

Response to “Oops, that was a mistake: Examining the effects and implications of changing assessment policies

By Joan M. Green
Discussant

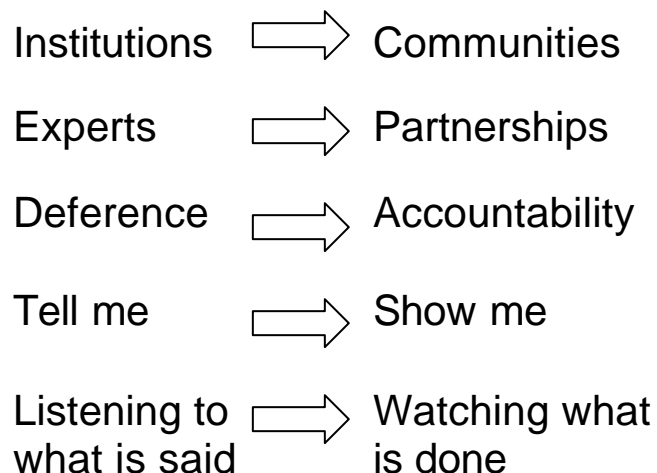
School improvement preoccupies everyone involved in education, whether as a policy maker, educator, parent, student, or citizen.

Around the world, the job of educators in this new century is huge and complex. Sometimes we feel that we live and work in the eye of the storm. Educators' challenges are acute, the demands high. Never has the scrutiny been more intense, or the fiscal pressures more fierce.

In Canada and in many other places around the world, the context in which educators work and live is profoundly different than it was twenty, ten or even five years ago. The parents in our communities and the students in our classrooms are more informed, more involved, more sceptical about accepted wisdom, more questioning of authority, and more intolerant of ambiguity. More is expected from teachers and schools than ever before.

From the outset, we must acknowledge that students want and need to know how well they are doing and where they stand, as well as what they need to do in order to improve. Teachers need informed and timely feedback on how their students are doing and on what teaching practices are working well. Parents need to be aware of the content taught and standards expected in our schools, they need specific information about their children's performance and progress. Citizens need evidence that education of a consistently high quality is available in all schools and that resources are being used effectively to maximize student learning.

The following illustration points out the shift in focus that has occurred in how people respond to information from public institutions. We have moved from



My overall response to Don Klinger's paper "Oops; that was a Mistake: Examining the Effects and Implications of Changing Assessment Policies," is to say yes, changes in large-scale assessment will sometimes result in misinterpretation. Sometimes however, those changes are unavoidable if assessment programs are to evolve constructively on the basis of sound advice from practitioners and learning over time as large-scale assessment wrinkles are ironed out. Klinger makes the point that "unfortunately, given the lack of assessment expertise of most teachers, administrators, district personnel and the public at large, inappropriate comparisons will be made between the two years and the differences attributed to a variety of factors, both real and imagined". (Klinger,p.11). He argues strongly for getting it right at the outset as far as possible through investment of appropriate time and resources in the original design and delivery. While no one could argue with that view, it is important to emphasize, as Klinger does in his paper, the vital contribution that the involvement of teachers and administrators can make to the development and implementation of the assessment program. Clearly, they will have experience over time, which will lead to necessary changes in the product and process of the testing exercise. In those cases, I would say that Klinger's argument that "work is also required to support the proper use and interpretation of these results" is well-founded. He asserts

"Given the general lack of expertise regarding the use and interpretation of assessment information, it is important that the foundations for assessment be strongly conceptualized and the need for subsequent changes minimized. As the expansion of large scale assessment programs continues across Canada, those responsible for the implementation of these programs need to ensure that sufficient time is provided to develop assessment programs that meet their stated goals when they are implemented." (Klinger,p.13).

In at least one example of changes to large-scale assessment practices that Klinger cites in his paper, one could argue that the changes, in the case of the British Columbia Foundation Skills Writing Assessment were rooted in good advice from educators. The addition of a second writing sample, the change from one prompt to grade-specific prompts, and the change in scoring to match emerging performance standards, may all have resulted in a better assessment. However, as Klinger points out, "modifications and changes to assessment programs have implications for students, teachers and policy-makers" (Klinger,p.10). I would add parents to this list and emphasize, as Klinger does, that when these changes are introduced, it is vital to provide adequate discussion of both the similarities and differences between the old and new assessments so errors of interpretation can be minimized, and misleading conclusions drawn from the results avoided. I wholeheartedly agree with Klinger's assertion that "it is important to examine the effects of changing assessment procedures from both a technical and policy perspective,"(Klinger, p.3) because it is clear that changes to

assessment programs can dramatically alter the impact of the assessment and create some unintended consequences. .

Klinger's title "Oops: that was a Mistake" sends the pulses of all of us involved in the scale assessment racing because we know that inaccurate reporting or ill-considered processes can "unfairly harm students or lead to misinformed policy directions." (Klinger,p.13). However, to paraphrase Dr Lorna Earl's observations in her address to the conference, we can't avoid mistakes in large- scale assessment programs; we need to make them faster, make their implications clear to all stakeholders and get over them and move on.

As several conference participants declared in their discussions about empirical evidence and the constructive use of data to improve educational outcomes for all students, the story is no longer about opportunity to learn; rather it is about higher achievement for all and transparency about how systems are progressing toward the realization of that vital goal. "If a school gives up on a kid in 2001, the price to be paid by the student and society is much higher than it was in 1950. Now that we succeed at keeping most students in school, we must figure out ways to educate everyone we keep there. This is where some of the most heated debate is occurring today." (Green,J.M., Speech to The Empire Club,1998)

There are clearly great differences in the way assessments are conducted and results reported from jurisdiction to jurisdiction and within districts over time. Carol Rolheiser; Associate Dean of OISE/UT suggests we would all do well to ask the following question: " What movie does the a nalysis of data about performance in your school or school district remind you of:

- Titanic?
- As Good as it Gets?
- What Lies Beneath?
- The Full Monty?"

In September 2001, along with three other Canadians, I participated in an international meeting on authentic assessment held in Chester, England. One of the books discussed at the conference, Assessment in Education: Principles, Policy and Practice, included an article by Ginette Delandshere who posed three questions that we need to continually ask about our assessment programs to examine the basis on which judgements are formed through educational assessments:

- Why is the assessment used, what are its purposes and justification?
- What is the form and nature of education assessment, or, more specifically, how is knowledge defined and by whom?
- What is the nature of the relationships between and among the various agents (e.g. teachers, students, parents, the government) involved in the

assessment process, and the social and political context within which they interact?

Ethical, sociological, and epistemological issues under-gird all three questions and need to be addressed as we develop, modify, and extend large-scale assessments.

I would now like to comment on the recommendations Klinger makes at the end of his paper.

1. "Given the expense of sound assessments, carefully consider the need for such new or expanded assessments."

Sound assessments that address both what students know and can do when presented with authentic tasks are indeed expensive. They are also rich in the data they provide to assist in the improvement process. These assessments are useful at key stages of schooling for both accountability and improvement purposes. However, they do not need to be annual events in every subject to provide the system, school and individual level information that augments and supports high quality daily classroom assessment.

2. "Be honest with respect to the use of assessment programs and implement procedures to prevent the misuse of the results."

Clear public statements about the intended use of assessments for fair accountability and constructive improvement purposes are essential to the integrity of any large-scale assessment process. For example, through its Chair and CEO, the Ontario Education Quality and Accountability Office consistently opposed the invidious comparisons which arise from simplistic rankings. EQAO also provided two methods for public reporting in order to take into account all the variables that affect achievement. The concept of "statistical neighbours" through which schools with similar challenges and successes can learn from each others' improvement efforts was another approach promoted to "prevent the misuse of results" which Klinger rightly deplors.

3. Develop a sound implementation plan, and provide time to implement procedures to prevent the misuse of the results.

The sound implementation plan that Klinger advocates depends on the involvement of educators in the design, marking and reporting on the assessments. The development of exemplars and anchors for use in

discussing expectations and results with students and parents does take the kind of time to which Klinger refers. It is well worth the investment.

4. Involve key stakeholder groups

Stakeholder involvement and understanding of key issues are crucial. The whole large-scale assessment process should demystify both expectations and results for parents, educators, students and the public. Outreach to key media representatives to explain purposes and processes is very useful in enlightening the public debate about the testing program.

5. "Identify potential problems before rather than after the fact."

As in any enterprise, the early identification of potential problems is a key contributor to success. For example, the development of appropriate strategies and procedures for involving students with special needs is imperative if their participation is to be meaningful and useful in supporting their academic achievement.

6. "Communicate results in a clear unambiguous manner."

The public reporting of both results and recommendations for improvement has to be transparent and accessible. On a provincial or district basis, it is often helpful to require schools and school boards to use a template that can serve as the building block to which their more specific local information can be added.

7. "Do not allow reduced costs to be a factor for changing assessment practices."

Klinger is absolutely right to admonish that appropriate funding of assessment processes is essential if the assessments are to genuinely examine both what students know and can do and if the results of these assessments are to be used productively for improvement purposes. Simple off-the-shelf commercial tests, incompletely matched to the curriculum, cannot achieve this goal.

8. "Carefully consider the range of implications associated with changes to existing assessment programs."

There is no question that changes can jeopardize the validity of the assessment for the purposes that were originally intended. Certainly, comparability between one assessment and another can be compromised if the assessment changes radically from year to year. When target-

setting for achievement improvements are based on test results, it is particularly important to reflect on the impacts of changes to the process. However, the assessment business is all about learning to do better and there are circumstances where, after “the careful consideration” that Klinger argues for, changes are necessary and lead to more effective assessments.

9. “If changes are made, examine procedures to maximize comparability or develop policies to prevent inappropriate comparisons.” (Klinger,p.13).

There is no question that when changes are made every effort must be made to determine what conclusions can be sensibly drawn for the year-to-year results. The credibility of data-driven decision-making for improvement planning depends on this thorough scrutiny.

By way of conclusion and to reflect the spirit of Klinger’s paper, I want to refer to the core values which guide the work of the Education Quality and Accountability Office in Ontario of which I was the founding CEO. *EQAO’s Index to Effective Assessment* defines the essential determinants of a useful assessment system that can contribute to improved teaching and learning in all schools:

- 1. Large-scale assessment must give all students a chance to demonstrate what they know and can do.**
- 2. Assessment processes must provide capacity-building opportunities for teachers.**
- 3. The accountability function must be active, moving beyond statistics to plans for improvement.**
- 4. Large-scale assessment must be performance-based.**
- 5. Large-scale assessment must be standards-based.**
- 6. The reporting of assessment information must lead to improvement for the individual learner and for the system as a whole.**
- 7. Each assessment must provide contextual information that allows educators, parents and the public to interpret results in a sensible and constructive manner.**
- 8. A quality indicators system must be developed to provide a richer context for reporting student achievement and for assessing quality in the school system.**

9. The testing authority must be seen as independent of any government agenda and accountable itself.

10. Assessment initiatives must recognize that all large-scale change is ultimately the result of local implementation. (Green, J.M., 1998)

As Klinger says in the conclusion of his paper, “the importance and use of large-scale assessments is increasing, resulting in the development and expansion of assessment programs.” (Klinger, p.13). I agree that, wherever possible, we need to get it right the first time in order to minimize misinterpretations of the data or misinformed policy directions. However, it is essential to remember that sometimes assessments change because people figure out how to make them better, and these decisions are based on experience that was not available at the outset.

References:

Green, J.M. (Fall,1998). Constructing The Way Forward for All Students.
Education Canada : Education Reform: Are We On the Right Track?,p.10-11.

Green, J.M. (1998)). Greater Accountability and Better Quality in Ontario Schools.
The Empire Club of Canada: Addresses 1998-1999, p.275.

Klinger, Don. Oops, that was a mistake: Examining the effects and implications of changing assessment policies. *Paper presented at the Empirical Issues in Canadian Education Conference,Nov.2001*