SMC&PA Submission 175
Received 13 May 2013

To the Education and Innovation Committee

Dear Committee Members,

I am a Secondary teacher with B.A(Maths UQ)'B.ED and Grad Dip qualifications.
I have taught for Ed Queensland for over 30 years and worked as a Maths B panelist for at least 10 years.
I have also acted in the role of Maths HOD for over 6 months on several occasions.

Sincerely,
Anthony Lewis

Under the Terms of Reference Three – VALIDITY AND RELIABILITY

SUMMARY

Concrete evidence is provided here that shows the QSA avoids conventional methods, is 'loose with the truth', and does not support valid and reliable assessment. Under the headings of:

Term of Reference Three – Reliability, and ,
Term of Reference Three – Validity,

these points will be made:

- Evidence, with **references**, that the QSA's claims that they happily permit teachers to use numerical marking is false
- Explanation of how lack of numerical marking, in turn, prevents a reliable assessment system,
- Crucial passages that shows disregard for valid assessment of knowledge.

Recommendations are listed at the end.

Introduction

I wish to elaborate on how the QSA assessment methods do not support **validity and reliablity** as part of Term of Reference Three.

> -The QSA body has failed to convince that its assessment management is ensuring our schools are **validly** measuring the expected standards of knowledge that the common man on the street would expect, let alone what students need for further trades or challenging tertiary study. No robust evidence can be found to support either valid and reliable assessment in their advisory documents.
>
> -By way of damning and contradictory evidence provided to the inquiry and elsewhere, the QSA shows also that it **cannot be trusted** to provide teachers with a **reliable system** of marking, nor can it be trusted to **provide consistent advice** to teachers **and the public** about its approach. This throws into question both its system and its function.

QSA advice differs from that in parliament and proves its approach is unreliable

The following extracts are from perhaps the most up-to-date publication you may read on QSA assessment on their website. It is a document, dated **March 2013** from esteemed State panel chairs of the QSA. It is called the **State Review Panel Reports** (from) 2012 (**March 2013**)

http://www.qsa.qld.edu.au/downloads/senior/assess_snr_rpt_panel_2013.pdf
(Accessed May 6, 2013)

Comparisons will be made showing conflicting testimonial provided by the QSA to parliament in the same month on the 20th March, 2013.

Why this document is critical

This "State Review Panel Reports" document shows how teachers have been advised by powerful state panel chairs to avoid numerical marking. State panel chairs are on top of the moderation process and provide a summary like this every year. This was summary advice given to teachers throughout last year of 2012 and serves as a definitive guide for this year.

The social moderation process in Queensland is extremely unusual. It relies on adherence to follow policy about what to bring to moderation. At the same time, the student's actual achievement in the nuts-and-bolts (the content) of the subjects is subordinated. That is because QSA has **no subject-specific external exams** to statistically check any common statewide subject results, as used in other jurisdictions nationally or internationally. They are completely lacking.

Therefore, in Queensland, human judgment is the key. So, instructions to teachers on how to mark their own test results, and what matters in moderation meetings, are absolutely pivotal. What is said in documents to teachers, available to the public, exposes the obvious mistruths stated by the QSA in parliament, whereby the QSA sought to deflect concerns and place blame on teachers.

Term of Reference three – Reliability

On the 20th March, 2013, the QSA made breath-taking claims. On one hand, the QSA attempted to convince the committee that their unsubstantiated methods are reliable while, on the other hand, denying that they have ever discouraged teachers from using conventional numerical marking used in other states and countries. In the documents published on the QSA website, detailed further below, you will see that teachers have, in fact, been advised to avoid using numerical marks, albeit in many different ways. Therefore, by default, teachers are forced to use the QSA experimental system utilising multiple alphabet letters, which has no scientific evidence for reliability in any way. This is far from the rosy picture painted under parliamentary privilege, where the QSA State Panel Chair for Mathematics C (p 2) stated on 20th March:

> *"I would like to say there is no rule about how you collect that information. So if you wish to use a marking scheme with numbers, feel free. When the work arrives at the panels for review, it will not matter whether the school used that or something similar. It will not matter."*
>
> - Bevan Penrose, QSA State Panel Chair for Mathematics C.
> from page.6 of the Public Briefing – Inquiry into the assessment of maths, chemistry and physics in Queensland Schools – Transcript of Proceedings 20th March 2013) Education and Innovation Committee, Qld Government. Compare with documentation detailed later.

It should be noted here that the three esteemed QSA spokespersons chosen by the director to explain the QSA system as one of them did in the above quote on 20th March

> *"are three of the 500 chairs of the state and district panels who coordinate the work of over 4000 QSA panellists responsible for reviewing other teachers' assessment of student work, to ensure it matches the requirement of the syllabus – that is, the moderation process."*
>
> – Patrea Walton, Director of the QSA.
> Public Briefing – Inquiry into the assessment of maths, chemistry and physics in Queensland Schools – Transcript of Proceedings 20th March 2013) Education and

Innovation Committee, Qld Government

Why is avoiding numerical marking, the cornerstone of QSA's approach, so wrong?

The evidence that the QSA has been disingenuous in parliament about how they actually instruct teachers on avoiding marks will be provided in detail further on. One cannot overstate the importance of disallowing numerical weighting to test papers and also to the composition of many test papers by the students. The students have no way to proportion their efforts to each task or test. There simply is no 40% for this task and 60% for the final exam without numbers. This is virtually unheard of in other states and places around the world. Donnelly K, 2013, "External Beats Internal. Online Opinion).

In addition to standing out like a sore thumb from other states, QSA refusal to having a robust statistically analysed exam system is simply out of touch and ethically wrong, from the point of view of known commercial and government international testing standards

Some references for international standards illustrating the need for summing up numerical marks:

- Hanushek, E., & Woessman, L., (2010). *The Economics of International Differences in Educational Achievement*. Discussion Paper No. 4925*. Hoover Institution, Stanford University, University of Munich, Ifo Institute, CESifo and IZA;*
- Kipp Charter Schools (2013*) What you Measure Matters: http://www.kipp.org/reportcard;*
- Clarke M (2012) *Measuring Learning: How effective student assessment systems can help achievement for all. world Bank Education Advisory Service in conjunction with Russia Education Aid for Development (READ) program http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2012/03/16/000333038_20120316011745/Rendered/PDF/673460revised000110Assessment0Press.pdf;*
- Submission 58 by ACER to this inquiry

Without numerical marks, there is no reliable way to rank students, or check numbers against each other, scale harder test results up and easier ones down. The QSA is thus incapable of analyzing complex data about students that is fundamentally required of teaching. (Donnelly K 2013 External beats internal; Caldwell B, 2012).

> "*Teaching is a highly sophisticated profession that calls for a high capability to analyse complex data about students and diagnose the kind of teaching support that they need."*
>
> - Prof Caldwell. Excerpt from Courier Mail (2013) Education Review Leader Professor Brian Caldwell claims teacher quality remains key to improving student-outcomes. http://www.couriermail.com.au/news/queensland/education-review-leader-professor-brian-caldwell-claims-teacher-quality-remains-key-to-improving-student-outcomes/story-e6freoof-1226391647715

This confirms the impossibility to use a statistical analysis to check any methods used in Queensland as is normally expected in other states and countries of assessment systems. So, inherent self-reviews of the Queensland system cannot be done in a robust way. Thus, in turn, the Queensland methods are unreliable methods.

This belies what the QSA stated under parliamentary privilege on both 7th and 20th March.

In the QSA hearings under parliamentary privilege, the QSA officials hand-chosen to convey the workings of the QSA stated that teachers were indeed free to use numerical marks and that they can even bring them to the moderation panel meetings so long as the students' work matches the criteria-standards, which in turn are indeed headed up by ABCDE letters.

How this really fails to support Term of Reference three - reliability

Contradictory advice difficult to follow and is unreliable

The first problem with that is that it is extremely difficult to manage two different marking systems or have the time for it. It defeats the purpose of marking the work directly with normal marks if the QSA still expects us to be checking off the boxes in the 'standards' paragraphs of their marking grids while at the same time accumulating letters on profile sheets. Work overload leads to teachers quitting.

The second more serious problem relating to reliability is that teachers must present their graded papers to a human judge, negotiate with a human judge if their students' grades are not accepted, and then have to accept what the 'umpire' says. This, therefore, is an entirely subjective process.

So, while it may not be intentional, a teacher may be relying on the good-will of a busy moderator to adjust their students' grade back up to the grade they originally handed in and believe was deserved.

In Queensland, therefore, it is not the fixed yardstick of a state subject-specific exam that determines the recalibration of a grade. In Queensland, the umpire is a human moderator who decides if a child's grade deserves to go up or down.

With subjective judgments, the teachers rely on good-will for their students' final grades

Therefore, considering the above need to remain in favour with these umpires, because everyone is after all only human, any document that states what teachers *should* do or, conversely, what is *discouraged*, has to be taken very seriously by teachers who by law have no choice but to follow policy of the QSA, in the form of documents that instruct teachers (see policies in for example: Policy Statement: Assessment 2013 DETE.
Accessed: http://education.qld.gov.au/curriculum/framework/p-12/docs/policy-assessment.pdf; also see http://ppr.det.qld.gov.au/Pages/default.aspx

See Section 3.4. Years 11-12. "Teachers plan for assessment by complying with Queensland Studies Authority processes"; http://ppr.det.qld.gov.au/Pages/default.aspx.)

Therefore, the following is evidence of stern policy advice - that cannot be breached. Basically, teachers must avoid numerical marking. Thus, what the QSA told the committee stands in stark contrast to published instructions to teachers, and what is, in turn, used on students.

Proof that numerical marking not accepted despite QSA denials. QSA misled parliament

MATHEMATICS A

In this State Review Panel Report document, the QSA actually says that using a grading system that uses aggregated marks alone could possibly *"disadvantage students"*. That could be a warning. Just aside, "aggregated marks" is another term for numerical marking. Please notice there is a distortion of terminology.

QSA is aware that their system of avoiding the use of 'numerical marking', has made headline news locally and nationally and it seems mischievous that they use synonyms all over this document as you shall see. (http://www.platogld.com/wp-content/uploads/2012/05/Media-1.pdf ; http://www.theaustralian.com.au/news/nation/teachers-told-to-stick-to-letters/story-e6frg6nf-1225809219480

. *" Feedback from districts indicated that there still seems to be **an issue with the on-balance judgments** at the Limited Level of Achievement and the threshold Sound Level of Achievement. In some cases, grading systems that **aggregate results** and/or use arbitrary cut-offs with no explicit links to criteria and standards descriptors, may disadvantage some students. When evidence in these student responses was matched to the syllabus descriptors it often displayed C standard qualities, particularly in the Knowledge and procedures and Communication and justification criteria. In some cases, this provided sufficient evidence of an overall Sound Level of Achievement placement. This highlights the **importance of making on-balance** achievement decisions across a folio of responses by matching the qualities in the student responses and the standard descriptors in the syllabus. "*

This is again pressuring teachers to only use their unusual alphabet-letter method; remember, they have the final say on student grades, so therefore using numerical marks will be taking a risk.

MATHEMATICS B

Here the QSA avoids the use of the term 'numerical marking' and instead says 'mechanical' or' formulaic' method. But it is the same. They also state that "*An on-balance judgment*" **should** use be used.

. *"A concern about a small but significant number of level of achievement decisions arose when decisions were not made using **on-balance** judgments. Where a **mechanical or formulaic method was used** to determine standards in a criterion, there was little alignment to the specific standard descriptors. An on-balance judgment **should** be used to determine the standard awarded in each criterion."*

A question that should be asked at this point is: 'What is an 'on-balance' judgment?'

The answer was already given at the 20th March hearing by QSA to parliament: When a committee member inquired as to: *"how the teacher finds the mean mark or level for the entire paper?"*, the following answer was given by a QSA State Panel Chair for Mathematics C:

"*That is clear because the work of students does seem to fit one of those standards...I am pretty good at judging, before the test, 'Well, I guess this student is a B or an A or whatever'."*

and similarly on page 5:

"*I look at the student's work. I can tell, as I said, for most students, it is a reasonably simple process, where that student falls — or the work of that student, I should say — in that matrix"*

and, similarly on page 8:

"The teacher would look at those five results and think, 'Possibly a B standard'"

(- Excerpt from QSA parliamentary briefing to parliament on how total grades are arrived at.)

This may well be as far removed from reliable marking as you can get. It is completely unacceptable.

.

MATHEMATICS C

Again the use of numerical marking is advised against and is similar to the quote for Maths B:

> "*A concern about a small but significant number of level of achievement decisions arose when decisions were **not based on on-balanced** judgments. Where a mechanical or formulaic method was used to determine standards in a criterion, there was little alignment to the specific standard descriptors. On-balance judgments should be used to determine the grade awarded in each criterion"*

This State Panel advice to teachers, published March 2013, is authored by none other than Bevan Penrose, State Panel Chair and Wayne Stevens, Senior Education Officer, the latter gentleman many regard as the designer of the letter-grading system as he has publicised and given the final say on these marking matters.

Compare this with what was said in parliament by Bevan Penrose. Teachers are indeed told to avoid using numerical marks and instead are strongly advised to make an 'on-balance' judgment. How? Put their hands on their hearts? This is a nonsensical method and certainly not reliable grading in any way.

A moment should be taken to reflect on how, what is selectively said (and omitted) by this assessment board to parliamentarians is one thing, while what is said to teachers and then used in judging students is another.

Please keep in mind that thousands of students' grades are at the whim of, the end of the chain of, this continued, subjective wrangling.

BIOLOGY

Biology is a science subject. It has in common with physics and chemistry that it includes a range of structured categorical information that is also ideally taught and tested with methodologically sound methods. Yet, again, on page 12, the QSA directs teachers to **not** use normal marking schemes that are accepted worldwide.

This again shows that the QSA was saying wildly contradictory statements during the 7th and 20th March: **they said that they didn't have a problem with teachers using numerical marking or with bringing those to the panel**. (the latter said at the 20th March QSA briefing).

The QSA personnel are paid to run a public good, perhaps the most important and sensitive public service of all, the education of children, and it is their duty to do this ethically and effectively. Everyone from top down should know and understand the assessment approach and it should be clear in layperson's terms.

The excuse that one hand doesn't know what the other is doing does not wash: The director said that she is new to the position and therefore in some parts, passed the QSA assessment policy explanations mainly to other experienced QSA staff who, under parliamentary privilege, said the opposite of the QSA direct instructions below:

> *"Criteria sheets need to be instrument-specific and reflect the syllabus exit standards. They should **not** resemble a **marking scheme** nor should they contain descriptors that prescribe a "**quantity**" of some element. "*

Remember, there is no external examination to keep a check on this process. What the QSA says, goes. We teachers are completely dependent on staying on their good side and if they say 'numerical marking is not acceptable' (which is the gist of all these proclamations), then teachers

have no choice legally but to follow this policy.

Apart from being caught out saying the opposite to teachers as to what they say to authorities, the QSA has also stubbornly refused to take advice from experts to restore reliability to the system.

Without any numerical marking, percentage weights of various tests and tasks cannot be 'added up' towards the graduating students' final exit grade. This failure to provide an upfront composition of the worth of the components of the entire 2 year course is inexcusable. It has already been pointed out in a comprehensive review that declared compositions was needed (Matters G ACER 2008 'Realising and releasing potential: 40 years on).

Significantly, without any of these reliable numerical scores, the quality of tests from one corner of Queensland to another can vary widely without any checks or balances. The need for subject-specific exams is acute. It would be judicious to implement these as soon as possible. Without them, there is no ability to do state-wise recalibration of internal marking (Donnelly April 2013, External beats internal)

Term of Reference three - Validity

Evidence is provided below that, for the QSA, 'knowledge' in a subject is not a priority for assessing students and that, also, schools are being blamed for QSA's own invalid approaches.

Below is evidence that the QSA is:

a) Blaming schools for setting too long essays (i.e. invalid essays). It is evident that lengthy English does not test physics, so QSA, therefore, is blaming teachers for invalid methods. Yet it is the QSA itself that has set these **invalid** long exemplars.
b) Directing teachers to *not* focus on the knowledge and skills of a subject. Therefore, the QSA are not assessing what they purport to assess (ie, the knowledge and skills of a subject that most citizens expect to be tested) and thus this approach is **invalid**

A) QSA sets invalid English assignments in the sciences, yet blames schools for setting too-lengthy (invalid) assignments

Here is evidence for the QSA blaming schools both at the inquiry briefings and in the most recent documents published by them in the details below.

Blaming schools at the inquiry briefings by QSA:

> *"...there still appear to be instances where students exceed the word length. It is not the role of panels to enforce word lengths. They are primarily concerned with the way in which a student's work meets the standards in the syllabus. Schools consciously make this decision, not the Queensland Studies Authority".*
>
> – Patrea Walton, Director of the QSA.
> Public Briefing – Inquiry into the assessment of maths, chemistry and physics in Queensland Schools – Transcript of Proceedings 7th March 2013) Education and Innovation Committee, Qld Government

Blaming schools In QSA's own documents:

PHYSICS

> *"Some schools are implementing assessment instruments that are beyond what the syllabus requires in terms of length."*

To see a physics exemplar that the QSA published as an ideal 'A' grade assignment for physics, go to:

http://www.qsa.qld.edu.au/downloads/senior/snr_physics_07_as_eei_1209.pdf

This QSA assignment model is 6000 - 7000 words long. This is thesis level English writing and has no place in a Senior school physics subject. Children have been subjected to these expectations for years.

Note, some time after August last year when momentum was building to hold an inquiry and public disquiet was published in the newspapers, a caveat was added between last year and now as a badge on this exemplar that still can be seen on the QSA website.

> *"The assessment instrument and responses provided in this sample met syllabus assessment conditions current at the time when the assessment was completed."*

Despite a small, hardly noticeable announcement on 12th March, 2010, that the QSA recommended (not mandated) a word cap, this assignment nevertheless remained as is for years until just recently. A recommendation is useless because schools are free to go over it, and not only do they go over it in the game of gaining highest ranking in the state, the QSA itself does.

It must be pointed out here that a document of this massive assignment example - set by the QSA as ideal to be followed by teachers - that was downloaded from the QSA website and dated as late as 2/8/2012 **had no caveat on it**.

Yet, only after teachers repeatedly complained for years of the workload of such lengthy assignments, at expense to their own health with some leaving their jobs, has a caveat been recently added at the beginning claiming that enormous exemplar was acceptable at the time of publishing. It said it met syllabus conditions. If the syllabus implies that, then that syllabus should never have been published to suggest this kind of impossible expectation on children in the first instance.

This pivotal flawed approach – tinkering to plug gaps in an inadequate system after teachers complain for years - is here for the eyes to see, both in the inception of the cobbled-together advice given to teachers (described as a system 'in development' by ACER 2008) and the ad hoc tinkering around the edges after damage has already been done.

The sum total of concrete assessment pieces

It is law-binding policy that teachers **must** follow what the QSA advises for assessment. In the QSA's failure to provide central, reliable external exams, the assignments they provide online as ideal examples, therefore, should at least be accountable in adhering to QSA's own recommended word caps. Yet, even the QSA exemplar presented to parliament on 20th March - as a prime example of the benefits and accountability of the QSA system is was well over the word limit. That is, despite the instruction sheet asking for no more than 1000 words:

The total word count was over 2000 words

This is not by stretch the student's fault. The teacher gave a loophole (as does the QSA) and only **recommended** that two sections – the introduction and discussion – be limited to 1000 words. As it turns out, the student's introduction is about 500 words and the discussion is over 1200 words long. The student has perhaps sought to gain more 'A' standard tick-boxes by expanding the other sections also. In an accountable system an assignment that exceeds a word cap by 100% would be immediately disqualified and worth zero. (link). Again, it must be emphasised that the students

in this system are the victims and that the QSA's incapacity to set achievable goals and fixed common exams is the problem

In essence, the QSA has misled parliament with regard to the fact that the QSA has been extremely disingenuous to blame schools for setting lengthy essays, which by logic implicates schools as bein the ones setting invalid assessment. Critically, in the absence of tests with fixed answers in Queensland, the assignments published online are actually the most significant mandated assessment by the QSA. The fact that a top-ranking QSA employee, the District Review Panel Chair for Physics cannot not even weed out the grossly unfair inflated assignments, shows that the QSA is at fault for allowing what is essentially an escalating arms race.

Again, it must be pointed out, that even if this was regarded as a mere oversight, the online exemplars have stood for years. The previously described physics example, a 6000 to 7000 word essay unamended, was uploaded and left up intact until only recently in 2012 after much lobbying. This is despite the claim by QSA that it said it recommended to teachers to reduce word count in 2010.

In the absence of quality exams used in other states, QSA has provided not only poor assessment prototypes in the form of unreasonable thesis-like essays. QSA has also not addressed the fact that it is the job of this statutory body to ensure word counts are adhered to; it only made a recommendation not requirement and it only capped the words going into a few sections of the lengthy report. Finally, it has proved its own incompetence in showing to parliament an exemplar, used instead of a clear-cut examination, which in itself is majorly flawed by being approved with 'A' standards despite going over the word limit.

Schools will try to out-trump each other with sophisticated appendices or methods sections due to the QSA's failure to mandate the length of these. It is anyway difficult for students to cap their words due to the mountain of ridiculous requirements of them (stated by Professor Peter Ridd towards end of transcript on 6th March, 2013, p 18-28, http://www.parliament.qld.gov.au/documents/committees/EIC/2013/QldAssessment/trns-pb6Mar2013.pdf

Again, this suggests that under parliamentary privilege the QSA says one thing even though it has published other things under its own name, ie, it has indeed set inappropriate prototype tasks that must be followed by law by innocent teachers. This can only be rectified by an acknowledgment by outside authorities that the QSA's haphazard demands in assessing children has arisen from an unworkable and overwhelming philosophical system.

B) The sub-ordination of knowledge by the QSA and validity

The underpinning philosophies that have been trashed overseas and previously in Australian when OBE (Outcomes Based Education) was discarded, manifested clearly in a statement by a QSA appointed representative on 20th March briefing to the parliamentary committee.

The QSA District Review Panel Chair for Physics stated that knowledge was no longer important to the QSA assessment approach. Having said that, the director abruptly interrupted this QSA official and insisted that the QSA does assess for knowledge, but without demonstrating how.

His statement that the QSA does not care so much about testing (and thus teaching) **knowledge** is backed up by QSA admitting this in this same crucial QSA state panel chair publication described earlier in this paper. In this, the QSA effectively says that

1. **students should be tested on 'higher-order thinking skills**' – the 'higher-order thinking skills' framework is a discredited assessment approach that has already been thrown out of overseas states and already once before in Australia; and,
2. The QSA states also that, "***discriminating selection**, use and **presentation** of scientific data and ideas*" is important (Chemistry syllabus, p. 29)

CHEMISTRY

Under Chemistry, QSA is quoted as saying

> *"Investigative processes and Evaluating and concluding criteria had not been appropriately applied. The A standard descriptor for Investigative processes requires that data is analysed to identify* ***relationships*** *between patterns, trends, errors and anomalies. It is within the Evaluating and concluding dimension that these* ***interrelationships*** *are analysed and evaluated.*

Both of these approaches will not directly measure what the QSA director Patrea Walton claimed the QSA measures: knowledge in the subject. Quite surprisingly, the QSA has adopted a system that is similar to a one that was thrown out of all Washington schools. A senate inquiry heard that the system invasively assessed process over knowledge. The dumped Washington approach also had questionable 'standards' rubrics based on a fad philosophy to test how children '**represented'** their work and discovered '**relationships'**.

This overseas system was discredited for being invasive in testing student 'affective' behaviour and measuring the processing of technology, relationships and links between ideas instead of the facts themselves.

Note the similarity. The QSA syllabuses and moderation advice also mandate that a significant chunk of so-called 'standards' must include technology and 'representing relationships', a tiny aspect of the science… instead of 'where are the facts, the knowledge, the concepts and direct applications of that?

Indeed, Professor Matters of ACER stated in a QSA commissioned review of Senior science assessment in 2006 (Matters G ACER, 2006) concluded in advice to the QSA that rhetoric in QSA's rubrics should be amended in order to stop measuring behaviours of the student and instead focus on the features of their work. That was seven years ago.

Therefore these approaches are not valid.

The following provides further evidence that knowledge that needs to be tested may be very likely missed. That is

1. other features, such as 'technology' and 'intuition' are emphasised while important knowledge is sub-ordinated.
2. open-ended projects are advised. These leave students stranded with a limitless topic; they don't know what the teacher is looking for and they don't know how they will be judged on whatever they choose to research. This is why they desperately go over word limits, in the hope that they satisfy the pre-conceived ideas in the teacher's and QSA moderator's minds. This goes against educational and psychological assessment validity standards worldwide.

> Page 66 under Mathematics A
> *"Effective assessment packages provided opportunities for students to display evidence of* ***three principles of a balanced course****: application, technology, and initiative and complexity (syllabus, pp. 5–6).* ***Open-ended instruments*** *such as extended modelling and problem- solving tasks and reports provide the opportunity for the incorporation of these principles. "*

In conclusion, the QSA has failed its own standards. To quote them, “The A standard descriptor for *Investigative processes* requires that data is analysed to identify relationships between patterns, trends, errors and anomalies. “

With the QSA system, errors have been made repeatedly but not detected or amended. Moreover, the lack of a reliable fine-grade marking system underpinning the high-stakes OP ranking score is logically flawed, yet QSA has failed to grasp this. That is, the failure to use real data, ie, quantitative figures (i.e., numerical scoring and percentage compositions, scaled according to external exams) prevents the QSA from analysing, evaluating and justifying their own system.

The QSA staff have such an entrenched backward-looking culture that they are unable to see that their non-analytical approach, taken for granted in the real world, does not allow them to check for any trends or anomalies that might harm students, or conversely reveal which schools have helpful techniques.

Queensland is on borrowed time as this accountability measure was already recommended seven years ago and it is needed now to help any education organisation in charge to recognize where tasks are **invalid** and to analyse the appropriateness and **reliability** of various assessment methods of student grading.

For these reasons, I would recommend an **urgent directive** from this inquiry, by parliamentary legislation or other, to:

1. Reinstate numerical marking as a mandatory method at the conclusion of this inquiry. This would be essential to introduce a measure of accountability that is evidently lacking for students struggling to meet demands.
2. There is a need to immediately set achievable assignments with limited questions and fixed solutions. The foundational content should be set by subject-discipline experts with experienced classroom teachers. Such experts could be asked to amend the relevant syllabuses before the close of 2013 also.
3. Beginning no later than 1st Term, 2014, all senior students (including Year 12 students) should be advised of the worth of the tasks that they will undertake (for example, an assignment be worth no more than 5% and the remaining tests 95%). There should be no assignment in mathematics whatsoever.
4. Any assignments, whether used for learning or final subject grades should have total word count caps that are mandatory and exceeding this would incur penalties and, finally, disqualification depending on the percentage exceeded.
5. All senior maths, chemistry and physics subjects in the least should have external examinations in place by 2015 latest, perhaps replacing a final term exam.

Plentiful evidence is provided here to ask the inquiry to investigate if the QSA administration should show reason for why their positions should not be abolished. If legislation is necessary to establish a more appropriate testing authority as was set up last year in the UK, then that should be considered also. Those in put charge of setting examination papers should be mathematics, chemistry and physics university academics and senior teachers, with input from test-developers from other state boards and assessment professionals in testing companies such as ACER. The input of education academics would be a questionable option considering the inability of many such academics to further improvements over the years.