

---

# **Student Assessment as Public Learning**

**Report on Consultations and Recommendations for a  
Provincial Student Assessment Policy— Grade 6 to Senior 1**

December, 2004

Submitted to Hon Peter Bjornson  
Minister of Education  
Province of Manitoba

by

Ben Levin  
Faculty of Education  
The University of Manitoba

---



## **Background**

In June of 2004 I was asked by the Minister of Education, Hon Peter Bjornson, to prepare a discussion paper, undertake a series of regional meetings and other consultations, and submit a report with recommendations on a provincial policy for student assessment in the years between grades 6 and 9 (Senior 1). This request was a result of the decision by the Government to discontinue the optional provincial standards tests that had been taking place in grade 6 language and S1 mathematics for the past several years. This document reports on those consultations and includes my recommendations for a replacement policy. I also include a brief rationale for having such a provincial assessment.

---

## **Acknowledgments**

Thanks are due to many people who contributed to this process. I thank the Minister for giving me the opportunity to undertake this work and thereby to contribute to the education policy process in Manitoba. Staff of the Department of Education were very helpful at all stages including thinking through what the consultations would look like, providing background material and organizing all the logistics of the consultations. I am especially grateful for the thoughtful participation of so many people across the Province during the consultations, including educators, school board members, parents, students and other community members who gave their ideas willingly, listened carefully to the views of others, and worked hard to develop good proposals. I hope the result is adequate.

---

## **The Consultations**

After several meetings with staff of the Department to clarify terms of reference for this work, I conducted preliminary meetings with MAST, MTS, MASS and MAPC to gather initial perspectives on the issues involved. Based on these discussions I prepared a discussion paper outlining principles for a policy and providing a basic proposal with some options for consideration.

Staff of the Department then organized on my behalf 7 meetings of interested parties across the province – in Brandon, The Pas, Dauphin, and four in Winnipeg, of which one was aimed at French Immersion schools and another at Français schools and the division scolaire franco-Manitobaine. All Manitoba school divisions were invited to participate in these meetings by sending teams of people, including teachers, school and division administrators, parents, students, board members and other community members. At the request of MTS and MAPC, these organizations were invited to have representatives at all the consultations if they wished. In addition, MTS organized two meetings of teachers from grades 7 and 9 – one meeting in Winnipeg and another in Brandon. The consultations all took place between early August and November 29, 2004.

In all, well over 300 people took part in the consultations, including teachers, parents, students, trustees, school administrators and others. At each session participants were briefly introduced to the discussion paper and spent most of the time, often in small groups, working through the issues and suggestions in the paper. Notes were taken on all discussions and these notes were reviewed in preparation of this report. Participants were told that there was no requirement to come to consensus and that the expression of diverse views was welcome. The discussion paper noted that provincial assessment policy inevitably involves a series of choices and trade-offs, so that there would not be a single desirable approach.

The discussion at all the consultations was thoughtful and constructive. Participants engaged seriously with the issues, provided valuable and sincere perspectives and debated with each other respectfully. Many important points were made and issues raised that have substantially influenced this report. There was a very high degree of support overall for the basic approach outlined in the discussion paper, though many issues and suggestions regarding aspects of it.

---

## **Proposed Policy**

The overall intent of this policy is to create an approach to assessment that is truly educational for students, parents, educators and the general public, in that it stimulates dialogue and discussion around how we are doing and how we might do better. The policy is intended to provide reliable and meaningful information to all parties to support that dialogue.

*I recommend as follows:*

1. That the Department organize and support “The Middle Years Provincial Assessment”. This would be an assessment of all students in grade 7 in two key areas of mathematics and in regard to school engagement, and all students in grade 8 in reading and writing in the main school language and, at local option, in the other official language or in another locally-determined area of study.
2. The assessment would not involve a single test or single set of tasks done by all students, but would assess the same skills using the same standards based on teachers’ accumulated knowledge of students over the year.
3. The assessment would result in a report to the parents or guardians of each student, in a common format across the Province, using plain language and a small number of clearly understood rating categories. Although the assessment process would be ongoing, reporting to parents and official aggregation of results should take place early in the calendar year – between the middle of January and the end of February.
4. To ensure comparability in ratings of standards of performance, the Department should develop, working closely with teachers, students and parents, a set of rubrics or exemplars that would illustrate for all parties the kinds and levels of skills and

attributes represented by each scoring category. Information on these performance standards should be widely available to students, parents and the public to increase understanding of the assessment. Professional development of educators will be a key requirement to build reliability and validity in the assessment.

5. Students should be asked to do a self-assessment on the same key competencies. The results of these self-assessments should be used to support a dialogue in the school and community about ways to increase the commitment of students to their own educational success. Students and teachers should be supported in developing the skills to conduct this kind of self-assessment.
6. In addition to reporting to parents or guardians of each student, every school should include its overall results on this assessment in its annual school report. Divisions should report annually on aggregated divisional results on the assessment, as would the Department of Education at a provincial level. Reports at all levels should also include results by sex, for students whose first language is not English or French, and for students with identified disabilities, unless such disaggregation would threaten individual privacy on account of small numbers. School by school results should not be reported by the Department.

---

## **Rationale and discussion**

### *A word about purpose.*

The discussion paper noted that large scale assessment has multiple purposes and audiences, which inevitably means trade-offs in what is done, but that any assessment should have a worthwhile educational purpose. Not all participants saw the purpose of provincial assessment, because many educators are, rightly, focused on their work in schools and classrooms. However I believe that large-scale assessment contributes to effective schooling in two important ways that cannot be done only through local assessment at the classroom or school. First, provincial assessment can provide information at the system level that helps identify priority areas for attention that we would not otherwise notice. For example, the first year of the Grade Three Assessment showed that primary students' skills in adding and subtracting were not at the desired level. This finding, which would never have surfaced without a provincial aggregation of data, led to efforts at both local and provincial levels to strengthen the teaching of these skills including better professional development and better supports and resources. Identifying strengths and concerns at a system level thus allows us to focus our attention where it may be most needed.

Some consultation participants wondered why this kind of aggregation could not be done using existing data rather than through creating a new assessment. However reporting formats and data across Manitoba are very diverse, making it impossible to aggregate them with any validity. The alternative would then be to have a single provincial reporting format in various grades, which I regard as a much less desirable approach because it would greatly restrict local ability to set priorities and forms for reporting to parents in all areas.

The second purpose for provincial assessment is to provide to the general public relevant and readily understandable information on the performance of the education system as a whole. The people of Manitoba entrust their children to schools every day as well as putting about \$1.5 billion per year into the public school system. They are entitled, I believe, to some reporting on what is being achieved. It is vital for educators to take whatever steps we can to counteract the decline we have seen in public confidence in all our major institutions over the years; otherwise we run the risk of undermining the foundations of our public institutions. Moreover, sound assessment data can be part of a process of encouraging public dialogue and greater public understanding of educational issues and needs.

Yet there are, despite the development of the provincial annual report on student outcomes, relatively few public indicators of the health and success of our education system. I am not suggesting for a moment that this assessment provides a complete picture – it is clearly only a snapshot in some limited areas. Yet every area of public policy uses such limited indicators – think of GDP or unemployment rate in economics, or waiting times in health care, or recidivism rates in criminal justice. In every one of those fields practitioners would argue, rightly, that the indicators give only a limited, and sometimes even a distorted picture. But the reality of public life is that people only have limited time to consider these issues and they rely on such indicators to give them a general sense of what is happening. The task is to design an indicator that, as far as possible, gives public information that is accurate and also educationally sound, as well as one that stimulate public discussion about our hopes and expectations for our schools.

These considerations underlie the proposals in this report, and that is why I describe this assessment as having the goal of stimulating public dialogue about educational outcomes and options – in other words, as a learning activity for the people and schools of the Province.

---

## **Detailed discussion of recommendations**

1. That the Department organize and support “The Middle Years Provincial Assessment”. This would be an assessment of all students in grade 7 in two key areas of mathematics and in regard to school engagement, and all students in grade 8 in reading and writing in their primary school language and, at local option, in their second school language or in another locally-determined area of study.

This first recommendation speaks to two key points – the key areas to be assessed and the grade levels at which the assessment should take place.

The discussion paper suggested five key competencies or attributes for the assessment – two related to language skills, two related to mathematics skills and one regarding students’ level of engagement with schooling. There was

broad support for having a limited number of areas for assessment, and also for these five particular indicators, though many participants correctly noted that there were other important areas that could equally be considered, such as physical fitness, understanding of history or science, and so on. I believe that the original proposal for the five areas remains valid and desirable, and balances breadth with feasibility for teachers and levels of public interest. It is important to note that this proposed provincial policy is only one part of the continuum of student assessment practices in the school system. Schools will, quite properly, continue to assess students and report to parents, through report cards and parent-teacher meetings, on a much broader range of areas across the Manitoba curriculum. The provincial policy on school planning and reporting requires schools and school divisions to report regularly to their community on various student outcomes seen as important. The Department of Education should continue to issue its annual report on student outcomes using provincial data, and should gradually expand the range of indicators reported. Taken together, these efforts should provide students, parents and the general community a variety of evidence on a broad range of student outcomes.

There can be little doubt that most Manitobans regard a reasonable level of reading, writing and arithmetic competence as absolutely vital for all students. The inclusion of engagement requires some further rationale. Student engagement refers to the degree to which students are actively involved in and take responsibility for their education; whether, in short, they see schooling as ‘theirs’. A considerable body of research, as well as educators’ own experience, shows that students’ sense of involvement in their education is vital to their effort and success. Moreover, engagement with learning is critical to students’ capacity to be lifelong learners and is likely to be predictive of their ability to take on new challenges after they leave school. Finally, including engagement in the assessment will broaden its focus beyond specific curricular outcomes and, together with the student self-assessment discussed later, will support a dialogue about student commitment to learning among, students and parents that can itself have useful educational consequences.

Reading and writing are self-evident skills to include though considerable work will be involved, discussed later in this report, to determine what aspects and levels of reading and writing will be included. Mathematics is somewhat more complicated. Most people would take the view that students in grade 7 or 8 need to know how to do basic arithmetic operations (add, subtract, multiply and divide whole numbers, fractions and decimals) with speed and accuracy. However mathematics experts point out that these skills are embedded in a larger understanding of and comfort with numbers, and that one ought to assess computational skills while also looking at this larger understanding. The two proposed mathematics competencies are designed to address both aspects. “Operational skills” takes into account ability to do mathematical operations with fluency, and “number sense”, which refers to a broader level of comfort with and ‘feel for’ numbers.

There is no one right way or place to organize large scale assessment, as shown by the enormous variety of arrangements in place in Canada and elsewhere. Almost every possible set of grades, subjects and arrangements has been used or is being used somewhere. During the consultations many suggestions were made as to key areas for assessment, grades at which the assessment should be done, and form of reporting. Some participants thought that the assessment should happen in grade 6, others in grade 9. Some felt that the entire assessment should be done in the same year and others favoured a split over two years. There was general support for the choice of grades 7 and 8 as occurring towards the end of the elementary/middle years and before students begin high school. The largest body of opinion agreed that conducting the assessment over two years would provide more opportunity for teachers and parents to respond to concerns.

The main concern of participants in the consultations about grade levels was, very properly, that the assessment should be done at a point where it could still affect students' subsequent schooling. The proposal to assess mathematics in grade 7 is precisely to give more time (until the end of grade 8) to ensure that students' mathematics skills are adequate for the challenges of high school and particularly for the grade 9 course.

French Immersion schools would be expected to conduct the language (reading and writing) assessment in both French and English. Français schools should conduct the assessment in French and have the choice as to whether to assess students also in English.

The assessment process is flexible enough that a school or division may choose to add one more element to it that is particularly significant locally. For example, a school that had given considerable priority to social studies might want to include an assessment in that area as part of this process. However the Department of Education should not be expected to provide supports for such local choices.

I also recommend that the Department extend an invitation to First Nations schools in Manitoba to participate in the assessment if they wish to do so, including providing access to resource materials and professional development activities.

2. The assessment would not involve a single test or single set of tasks done by all students, but would assess the same skills using the same standards based on teachers' accumulated knowledge of students over the year.

This recommendation also balances between two competing elements of assessment. On the one hand, it is critical that an assessment addresses the same skills for all students with enough reliability so that parents and the public can be confident that reports from across the province are truly comparable. On the

other hand, an agenda of improvement relies on the understanding and commitment of teachers and students, and should use their deep knowledge of particular students and situations to this end.

I am recommending against provincial standards tests that would see all students do the same test at the same time because such tests are costly to develop and administer, are demotivating to many students and teachers, and do not necessarily produce more valid and reliable data on student outcomes. A discussion of the merits and limits of large scale testing would require its own extensive paper, so here I will simply assert that we can use teachers' accumulated knowledge of students combined with ongoing professional development to achieve sufficient reliability and comparability and also much higher commitment to use of the assessment results. The consultations showed strong support for this approach among all parties.

Teachers will also want assistance in ensuring that their professional judgment about students is being applied fairly and consistently – for example by being given examples of the kinds of tools and processes that can be used with confidence in their reliability to form judgments about students' skills. Ongoing professional development for teachers will be vital to the success of this policy, as it is to all efforts to improve student outcomes.

It will also be vital to involve school leaders in this policy in an active way, including appropriate professional development around how to support its various components.

In most schools it will be evident which teachers will conduct the assessment in language and mathematics; this will be the teacher responsible for that subject for the students involved. At the same time, I suggest that schools think about the assessment as a school team activity rather than as an assignment only for some teachers in grades 7 and 8. The greatest benefit from the assessment will come when it is treated as a responsibility of the school as a whole, so that there is the greatest commitment to using the results to inform school planning.

The importance of a team approach is most clearly evident in thinking about assessing student engagement, since this is an issue that spans all the subjects. In order to distribute the work among more teachers and, even more importantly, involve more teachers in the discussion around students' performance, I suggest that each school determine a teacher or teachers to be responsible for the assessment of student engagement. I further suggest that in making the assessment on engagement, the designated teachers consult with some other members of staff who teach the students involved to get a fuller picture of students' engagement. For example, a school might decide that in a given year the assessment of engagement will be the primary responsibility of physical education or music teachers, or guidance counselors, or science teachers, in each case working with colleagues in other subject areas. It would

not be reasonable to expect all teachers to be consulted, especially in schools where a grade 7 or 8 student might have half a dozen teachers, but it should be possible to involve at least 1 or 2 colleagues. The advantages of this suggestion are: that it will create more discussion in the school about the vital issue of student engagement, that it will give more teachers an opportunity to participate in the professional development around the assessment, and that it will extend the assessment to subject areas other than language and mathematics.

3. The assessment would result in a report to the parents or guardians of each student, in a common format across the Province, using plain language and a small number of clearly understood rating categories. Although the assessment process would be ongoing, reporting to parents and official aggregation of results should take place early in the calendar year – between the middle of January and the end of February.

It will be very important in this process to develop a set of descriptors of performance on the assessment that are readily meaningful to parents and other non-educators as well as to teachers. Educators often have specific technical language to describe skills but these terms do not carry very much meaning to the non-educators who are a primary audience for this assessment. Various options were tested in the consultations, such as ‘emerging’, ‘developing’ and ‘proficient’ or ‘struggling’, ‘some understanding’ or ‘mastery’. None of the proposals was entirely satisfactory. On the one hand, terms have to be accurate and meaningful to lay people. Participants in the consultations had a hard time distinguishing between ‘emerging’ and ‘developing’. On the other hand many participants were concerned that terms not be demotivating to students or parents, so did not care for words like ‘struggling’. The option that seems preferable to me would be to couch the terms in relation to grade standards such as ‘at or above the expected skill level for this grade’, ‘near the expected level’ and ‘significantly below the expected level’. These terms are entirely neutral but if accompanied by clear examples of what grade level performance looked like, would be meaningful to all audiences. It will therefore be important to provide to parents not only the ratings, but sufficient information about the scoring so that parents can form a reasonable judgment of what the ratings mean.

The rating descriptors should be exactly the same for reading, writing, and the two mathematics competencies. However student engagement will require a different set of ratings. I suggest that these should be: ‘Generally positive’, ‘Mixed or neutral’ and ‘Generally not positive’. Once again, it will be important to provide for parents a clear sense of the basis for these judgments.

There was some debate about whether to have 3 or 4 descriptors, with some participants wanting to distinguish students whose performance was exceptionally strong. Given the purpose of this assessment – to provide an overview of areas of strength and weakness on these key skills – I lean towards

3 categories as being simpler for all concerned. Parents will get fuller information on the children's level of performance through school report cards.

An issue raised several times during the consultations had to do with evaluation of students with special needs. Educators were concerned that some students with significant disabilities could be rated on the assessment as doing poorly even though they were actually making significant progress relative to their starting point and capacities. This is an important issue, but also a complex one. While an assessment should not set up unrealistic standards of performance, we must remember that historically we have consistently underestimated what various students could do and thus limited their opportunities. At one time women were thought incapable of benefiting from high levels of education. Much more recently, children with Down Syndrome were regarded as uneducable. Now we are learning that many students have much more capacity than we had expected.

I therefore recommend that all students, with no exceptions, be included in the assessment. (This is another reason why I favour neutral language in the rating categories). Where students have an individualized education plan, parents will already be aware of this and their expectations for the results of this assessment will be adjusted accordingly. Another option would be to add a fourth category to the assessment ratings, which would be "Performing significantly below normal grade expectations due to an identified disability" or, for reading and writing, "Performing significantly below normal grade expectations because English is a second language". I suggest that schools or divisions wishing to use these options discuss them with parents of special education or ESL students before making a decision to do so. The impact of these issues on school reports is discussed below.

4. To ensure similar ratings of standards of performance, the Department should develop, working closely with teachers, students and parents, a set of rubrics or exemplars that would illustrate for all parties the kinds and levels of skills and attributes represented by each scoring category. Information on these performance standards should be widely available to students, parents and the public to increase understanding of the assessment. Professional development of educators will be a key requirement to build reliability and validity in the assessment.

A critical element in the process will be the development with much input from teachers and students of a set of criteria and ratings scales that will constitute the assessment. Teachers across the province need to develop a common understanding of what it means to demonstrate reading or writing or number sense skills at a level expected of a grade 7 or 8 student. This means that teachers need to come together with external consultants to examine student work and debate standards. While teachers must play a central role in this process, I emphasize again the importance of extending the discussion to

students, parents and the community with the important goal of building very specific understanding across the community of the kinds of skills we are expecting from young people in the middle years. A school or division or provincial report on the outcomes of this assessment will be much more meaningful if readers can relate it to clear and understandable standards.

Particular attention will be needed to developing a set of ratings in the area of engagement. Students should play an active role in creating these descriptors and rubrics. The concept of engagement should include a range of behaviours. For example, although one element of engagement would include taking pride in one's work and getting work done on time, the concept should not be limited to compliance with school requirements. Engagement could also extend to active participation in the life of the school, both in and out of classes, to support for peers, and even to actions that are critical of the school as long as they are intended to be positive.

A particular effort should be made to reach all parents and guardians, including those who are least likely to take the initiative in these areas. Steps in this direction should include translation of the assessment into other languages where appropriate, ensuring that 'plain language' reports are available, organizing sessions at the school to explain the assessment and its results to parents, and using the media (such as local community newspapers) to inform people in advance about the process and later to inform about its results.

5. Students should be asked to do a self-assessment on the same key competencies. The results of these self-assessments should be used to support a dialogue in the school and community about ways to increase the commitment of students to their own educational success. Students and teachers should be supported in developing the skills to conduct this kind of self-assessment.

A central feature of schooling is that in the end, whatever efforts adults make, it is students who do the learning. An important element of any assessment, then, should be its connection to students' motivation and aspirations. Good assessment helps students see how they can do better, and why they should.

Powerful recent evidence on formative assessment shows that students can benefit enormously from thinking about their performance and how it might be improved. Nowhere is the need for this kind of reflection greater than in the middle years, when students are moving towards adulthood and a much greater level of responsibility for themselves. I am therefore recommending that as part of the middle years assessment, students should be asked to do a self-assessment on the same skills and competencies, using the same rubrics and in roughly the same time period. The goal of this activity is to engage students in thinking about their skills. A second goal is to create the possibility of discussion between teachers and students (individually or collectively) and between parents

and students about what constitutes reasonable performance and what might be required to help students do better.

For many students and teachers, student self-assessment will be a new activity. There is some evidence that students may be harsh critics of themselves, and some concern that they may not take the activity seriously. The Department should therefore assist teachers and students in this area by developing some resource materials (including some specifically aimed at students) around the purposes and value of self-assessment, and by providing professional development for educators on ways to foster meaningful self-assessment by students. I would hope that the Manitoba Teachers' Society might also engage itself in supporting this area of work.

The results of students' self-evaluation should be used in three ways. First, individual students should compare their self-assessment with their teacher's assessment and, where there are discrepancies, discuss further why these views are different. Parents and guardians can also be important participants in these discussions. The goal is always to help students see how they can improve their skills. Second, teachers should compare their assessment with that of their students, looking for systematic differences that might suggest new opportunities to inform or motivate students around success. The same should happen at a school-wide level. However student self-assessment results would not need to be reported by schools to the Province.

6. In addition to reporting to parents or guardians of each student, every school should include its overall results on this assessment in its annual school report. Divisions should report annually on aggregated divisional results on the assessment, as would the Department of Education at a provincial level. Reports at all levels should also include results by sex, for students whose first language is not English or French, and for students with identified disabilities, unless such disaggregation would threaten individual privacy on account of small numbers. School by school results should not be reported by the Department.

An important purpose of the assessment is to inform program development and policy at the school, district and provincial level. This can only happen when results of the assessment are aggregated and reported at these levels, as happened, as mentioned earlier, with the Grade Three Assessment and adding skills.

The best way for schools and divisions to report their assessment results is as part of their annual reports to the community. This format allows these results to be put in the context of a school or division's plans and the entire range of student outcomes being reported. It also means that schools are not simply reporting results, but connecting these results to the school's priorities for improvement. While each assessment result is potentially important, a full and accurate picture of student achievement depends on a range of different measures, not just on one. Similarly,

the Department of Education would include middle years assessment outcomes as one part of its annual report on student outcomes.

In reporting assessment results, schools and divisions should provide appropriate contextual information such as the proportion of students with identified special needs, the proportion of ESL students, the annual student turnover rate, and so on. These features of a school's context are important background to understanding achievement patterns. As well, given concerns about the relative achievement of various groups (such as boys in reading or girls in mathematics, or Aboriginal students or recent immigrants), schools and divisions should attempt to disaggregate assessment results for target groups unless numbers are small enough to make personal privacy a relevant consideration. The Department should discuss with divisions and schools how to report data provincially to allow some of the same analysis by subpopulation.

Ranking of schools is not a good approach to publication of results because it leads to unreasonable comparison among schools. The issue should never be where a school ranks on a list, but should always be what proportion of its students achieves the required standard. A significant change in ranking may or may not mean a significant change in actual student performance. Most schools will be near the middle of the distribution in assessments such as this, so trivial differences in performance could lead to very large differences in rankings. In other words, our approach to assessment results should be aimed at increasing performance everywhere, not on getting to the top of a pile.

---

## **Conclusion**

I believe that this policy provides an opportunity for Manitoba to be at the forefront of national and international efforts to use large-scale student assessment as a vehicle for improving student outcomes and building public understanding of education. We can avoid the excesses of endless testing while providing valid public information and focusing on fostering student success. I wish the province much success in moving in this direction.

Respectfully submitted,

Ben Levin  
Dec 27, 2004