Education and Innovation Committee

From: Mary-Ann Parlato [redacted]
Sent: Sunday, 5 May 2013 1:36 AM
To: Education and Innovation Committee
Subject: QLD Assessment Methods feedback

Follow Up Flag: Follow up
Flag Status: Flagged

Categories: Orange Category, Green Category

Education and Innovation Committee
Parliament House
Brisbane Qld 4000

27th April, 2013

Please withhold my contact details

To the Education and Innovation Committee,

Please see my feedback below in regards to the current assessment methods within Queensland senior schools.

Issue:
Student educational records
are relevant to the Term of Reference Two: Student Participation Levels

The current Queensland senior school system appears extremely subjective and qualitative in its measures. It begs the question, how can this be comparable to the rest of the country and ultimately how can this be comparable with the rest of the world? Any student that needs to travel interstate or across borders may have difficulties sharing information on their achievement in these subjects when entering another system. The subjective nature of the assessment allows for a massive margin of error/difference in opinion; The marking guidelines that I have viewed from the accredited curriculum body, Queensland Studies Authority (QSA). have an inherent design that could be very inadequate in reporting on the student's true knowledge and skills achievement. For example, a student will fully answer a question correctly, but because further detail was not provided, then that student is not seen as not knowing the "nature of and use of mathematical proof" (in the higher grading requirement).

I manage global market research programs for the largest telecommunication companies in the world: Apple, Samsung, HTC, Motorola and many others. In market research, particular focus on how questions are worded, and ways in which people take in a question can impact their answers. For instance, how would you rate your satisfaction with Apple on a scale of 1 to 10 where 1 is very dissatisfied and 10 is very satisfied. Even on this intuitive scale, to one person, 6 out of 10 may be an acceptable level of satisfaction while very satisfied would be rated 8 out of 10. However, for someone else, they may see 8 as an acceptable level of satisfaction and for them, to be very satisfied would be a 10 out of 10. Automatically, there is a need to calibrate responses and that is possible by adjusting the scores. That is necessary because we are asking a question that can be quite subjective and can lead to two very different answers for what can mean the same performance from a technology product.

So, by this curriculum body setting marking guidelines that do not directly test the cut-and-dried solutions of a mathematical calculation, they are instead asking for more subjective interpretations. If even only two teachers are asked to rate a child's ability to "know the nature and use of mathematical proof" , they will understandably come up with two very different answers. Are we considering that the consequence of this alternative approach is that students have to hope they have a teacher that has a lower level of expectation? How does this impact on their participation in a subject that is marked this way?

Recommendation: Student participation or results may be dampened by an intuitive sense that their assessment approaches seem 'unfair' (invalid or unreliable). However, results determined in invalid ways may mask the underlying system's faults and may advantage some students, who might not have issue. In the absence of randomised, scientifically-based surveys it would appear prudent to investigate community sense of the fairness or otherwise of the methods of this system. However, I would emphasise that regardless of the community input into, or approval of, this system, a critical reason for upholding a system is whether it fits international best practice standards for assessment in terms of validity and reliability and what it achieves for the clients, i.e., the students and citizens benefiting from the school system.

Issue:

The relevance of the measurement methods
refers to Term of Reference Three: The ability of the Assessment processes to support valid and reliable judgments of student outcomes

Qualitative measures, like the ones in this current assessment response system, do have a use when looking deep into particular issues and providing feedback for further purposes. In market research, we use qualitative research to explore a chosen area to help explain a certain trend that we are seeing. However, large and small brands are interested in the quantitative (numerical) measure i.e. Your mean score for satisfaction is 6 out of 10 versus your competitor which is 4 out of 10. A broad picture is provided by quantitatively collected data; in turn the necessary overall information is required before any finer precise information can be sought. The survey questions, analysed quantitatively, are thus seeking information that is valid (finds out what it purports to survey). Reliability, the ability to see the same information when the assessment results have been collated by another person or the same person at another time, is vastly more certain when the provided answers are defined objectively and the quantitative data entry and analyses follow non-ambiguous guidelines. By comparison, the current school assessment guidelines in Queensland appear so rubbery and so unreliable that it invites doubt that two people marking the same paper would agree. Indeed, even the same person marking the same student's paper at another point in time might not award the same result.

Ultimately, the Queensland marking system is using a technique that should not be used to provide results and ranking. Instead it may be used when a student's performance or a particular cohort (for e.g., an entire school, region or perhaps only a class) has been found to differ greatly in score compared with the average of all cohorts. Further research could be done into why that student or particular cohort have lower than expected scores in comparison with others. This, of course, also necessitates a common 'survey' at the start. So if, for example, a cohort received maximum about 60% of answers correct compared with another, then further exploring what answers they got wrong would enable greater understanding of where improvements need to be made. Simplistically put, this is analogous to how companies that employ market researchers elucidate where they need to improve to satisfy customers.

How does the QSA know if the Queensland system is in the forefront of marking systems across Australia? How can comparison even occur when the system is qualitative and all other school systems in Australia and possibly internationally have more rigorous quantitative measures?

In my opinion, if the world's largest brands are not satisfied to use a qualitative system to help them improve their business (because they know that statistical evidence is paramount), then why would the Queensland education system use a qualitative system to rate the all-important foundational knowledge and skills of current and future generations?

Recommendation: Employ professional statisticians to be a permanent part of the school assessment management team. They should be familiar with internationally accepted analytical (statistical) assessment to ensure school marking, and calibration of those results, provides equity for Queensland students.

Yours sincerely,
Mary-Ann Parlato

[redacted]