children (such as the matriculation exams that were last used in the early 1970's), many educators took little notice of external assessments. For the most part, in New Brunswick as in other jurisdictions, external testing was perceived to have little to do with learning, and therefore had a minimal impact on the education system.

Recently, however, many educational policy makers have developed a belief that high external test scores reflect high quality education. Therefore standardized, external testing within the province is again high-stakes; the difference being that the current reincarnation is one of high-stakes testing of the public education system, rather than high-stakes testing of individual children. (The exception is the successful completion of the Grade 9 Provincial English Language Proficiency Assessment that is required for graduation.)

In the United States, No Child Left Behind legislation has placed such intense pressure on schools by means of Adequate Yearly Progress goals that headlines have been made by school closings, staff firings, manipulated test scores, and the abandonment of federal funding for school and/or district budgets. Rather than blindly being influenced by such a disastrous model, it is time for New Brunswick educators at all levels carefully to consider external assessments.

The New Brunswick Teachers' Association believes that classroom teachers have primary responsibility for assessing and evaluating student achievement. However, given the importance that policy makers are placing on external assessments, the New Brunswick Teachers' Association concedes that external assessments will be part of the provincial education system for the foreseeable future. If the province intends to continue to use these external tests, then the tests must support, rather than detract from, a child's education. Provided all of the following conditions are met, and student learning is enhanced, the NBTA will not oppose the judicious use of external tests by the New Brunswick Department of Education.

1. External assessment should only be undertaken if it leads to improvements in student learning.

The preeminent purpose of student assessment is to support student learning; if learning is not enhanced as a result of the activity, the activity should not be undertaken. It is extremely inappropriate to use the children of New Brunswick as unsuspecting research subjects. Regardless of the perceived value of any data collected from external tests, such tests should not be administered unless the results can be used to inform instruction.

Unfortunately, many non-educators believe that external assessments monitor the quality of instruction. Thus, the perception is that provincial or district tests may be used as a “systems check”. However, using the data compiled from external assessments solely as a “systems check” is unacceptable. The NBTA asserts that writing any test is a stressful situation for a child. Writing an assessment that has been designed by an unknown entity, in a format that is unfamiliar, in some circumstances without designated supports and on material that is unfamiliar to a child, can traumatize a child.

As educators, we know there is a vast difference between improving student learning and improving student test scores. In the absence of individualized, constructive feedback of results, and in the face of mounting pressure to achieve higher “average” test scores, schools can employ strategies that will improve test scores without improving meaningful learning. Teaching to the test is one such example.

Testing does not make kids smarter; teaching does. If the results of an appropriately-designed external assessment are compiled and accurately reported in a timely manner, if they are presented in such a way that teachers are able to identify student academic strengths and weaknesses, and if appropriate supports are offered to ensure interventions for struggling students, then external testing may aid in improving in student learning.

2. External assessments must be instructionally sensitive.

According to James Popham, “ An instructionally sensitive accountability test (1) measures only a reasonable number of genuinely significant curricular aims; (2) provides clear, teacher-palatable descriptions of what will be assessed; and (3) reports students’ results in a fashion that allows teachers to discern which of their instructional activities are effective or ineffective.” (Popham, 2006, p 356)

Popham asserts the “root cause of instructional insensitivity is almost without exception because of the staggering number of curricular aims ...” (Popham, 2006, p 355). The large number of curriculum outcomes is certainly a problem in the New Brunswick education system, particularly in the areas of mathematics and science. Curricular expectations in some subject areas demand that students master literally hundreds of outcomes. Although the outcomes may represent what may be theoretically achieved under ideal classroom conditions, the expectation does not reflect the reality of New Brunswick classrooms.

The provincial testing philosophy is incompatible with the provincial teaching philosophy. New Brunswick teachers are expected to differentiate instruction and assessment; they are expected to treat each child as an individual and to tailor instruction to meet the needs of the unique individuals that make up each unique class. This requires vast amounts of time – not only teachers’ non-instructional time outside the classroom, but also instructional time within the classroom. While knowledge may be transmitted quickly, understanding takes time.

Some teachers have been placed in an untenable position because of the dichotomy between teaching and testing philosophies. Given their classroom compositions, teachers can either work to ensure meaningful student learning occurs, knowing that only a portion of the assigned outcomes may be mastered, and that external test results will reflect the lack of coverage, or

they can superficially cover all outcomes, knowing that the external test results may reflect the students' lack of understanding. In some cases, this dilemma is resolved by reassigning instructional time from subjects that are not being tested provincially. This would serve as another example of how test scores might be improved without a concomitant improvement in learning.

If provincial assessments are to be administered, it is imperative that only a reasonable number of genuinely significant curricular outcomes form the basis of these assessments. The New Brunswick Teachers' Association would further postulate much of the learning that would be achieved in meeting genuinely significant curricular outcomes would be broad, and contextually embedded. Conventional accountability testing best measures discrete bits of knowledge, not broad understandings. If curriculum outcomes are designed to take into account learning styles, academic growth, and relevance, then the assessment of these outcomes should also take these factors into account.

Another factor that is integral to the design of an instructionally sensitive external assessment is also related to time. The scheduling of the assessments should be arranged so that conditions for the students writing the assessments are optimized. For example, teachers decry the scheduling of the Grade 9 Provincial English Language Proficiency Assessment. For a number of reasons, not the least of which is the sometimes difficult transition between middle and high school, teachers believe November of Grade 9 is a very poor time for students to be writing a high-stakes test.

The best designers of instructionally sensitive tests are those who currently provide instruction. Experienced New Brunswick classroom teachers, including NBTA representatives, should be involved in the development of provincial assessments. Those responsible for ensuring outcomes are met – teachers – are those who would best identify appropriate assessment items. Assessments are far more likely to be instructionally sensitive if instructors actually design them. Conversely, professional test developers design assessments to ensure results provide “score spread”. In order to attain score spread, items on which students perform well are often excluded from

the tests. “Items on which students perform well, however, often cover content that, because of the content’s importance, teachers stress instructionally. Thus, the better the job that teachers do in teaching important knowledge and/or skills, the less likely it is that there will be items on a standardized achievement test ...” (Popham, 2006, pp 337-338). In contrast, teachers design tests so students are able to demonstrate learning.

Finally, instructionally sensitive standardized tests must contain no surprises. This does not mean that students have prior knowledge of all test items, but rather that teachers have prior knowledge of the outcomes to be tested, the format and design of the assessment tool, and the scoring rubrics, so these can be communicated to students. Expectations must be clearly stated.

3. The Department of Education, including the Evaluations Branch, must place high importance on effective communication with teachers.

Communication is a key component of an effective evaluation system. Outcomes must be well identified, expectations must be clearly articulated, useful results must be accurately reported, and responses to inquiries and results must be communicated in a timely manner.

While teachers struggle to cover the numerous outcomes in mathematics and science curriculum documents, they struggle to interpret the vague outcomes outlined within the language arts documents. The Department has begun to address this by developing *Standards Documents*, but expectations need to be articulated more effectively by the Evaluations Branch as well.

Articulated expectations must include not only outcomes, but also scoring rubrics. The 2006-2007 results from the Grade 9 Provincial English Language Proficiency Assessment did not show improvement, in part because the assessments were scored with a new rubric that had not been communicated to teachers. Students did not have a clear understanding of exactly what they were expected to demonstrate in the writing assessment. That lack of communication with teachers led to unjust treatment of students. Scoring

rubrics must be clearly communicated to all teachers. Additionally, teachers need to be employed in the marking of the provincial assessments. Teachers report that the opportunity to participate in the marking of provincial assessments is an enriching professional development activity.

Another important component of pre-assessment communication concerns the registration process. Communication problems during the registration process have also caused results to be inaccurate. It was reported, for example, that in a small school 10 children wrote a Grade 3 Provincial Mathematics Assessment. However, the Evaluations Branch had sixteen children listed. Despite multiple attempts by the school to have the 6 children who were no longer attending the school removed from the registration, this was never done. The results for the school were abysmal – but meaningless. The process was an exercise in futility because of poor communication.

Once the provincial assessments have been graded and analyzed, accurate results must be reported in a timely manner, in a format that is user-friendly. Student results should be available to be used to inform instruction. In order for the results to be used in this manner, however, individual student results must be communicated. Teachers seek to use the results as a diagnostic tool better to identify needed interventions. Furthermore, the NBTA posits that longitudinal individual results should be provided, so effective (and ineffective) interventions may be identified, and instruction improved accordingly.

Lastly, privacy protocols with regard to the reporting of results must be developed and maintained. Just as individual student results must never be made public, so too should individual teacher and school results. As professionals, teachers and administrators are expected to analyze results to determine what aspects of instruction should be altered, but such critique should not be undertaken in any public venue. Such uninformed public criticism is damaging to the entire school community.

4. Testing of accommodated, modified, and individualized students must be reconsidered.

Once again, the dichotomy between the provincial testing philosophy and the provincial teaching philosophy is problematic.

In New Brunswick approximately twenty percent of the Anglophone student population is accommodated, modified, or individualized. These students have programs that are tailored to their needs. The resulting accommodations and modifications range from very minor changes, such as particular seating arrangements or specified times for producing written work, to profound differences such as studying Grade 2 outcomes in a Grade 9 classroom. Regardless of the alterations required, teachers work every day to ensure those students feel successful.

External assessments do not take accommodations, modifications, or individualizations into account. Though the exemption process permits virtually all individualized students to be excluded from writing the provincial assessment tests, such students are classified as “experiencing difficulty” and their “experiencing difficulty” ratings are included in the school results. Accommodated and modified children are seldom exempted. Though the accommodated child may not need an exemption, required adjustments must be provided. For example, a child may be highly literate, but physically unable to produce hand - written words. However, no scribes are permitted within the external assessments, so the child would be categorized as “experiencing difficulty”.

A modified student would experience difficulty as well because the outcomes being studied by the child are not those being tested on the grade level assessment. The modified student is forced to write a test on material that has not been taught. Such a situation can be traumatic for the child. Teachers forced to administer such tests to these children feel it is unethical to do so. Teachers speak of high school students who must be successful in the Grade 9 Provincial English Language Proficiency Assessment in order to graduate,

rewriting the assessment four (or more) times, with each experience leading to greater trauma for the child.

The Evaluations Branch must develop and communicate a clearly defined exemption process that allows all students who write the provincial assessments to have some expectation of success. Furthermore, school results should include only those students who actually wrote the assessment in any reports. Continuing to force virtually all children to write inappropriate tests for statistical purposes is unacceptable.

In conclusion, the New Brunswick Teachers' Association asserts that classroom teachers are those who can best evaluate the achievement of their students. However, the Association also concedes that a limited number of external assessments could be helpful provided the following conditions are met: 1) The assessments lead to improvements in student learning, 2) The assessments are instructionally sensitive, 3) Effective communication systems are put in place, and 4) Testing of accommodated, modified, and individualized students is reconsidered. For the benefit of all students in New Brunswick, the Department of Education is strongly urged to ensure these conditions are met.

Popham, W. James, (2006) Assessment for Educational Leaders,
Pearson Education Inc.