SMC&PA Submission 217
Received: 13 May 2013

SCHOOL ASSESSMENT GUIDE FOR DUMMIES

(or for intelligent people who want to know more)

This whimsical document has a serious purpose to explain assessment concepts to a person without an educational or assessment background. All people deserve to engage in this critical debate. it also illustrates the usefulness of mathematics.

*There are, not one but, **three** fundamental things compounding problems with Queensland school assessment. In Queensland, student results in mathematics and mathematical-based sciences (physics and chemistry):*

- *are **not** checked with common statewide exams, unlike other states; Instead large parts of assessment*
- *are judged by extensive **English**, i.e., around 2000 words written essays; and, worse still,*
- *are graded by **avoiding the use of number marks***

The use of normal marks is just good record-keeping. The use instead by our curriculum board of matching student answers with paragraphs inside many checkboxes, yet not keeping a numerical record of what was right or wrong is pivotal here. It is clear in maths, physics and chemistry that the answers are cut-and-dried and need to be marked accurately.

Yet, most of the issues of this inquiry also apply to all subjects tested in Queensland and even to primary school where the same principles of assessment are being rolled out right now in the form of 'standards matrices', which will invite the same problems. If a child gets many answers in history correct, they should be tallied up. It is silly to avoid marks and instead give the child a profile sheet full of Ds and Cs just because the information they got right was defined as lowly 'statements' and were not accompanied by 'explanations about the reasonableness of their own answers' .(That was an example of similar A standard requirements by the QSA in other subjects).

- ❖Responsible states and countries **use data**, especially to check school results
- ❖Data, put simply, is **keeping records**.
- ❖**Every** issue of importance in modern society is counted or examined with **data** (*)

➢ (*) With the exception of Queensland. Queensland is the ONLY state that does not use or examine the data of student school results, which is possibly why Qld is the only state that cannot give an ATAR (pre-tertiary rank) to our kids. It needs a mathematical kick.

- Yet, education is **the biggest** portfolio in Queensland, counting employees
- Buckets of money to employ more people to reform poor outcomes has a record of failing and will not fix this.
- Instead, removing the roots of the problem alluded to here and replacing them with objective exams and good data book-keeping would do a world of good. It is backed by overseas findings.

To confirm, collecting data does even MORE than that count things, like how many correct answers in a test. It gives a picture, it **compares** things.

- Counting the **number** of car accidents over Easter *compared* with previous years, found a sudden jump up in 2012, so road-safety actions in 2013 is responsible government.

- Recording the **percentage** of disabled people there are *out of* the total community numbers, to decide on how many disabled car parks in shops, is a fair thing to do.

- This is using **maths** to analyse stuff to **compare** AND to make wise **decisions**

❖But some people, quite rightly, think that life is about more spiritual or emotional things that cannot be measured. On a superficial level you might agree with them. But that doesn't mean we don't need numbers for maths and science.

❖All experts in any area of life use maths to keep data. Any elected representative that conducts a system that serves the public must keep data and analyse such records. But what about less concrete things? Religion, education and relationships? The Australian Bureau of Statistics can tell us how many people practice religion and if it is going up or down. Recording data can show how many people live alone, and how education level is associated with that or their income.

❖Psychologists and assessment experts conduct studies on REASONS people quit work, HOW or WHY employers discriminate against some races, WHAT makes us happy, and even how people CHEAT to achieve certain objectives. Surely, those interpretations are subjective? Yes, but by collecting many responses and scoring them with numbers, the judgements can be analysed with number-crunching.

❖All good boards (psychologists' boards, medical boards, etc) review studies and surveys that have been tested over and over to look for trends. Even questions that ask, "Are you happy at work?: Choose one answer: Very happy, happy, OK, sad or very sad." This information is scored with numbers, eg 5, 4, 3, 2, 1. Then add, "Where do you work?" It doesn't matter if thousands or millions of people answer, those **numbers** will add up perfectly, do a bundle of athletic tricks for us and will give an accurate picture of a subjective thingamabob that you don't normally associate with maths or science.

- Here are some **hypothetical** survey results from the above survey: Many different scores came in but they simply got all the numbers and divided by the number of people and got an average for the whole survey. The whole population in the survey scored 3 marks on average, so Australians are OK on average (remember a 3 was a score for feeling OK).

- But 90% of those who work in gardens, building, mining or shops scored a 5. So, they were very happy and should have pumped the happiness results up. I wonder what jobs scored so poorly that it brought the average back down? Do we need to check? What if 90% of teachers are unhappy? Will it affect our schoolchildren? What if we don't know the true percentage of teachers who are deeply unhappy **because many have quit already**?

- Notice the use of numbers to report on **qualitative** things and even predict qualitative things, not just maths or science.

  So, you measure stuff with numbers, add them accurately and do other things (some people call this manipulating the numbers but it means a good accurate thing) and get the picture. However, the reverse is not possible. You cannot record a fluffy letter for each feeling and add them up. If you put down A for very happy, B for happy, C for OK , D for sad and E for very sad you will have nothing to work with once you go past about 3 respondents. Imagine 50 pages full of letters - how do you see the trends? Which ones are most common? You can't tell without laboriously counting hundreds or more of each of the letters up, which is of stupid of course when you have numbers to do the job instantly.

  That is why we use maths and science based on maths. We can even measure fluffy things that could affect our lives in a big way. Our scientific skills are also used every day when we adjust our own paracetamol dosage depending on the severity of our headaches. A working application of physics helps our mechanics keep our cars on the road. But just how important is it to use maths to keep data on school results when Queensland refuses to do so? Answer: using maths to mark grades is as fundamental as it is to use maths for accountable analysis of every other thing.

- For example, all over the world, we know that only about 5% (*) of kids really enjoy eating broccoli

  (*) I made that figure up. But a legion of dietiticians, and common knowledge, would probably confirm I'm close).

- Favourite Dessert Survey - **hypothetical – to show how numbers reveal things**

However, a recent hypothetical survey in Queensland said that 90% of 200 children surveyed were found to prefer ‘Krafty’ frozen broccoli over ‘Yum’ chocolate ice-blocks. An assessment expert (familiar with statistics and/or psychometrics) was brought in because, intuitively, this was a majorly flawed result! They found:

- All the respondents were concentrated in a particular region. The survey was obviously not wide enough or random enough.

- By separating the answers into subgroups of people whose families worked at the local ‘Krafty’ broccoli factory and those kids whose parents worked elsewhere, an interesting trend was noticed. That is:

- Almost every single response (not every but practically all of them) that gave the nod to broccoli over ice-cream came from kids whose parents worked at 'Krafty' broccoli factory.

- To further confirm the results, the other 10% that said they preferred ice-cream - as is expected of kids - were indeed found to be made up of kids whose parents did *not* work at the broccoli factory, except a few.

- Those few children would be either true broccoli lovers (and they also inexplicably like baths and chores) or had an uncle at the broccoli factory. This could be checked with good data questions also and seeing what percentage of responses matched up to what sub-groups.

- Another completely anonymous survey was conducted and it was found that **85% of parents of broccoli-lovers** helped fill out the forms **for their children**. No surprise there but it was nice that the figures confirmed that trend.

- And of those parents: 50% said they were too embarrased to give a reason but the other 50% of parents admitted that the reason they did this was to seemingly,naively uplift the popularity of broccoli because their factory jobs depended on it. Little did they know that by helping their kids fill out the forms to say how fantastic broccoli was, the end result for all children in Queensland was 24-hour ice-cream advertisements, which led to poorer nutrition choices overall and the broccoli factory went bankrupt and the Queensland farmers supplying it lost out. Humorous story. But what is the point?

▪ Use of data: If all dieticians and policy-makers followed this flawed study without analyzing the data mathematically with attention to factors being tested, then poor decisions about TV advertising might be made. That is: " If every kid prefers broccoli anyway, then it's not harmful to allow 24-hour, round-the-clock ice-cream adverts on a kids' TV channel as they love veges more anyway"!

❖ So, using data (counting and comparing, etc) reveals trends and allows better decision-making that might even benefit thousands of people. The original survey can be adjusted to be more accountable. NOTE: We found out that the percentage of kids that liked broccoli better than ice-cream (90%) was a majorly confounded result. That did NOT mean that the children or parents should be lectured on keeping truthful diaries. That also did NOT mean that the process of number-crunching was bad. Quite the opposite: by adding up the results and checking the percentage in each sub-group, some really interesting trends were found. The data was incredibly useful when numbers were used

and analysed.

- Let's contrast that with the way the curriculum board here (the QSA) talks about education of children in Queensland. It insists that student's tests or tasks be recorded with alphabet letters and forces teachers to match those letters to some funny criteria-standards that have been thrown out of a few states in the U.S. Extrapolating that process to a joke situation: Had an interviewer from the hypothetical 'Health Board' left some Favourite Dessert survey forms in people's houses and came back later with results for the Health Department, they might have said this:

"I gave an A standard for every child who chose broccoli, and could cook it, and had excellent hand-writing (especially if typed **and** presented with **various** technologies, for e.g. Excel Wizard spreadsheets). I gave a B for broccoli and bad-handwriting. I gave a C standard for those that liked ice-cream with good-handwriting and a D for those who liked ice-cream and also had bad hand-writing." (So, the presentation was an equal criterion with the food and the A to E standards cut-offs were based on apples and oranges instead of counting what the heck needed to be counted.) "Eureka!", the Health Board official shouted, "Look at all the As for nutrition of children in Queensland!"

In Queensland, there seems to be utter disregard for using data (funnily that's maths) in checking individual students' school results, even for maths.

In addressing the terms of referenc this submission is about validity and reliability. But this inquiry is not just dealing with the maths and mathematical-based science subjects by themselves. That is because checking to see if their assessment is doing the right thing, will boil down to seeing that the raft of complaints stem from the ***blanket avoidance of using data*** *(again, that's the usefulness of maths) as normal book-keeping in* ***all*** *subjects in Queensland. Record-keeping in the form of transparent results from a child's test-paper and how those results compare to someone from another corner of the state simply cannot be carried out in Queensland, even though it's accepted practice in other states. For example:*

If a child gets 90% at 'EASY COLLEGE' and another child gets only 40% at "HARDER COLLEGE", even if the numbers are different, they are easily scaled up or down when someone compares the degree of difficulty of two tests. Even better is if everyone in the state gets the same test - then the two children will get the same indesputable results and no scaling need be done. This is already happening at Hubbards School. It uses 100% external exams that are the only 'secret' statewide common exams. Every other test or task in Queensland schools is comparing apples with oranges. You can try moderating them by looking at them but even if you wanted to scale one grade up and another one down, you can't becuase they are reported on with only 5 course-grade letters. Basically every kid gets only 1 or 2 or 3 or 4 or 5.

The QSA's claim that it is OK for schools to mark the tests with A+ or A- is contradicting their own argument and showing their near complete ignorance about mathematics.

*Numbers are not just used to count objects with, they are used to position things on a scale. Assigning positives and negatives to **letters** is simply an attempt to expand the letters by using a numerical scale that works out to 15 marks out of a possible 15 (A+, A, A-, B+, B, B-, C+, C, C-, D+, D, D-, E+, E, E-). Everyone knows that A+ is higher than A, in turn higher than an A-. The wondrous irony here is that this is a numerical scale! Thus eliminating the need to use letters in the first place. This is just one of the flaws that the QSA is defending and it is revealing. QSA cannot see that they are completely contradicting their own arguments when they defend the use of this scale when they got rid of numerical marks in the first place due to a misplaced distrust of numbers. Furthermore, the damage they are doing to the minds of children is unfathomable due to them receiving Cs and Ds (that stand for failure) when they have answers completely or partly correct.*

Then, to add another layer of confoundedness that applies to all subjects marked with letters: These, in turn, pigeon-hole children's results into the failure-end of the report card even though the answers were right or partly right. It is abusing the child's achievements in each question by changing 'correct' to 'wrong' with a D. This is a flaw of such monumental proportions that experts in the field of statistics think it is from another planet.

Yet, data-keeping should be carried out for all subjects as it is done in places around the world. The validity and reliability of testing our students' high-stakes school results - that can alter a child's life - is critical. It's comparable to the bread and filling in the definition of a sandwich. Without them you don't have a sandwich. You might have some salt and pepper, or a dash of sauce, but you cannot define that as sandwich by any measure. In turn, without a set of data, the twin pillars of validity and reliability cannot be checked by any assessment specialist's standards.

*But wait, there's more. This is not just an anomaly compared with other states and countries. This is a glaring anomaly compared with mostly anything we observe in modern society! Picture this: How many burgers did the new teen kitchen-hand ("Slow-but-steady") make at the hamburger joint? If it was a low number, shouldn't we check on how many of his burgers had the requisite 5 ingredients compared to the other new employee "Sloppy Sam", who might have made more burgers, but which completely missed some ingredients. You have to count not only all the burgers but compare what's **inside** the burgers so you can compare your employees. Likewise, the cab rank of Queensland's school graduates in the form of O.P. scores is meaningless if you cannot check what is inside their internal school results before getting to that figure.*

*Everyone's gut feeling is that the above hamburger skills assessment is fair. And EVERYONE, even people who hated maths at school, have a good gut feeling that it is a duty to teach maths to every citizen… but why? Well, it's essential to each individual (not living in a cave) to manage shopping, wages and taxes for a start. But there is something far, far bigger; The irony of the current inquiry into the lack of maths to check maths is that we teach maths and quantitative sciences in **order to use** those mathematical-based methods **in every day life**, not just to go to university, to measure and compare things. This is including using those methods*

on a massive scale during good governance itself.

Think payroll scheme analysis, think road network planning, elderly care distribution, number of houses available for the number of needy people. Data analysis is not just a responsibility of a representative of the people; each citizen has a right to know how to look at the data for themselves and participate in decision-making. This doesn't even touch the sides of maths and science used in general business, industry or medicine. The use of maths and sciences is surprisingly important in so many aspects of life.

GLOSSARY: for some light relief and a little useful information

Anomaly
– a complete oddity not normally seen.
You buy a tool set expecting a tape measure, screwdrivers, wrenches, etc and find cornflakes inside instead. No matter how you squish them together to look like tools, they cannot be used to measure or fix your car. An anomaly is not just a bit different to the usual, it is so unusual that it stands out like the proverbial.

Data
– the collection of information, most often with numbers.
Hence, the term number-crunching to interpret things in a useful way; perhaps to see if the cornflakes can be crunched down into a smaller packet that is useful for packing out the dent in your side-door, seeing as the Tool-kit Standards Board allowed tool-kits to be sold containing cornflakes instead of the real tools that the average person understands are needed to fix cars.

Quantitative
is using numbers (see the 'quantity' bit in the word). "There are 20 rats."

Qualitative
. This is compared with quantitative things (carefully check for the letter 'l' instead of 'nt' and see the word 'quality' in this word): "Gee, the weather is good today" is a qualitative answer to perhaps a test or survey question. Doesn't mean it's a 'good quality' answer, just something that this answer is about a **thing**, not a number.
So, a qualitative answer is important (the world is not just a series of digits) but, hey, just because someone said it, does this answer really mean that for the average person, it is good weather on a chosen day?
Aha! This is where quantitative data often trumps them all by giving useful information on just about anything in life…. By using numbers and analysing those numbers we found that on March 24th this year:
"
As of 2.45pm, 2951 homes and businesses were without power, with the majority (also numbers) of the affected homes in Logan.
During the severe weather, power to 58,000 homes and businesses was disrupted.
Energex recorded winds of more than 90km/h and 16,000 lightning strikes.
More than 360 fallen power lines have been reported to Energex. (news.com.au)"

There is no doubt, that the numbers can be used to paint a useful picture. The weather was indeed terrible for many people in the dark in homes or working, whether the boss or employees. Good for storm-chasers, that's about it.

** Please see over*

Signed: A concerned parent

Please **withhold my name and contact details to protect my child from any undue effects due the current subjective nature of assessment in Queensland schools.*