Developing a Diagnostic Assessment Model

James Wagner

The recent implementation of annual testing at the grade 3, 6 and 9 levels by the Education Quality Assessment Office has forced the school boards in Ontario to try to come to grips with the problem of measuring student achievement. This problem is not a new one in that teachers have been concerned about evaluating student progress since the advent of schooling. In this respect, the literature suggests that most teachers are quite capable of identifying the students in their classrooms who are or are not making sufficient progress. Why then do we need formal testing and in particular, large scale high stakes assessments?

The answer to this question goes to the purpose of assessing student progress. No one would argue against the notion that it is important for teachers to know whether or not their students are making progress. Principals and parents also need to know this. If one accepts the idea that higher-level administrators are responsible for student learning, then they too need to have access to such information. Recently, this logic has been extended to the Ministry of Education. This, in turn, leads to a second issue, which is of equal importance. What form should this information take? For example, anecdotal reports become less feasible the higher one goes in an administrative hierarchy, if for no other reason than they generate too much information. Thus there is a tendency to move towards assessment instruments which yield numerical data. There is also a tendency to aggregate results the higher one goes. Aggregation in turn leads to the need for standardization (one cannot aggregate the results of different tests) and a focus on group performance. Moreover, administrators often impose these constraints on an assessment practice so that they can evaluate results relevant to their level of authority. Thus it should come as no surprise that the Ministry of Education has imposed standardized tests with quantifiable results on the province.

However, there is a second purpose to assessment that appears to be missing in this approach. Not only does the classroom teacher need to identify those students who are not making progress but equally important, he/she must ask why. In other words he/she must attempt to learn what is holding the student back and of course, what pedagogical strategies would be required to change this. In this respect, assessments that do not yield this type of diagnostic information only do half of the job and are of limited value.

Of course, classroom teachers may often identify the causes of a student’s failure to learn but in Ontario they are hampered by the lack of explicit causal models in the curriculum and EQAO tests. There is very little in the provincial curriculum or the EQAO Reading Test for example, that yields useful information on whether or not the problems a student is having in word recognition are caused by developmental lag in phonemic awareness or oral receptive vocabulary.

In response to this issue, the Reading Clinic in the Faculty of Education at Brock University has been involved in some exploratory projects with the Niagara District, Hamilton-Wentworth District and Halton District School Boards designed to develop assessment models that can be utilized in diagnosing the progress or lack of progress of their students. The idea behind these projects is that tests are of little educational value unless they yield diagnostic information and to do so they must be based on sound models of the learning they purport to measure. This in turn implies that the learning must be made explicit in a model that reflects not only curriculum outcomes but also the causal factors underlying those outcomes. Constructing such a model is not easy as there is often no consensus on what these factors are and how they fit into a general framework. Moreover, educators sometimes confuse assessment models with teaching models. For example, a causal assessment model is meant to determine to what degree a factor such as oral vocabulary (word meaning) develops. However, it will not tell us what teaching strategy should be used to develop oral vocabulary if it is identified as a problem. In this respect the fact that one isolates oral vocabulary development in an assessment model does not imply that one should isolate it in teaching.

The causal assessment models that we have explored were meant to identify those factors that might affect early reaching success. Figure 1 is an example of one such model. Note that it is not claimed that this is an exhaustive model or that it is the only model one might construct. Rather, the purpose of our model is heuristic in that it is meant as a guide in di-
agnosing children’s reading progress and in selecting assessment instruments (individualized or standardized) which yield the diagnostic information we want. Note too, that the model can be used at both the individual and the group level. Thus a teacher might fit assessment results to the model to identify whether or not a particular student was lagging in phonemic awareness and whether or not this appeared to be a causal factor in his/her word recognition. However, teachers and administrators might also use classroom results to identify whether or not there was a more general problem in phonemic awareness, which in turn would lead to the question of whether or not program changes were required. The data base a principal might establish in such cases, could be stored and analyzed in Excel or SPSS and once established, could be used to chart and evaluate changes or improvements as new programs are implemented. Such changes would be assessed not only by comparing the pre and post-test scores of children on a skill such as phonemic awareness but also by looking at the effect of such changes on word recognition, which the model predicts should improve as well. These ideas have been implemented in a number of schools in the Niagara and Hamilton regions and are beginning to yield diagnostic information. One school in Hamilton for example, has identified listening comprehension as a critical skill in which many of its grade one children appear to be lagging. It is too soon to tell however, whether or not program changes in response to this information are having an effect. What is implied here is that changes as a result of new programs take time and that pre-posttest designs are required to measure them.

One of the problems we have run into is that the EQAO tests do not appear to be based on a pre-post design. As a new cohort of students is tested each year, there is no way of measuring growth as a result of program innovations. Perhaps more serious, however, is the difficulty we have encountered in discerning what if any, causal model underlies the EQAO tests and how it might be connected to traditional causal models of subjects such as Reading, as illustrated in Figure 1. There may of course, be some diagnostic value in the response categories (Reasoning, Communications, Organization, and Conventions) used in the EQAO tests but it is not clear how these categories would fit together in a causal model of reading. Put another way, there appears to have been no attempt to make explicit the relationship between these categories and the outcome category (reading) apart from adding them up. Equally critical, we do not know how factors such as phonemic awareness or word recognition are meant to be related to these response categories. Until the Ministry of Education or EQAO makes these relationships explicit in their model of reading, we are left with making program changes based on global or holistic comparisons with EQAO results, which may or may not yield the progress we want.

**References**


