Supporting New Senior Assessment
Designing for validity in local intelligent accountability systems

A RESOURCE FOR SCHOOLS

Dr Jill Willis,
Dr Kelli McGraw,
Professor Linda Graham

This research was conducted with the support of the Queensland Department of Education and Training through an Education Horizon research grant but does not necessarily reflect the views of the Department or the Queensland Government.
Suggested citation:


Acknowledgements: Our sincere thanks and appreciation go to the teachers and students in our two partner schools. You were inspiring learning partners. We have also valued the ongoing discussions with policy officers in the Queensland Curriculum and Assessment Authority, English teachers as part of the English Teachers Association, and school leaders in various seminars while the work was in progress. The research team also thanks Kylie Andrews for research assistance, Simon Gardiner for video support and Xing Lee for web and graphic design.

Unless otherwise noted, this report is licenced under the Creative Commons Attribution- ShareAlike 4.0 International Licence.
# Table of Contents

Introduction .................................................................................................................................................. 4  
Why focus on local validity? ....................................................................................................................... 5  
Five validity inquiries ........................................................................................................................................ 6  
The five validity inquiries explained .............................................................................................................. 7  
1. What big purposes will shape your curriculum and assessment design? ...................................................... 7  
2. Is there a shared understanding of assessment language? .......................................................................... 9  
3. How equitable is the summative assessment design? ............................................................................... 11  
4. Are responsive formative pedagogies aligned to the purpose and summative assessment? ................. 13  
5. How can feedback inform assessment quality? ......................................................................................... 15  
A summary of project findings ..................................................................................................................... 17 
Everyday evidence of learner agency .......................................................................................................... 18  
Recommendations for school managers of a system of local intelligent accountability ................. 19  
a) Clarity about roles and responsibilities ................................................................................................. 20  
b) Routines for accountability processes .................................................................................................. 22  
c) Resources .................................................................................................................................................. 22  
d) Information .............................................................................................................................................. 23  
Monitoring and improving .......................................................................................................................... 23 
Conclusion .................................................................................................................................................... 24  
Glossary ......................................................................................................................................................... 25  
References ..................................................................................................................................................... 26
Introduction

The new Senior Curriculum and Assessment being introduced in Queensland from 2019 will bring substantial changes to many practices in schools. The introduction of new processes around high stakes assessment has created change to all of the subjects in the senior curriculum, with all new syllabus documents having new language, scope and functions. The changes will mean new transitions between junior and senior years of school, and senior years and pathways beyond. The changes will be evident in everyday management of school timetables and calendars, camps and school culture. Importantly also, there are changes in how teachers will design and grade assessment and the way that schools will establish the quality and validity of their assessment judgments.

As schools adjust to all of these changes, it is also important look for ways to maintain or update existing quality practices. The practice of teachers having professional conversations with peer reviewers where they can give an account of how their assessment designs are enacted with fidelity is one practice that is vital to continue. Professional conversation is the essence of intelligent accountability.

Previously teachers were able to seek formative feedback on the validity of their assessment judgments in the year prior to Year 12 through a process of inter-school social moderation organized by the Queensland Curriculum and Assessment Authority. In the new Senior Assessment and Tertiary Entrance system in Queensland, schools will now take responsibility for establishing the quality of their assessment judgments in local systems of intelligent accountability. Well before Year 12 assessment, students, parents and teachers will want to be confident that the assessment decisions that are being made in their local school are fair and valid.

This document outlines a process that can be used by teachers or schools to design their own local systems of intelligent accountability. It is based on the ideas trialed in two schools in Semester 1, 2017. The teachers and students were partners with researchers from QUT in an action research process. This process involved five inquiries into the validity of assessment through conversations with peers and expert reviewers about everyday evidence of practice gathered in day-to-day teaching.

This project focused on enhancing:

- **learner agency** - the idea of students as active decision makers in their learning and assessment processes, and
- **equity** - the full participation of all learners.

It involved year 11 students in Senior English classes. We think the process can be adopted and adapted to suit more subjects, grades and priorities.

Ultimately, we hope that this will guide you to design assessment that makes a positive difference with students to learn through assessment.
Why focus on local validity?

Validity is the overarching assessment principle that is strongly associated with the fairness of assessment. It is most commonly defined as the extent to which the inferences that are made about the outcomes from assessment tasks can be justified (Newton & Shaw, 2013). Validity is evident when assessment designs and decisions are trusted, and the assessment fits the purpose for which it was designed (Stobart, 2009).

As those who are able to observe most closely the impact of assessment outcomes, teachers need to be included as key agents in discussions about the validity of assessment (Sellar, Rutkowski, & Thompson 2017; Wiliam, 2014; Moss, 2016). Yet as validity is often written about in technical terms, teachers can feel like it is not something they have control over.

*This proposed system for the validity of local assessment is designed to support a vision for assessment that promotes student equity and the agency of teachers and students.*

Often validity is spoken about like it is an ideal or perfect state, and assessment designs seek to avoid risks to that validity. One of the problems is that when validity is described in deficit terms, teachers and systems will inevitably fall short of achieving that perfection. Perfect validity is not a realistic target as validity often relies on authentic and meaningful tasks that may not be easily replicated and *reliability* is achieved through replication. These two principles are often in tension, and like a see-saw are not able to be fulfilled to the same degree at the same time (Harlen, 2013). Also, our expectations about the validity of assessment ideally will change. As we become more innovative in assessment designs, and achieve more equitable outcomes with students, our ideas of good quality assessment processes and outcomes will also change and develop. As an ever-changing concept, validity is therefore a great topic for teacher inquiry.

This project explored validity as a dynamic process. The extent to which an assessment design is valid can be discussed when the inferences inherent in the design are made more visible. Responsive teaching means that adjustments to the formative assessment and pedagogic designs are made along the way to maintain fidelity to the overall assessment intent, and these actions can be considered and justified through peer reflection. Teachers are able to see how the lives of students are impacted by assessment inferences and decisions and take action in response to address issues of equity. This has particular significance for Queensland teachers who take responsibility for designing and grading most of the summative assessment tasks. Each design decision will shape the outcome for students and is worth talking about.

An argument for the validity of teacher assessment can be established through a practical process of inquiry into the coherence, completeness and assumptions that are inferred at each stage of the assessment process (Kane, 2016). *The following proposal for a system of local, intelligent accountability builds on the ideas of Wyatt-Smith & Klenowski (2014) who advocate for system level support for teachers to engage in local inquiries and Crooks, Kane & Cohen (1996) where validity is represented as testing the strength of the links in the chain of assessment design decisions.* This proposed system includes a process of five linked inquiries that support teachers to create an argument for the validity of their assessment designs and practice within local communities of inquiry (see figure 1).
Five validity inquiries

**Discipline team: Pre planning**
Identify which first order priorities will drive design of assessment and teaching drawing from syllabus, student needs, big ideas, pedagogic framework.

**Discipline teams: Unit planning.**
Annotate exemplars to check that there are shared understandings of the quality of key concepts being assessed. Annotate or use checklist to show validity of teaching plan.

**Discipline teams: Task design.**
Design a task that has content and construct alignment with the syllabus and also the first order purpose. Remove barriers of access through simple layout.

**Teachers: Involve students**
Link purposes to learning activities. Support students to gain control over assessment criteria through peer review, and active discussions. Collect artifacts from signature activities to share.

**Either cross discipline teams, or partner teams in another school: After assessment.**

**Creating everyday evidence**
- Create a design overview to align teaching plan and assessment
- IF – THEN - SO
- Alignment between evaluators of exemplars; Teachers prepare argument for validity/alignment of teaching plan
- Student and peer feedback on draft design. Checklist for validity.
- Student focus group interview prompted by artifacts. Review peer comments.
- Feedback reports from moderation. Identification of consequences for students & new actions

**Figure 1** Knowledge map for local system of intelligent accountability
The five validity inquiries explained

The five steps in the system reflect familiar assessment design practices for teachers, with the added opportunity for the design decisions to be evaluated and validated at each of those stages through an inquiry process. The inquiry focuses on the everyday evidence that is generated in the process. The overall validity of assessment systems occurs when each link in the chain is strengthened so that the connections can be trusted.

1. What big purposes will shape your curriculum and assessment design?

Validity occurs when there is fitness for purpose (Stobart, 2009). The challenge is that assessment is used for more than 18 educative purposes, including informing learning, certification for the student and accountability for the school, with each purpose subtly altering the learning that precedes it (Newton, 2007). It is not surprising that one of the greatest risks to the validity of assessment is the confusion of purposes (Popham, 2014). Identifying the first order or the main purpose of assessment is essential for valid and equitable assessment (Cumming, Dickson, & Webster, 2013; Wyatt-Smith & Klenowski, 2014).

Validity inquiry 1 - This inquiry invites teachers to identify a first order priority that can guide the assessment and curriculum design with their teaching team. The inquiry starts with teachers identifying the competing priorities, and then moves to teachers seeking synergies between the purposes outlined in the syllabus and instrument specific marking guides, the teachers’ knowledge of students, their development and their community, the big ideas of the discipline and the teachers’ expertise, to arrive at a theory of action or design brief.

Curriculum design is a creative and cognitively demanding process. It often occurs in the everyday/night work of teachers, and when there are not planned opportunities for shared planning, this isolation can lead to designs that miss some essential elements (Gallagher, 2016). Agreement about the collective purpose, signature evidence and learning opportunities enables a coordinated approach between teachers. It also supports teachers to know how to prioritise learning experiences when there is inevitable competition for learning time and not all of the planned learning activities can take place (Willis & Adie, 2013).

In this project the big purpose was the development of students’ critical and creative agency. Critical and creative thinking are identified in the English syllabus as underpinning concepts. Learner agency adds an additional dimension that of the opportunity for students to take action based on their thinking. The capacity of individuals to reflexively discern, deliberate and decide to take action is essential in our fast-paced, information rich daily life (Archer, 2003). It is a goal for everyone, hence the focus also on equity, or full participation.

Everyday evidence of collective agreement about the purposes of assessment - Once the first order purpose is identified it can inform the first stage in an argument for the validity of the design, where teachers document their theory of action in a paragraph that forms an introduction to their curriculum plan: ‘if’ we would like to see students’ critical and creative agency as a priority, drawing from these sources of priority... ‘then’ students need to have opportunity to learn and demonstrate these success criteria... ‘so’ the teaching priorities to support them to achieve this would include the following key learning activities... Peers can give feedback on the strength of the connections in the theory of action before the rest of the assessment or curriculum design process occurs.
This is going to sound really cliché but it's [my planning is] happening all the time in my head. I think, when I'm at the movies, when I'm reading a book - wherever I am, I'm thinking and I'm often thinking, oh this might actually work. If I can get this to do this, then this might work. Films, for example - when I see things that are a big source of information for me, especially for creative writing and for reflection and things like that. Every day. Every single day, yeah.

I've always got a unit plan. I've got a skeleton of an idea of things I'm going to do but I definitely will mould with my students, which way they're going and I will push them another way. I feel like I'm kind of steering. So it's not - I don't feel like I've got to stay on this path, otherwise people are going to miss out. I'm just going to - I'm just going to go with the flow.
2. Is there a shared understanding of assessment language?

To have a robust conversation about the validity of an assessment design, teachers need a shared metalanguage and an understanding of how assessment principles are contested or negotiated in practice. While definitions and checklists go some way to establishing an understanding of technical ideas of validity, assessment literacies are fully developed through contextualised examples and reflection on assessment experiences and outcomes (Willis, Adie, Klenowski, 2013; De Luca et al., 2016).

Social moderation has been a well-established process for teachers to critically examine a range of assessment designs and offer feedback that would then inform their own assessment literacy, in face-to-face or online forums (Adie, 2013). Online assessment modules prepared by the QCAA are another emerging approach to support the development of a shared assessment language. Engaging in moderation conversations about annotated example scripts prior to planning is another approach (Willis & Adie, 2014). In this project, the teachers were prompted by a checklist of assessment terms to consider how their teaching plans demonstrated these qualities.

*Validity inquiry 2* - Teachers were asked to create an argument to show that their curriculum and assessment plans had content and construct validity, by drawing a reviewer’s attention to how the signature activities linked back to the overall theory of action. In an iterative process, they checked back as they planned to make sure there were sufficient opportunities for learning about the content and skills required to be factored into the learning plans. This checking for alignment occurred with peers, as well as with external expert reviewers who may be peers in another department, another school, a professional association, or in this project a local university. Feedback from the reviewers enabled the links to be confirmed, questions or new possibilities to be considered, or gaps to be identified and addressed before teaching commenced.

*Everyday evidence of shared assessment language* – In one school the teachers used track change comments to highlight in their curriculum plan where they designed teaching and learning opportunities that serve the first order purposes of learning. In another school, they used a checklist and presented their argument within that checklist. The teaching plans, assessment task and argument were emailed to a reviewer for feedback before the commencement of teaching.

Feedback from the reviewers included commentary about where there was evidence of alignment and where there may be greater opportunities for validity. You can [view these at the project website](#).
The teachers valued comments from teachers on the reviewer feedback saying “it gave us a lot to think about...from different perspectives”, it showed “we needed to be clearer...[and] where we needed to go into further depth”. It was also helpful for the teachers to see where they didn’t agree with the reviewer, or realized how much their time or freedom to take a different approach was constrained as “there’s things we can’t do”.

**So validity** - I guess the closest thing that I can think of is authenticity or authentic, that it’s a valid judgment, that it’s an authentic judgment, that the judgment has integrity, that it is true,

We end up teaching a lot of life skills, whether we like it or not, particularly because of the text - so that idea, is this valid and useful I think they sort of link together

**Critical agency** - I definitely understand it as their ability to be empowered by making choices that are sometimes challenging - yeah - and not wanting me to make those choices for them. Yes and I’m going to back their decisions.

I’d given them quotations on diabolism, so we’ve looked at the theme, and I gave them the quotations but I didn’t want to do the analysis for them, even though they want me to. Because I really don’t like to read my own words and essays that I’m marking. I think with looking at that, in critical agency, if they’re just regurgitating what you’re saying, it’s not very helpful to them because they don’t actually learn anything from it... I was jumping around with the girls and just asking them obnoxious questions, like, are you sure, really, really. Just to question them and make them actually think beyond, I’ve written this now, so it’s right.

**Creative agency** - Having that ability to do things differently, to create differently and to make things differently. Also I guess that even a teacher agency to have sort of different creative abilities that you can bring to students and to a staffroom.

When I think validity, I think a student is going to ask me why, why do I have to do this? If I can’t say to them, this is a text that we see in real life, let’s go do it. It’s something we’ve created. They’ll be just like, why am I doing this? So I think validity - it has to have relevance to the students’ lives and to our world, true, that it’s meaningful, yeah.
3. How equitable is the summative assessment design?

By being aware of the criteria against which they will be judged, teachers hope students will grow their capacity for self-assessment and active participation. Thus, in their ideal state, these task sheets allow students to engage in self-evaluative judgments as they check their developing work against the specifications. Lots of detailed instructions and information may inadvertently create other barriers for students to fully participate in the assessment performance.

Assessment task sheets are invitations to students to demonstrate their mastery of new understandings, however these invitations must be evaluated in terms of the affordances they provide for different learners, and whether learners can transform those affordances into action (Gee, 2008). It involves critical inquiry so that we do not blame students for not achieving when it is the clarity of the invitation that is at issue (Torrance, 2017). This validity inquiry focuses on whether students can access the instructions to know what the task requires. Those instructions will often occur via the assessment task sheet, but also are delivered verbally in daily classroom interactions. For many students with high incidence language or attention issues, or have other barriers to ‘getting’ the task, the equitable design of the task is an urgent issue that needs to be addressed.

Equity in assessment is about reducing barriers that lead to direct or indirect discrimination so that all students have the opportunity to “demonstrate their abilities and what they know and can do”, so that they can be active participants in creating a “socially just, equitable and democratic global society” (Queensland Government, 2010, p. 12). Attending to assessment accessibility issues to enable full access is an Australian federally legislated requirement for reasonable adjustment (Australian Government, 2005) yet there is a need for some more practical guidance for teachers (Graham, et al, in press). Recommendations for designing accessible assessment need to consider not only accessible design elements, but also the experiences of learners, which are central to the definition of fair assessment (Klenowski, 2014).

So, how do we make assessment equitable? One answer is to focus on the first-order expectations of assessment (the core knowledge the task is designed to assess) and to eliminate “wherever possible” second-order expectations that unduly complicate the task (Cumming, Dickson, & Webster, 2013).

Validity inquiry 3- Teachers were asked to draw on their knowledge of the students and likely barriers they might face with the summative task, and to design support into their curriculum plans. Where possible, they sought feedback from students on the clarity of the draft versions of the task sheets for assignments, or afterwards, to highlight areas that seemed important and also unclear or confusing. The teachers also used an accessible design checklist to argue for the equity of their design. The teachers were easily able to adjust and align their task designs with positive outcomes for students, and greater clarity for their own teaching.

Everyday evidence of equitable summative design

Evidence for the validity of the teachers’ design can be found in the annotated teaching plan where barriers have been identified, supportive inclusive strategies included, or student’s strengths have been built on. Additionally the checklists that the teachers used to self and peer review their summative designs, provide evidence about the links between the intent and the design. Student feedback about the task sheets, and critical review of student outcomes is a further source of evidence for consequential validity.
When using the validity checklist...We sat and went, why is that even on there? So many years of tasks that involved - you've got dregs from times before. We're like, that's not even that task anymore, quick delete that. That's just, again, endemic of schools. We just don't get that collegial time. We work really well in a space; we're like, right you do this, I'll do this, and get it done. But I think our whole faculty should be doing it because I think - yeah, it's just finding the time where it's not going to be people feeling like they're being disrespected for the lack of time they have. “

“We want to do the sheets you gave us, the tick sheets, on our units for next year when we get time up our sleeve at the end of the year. Because we found it really good to cut back some of the rubbish out of those task sheets and just make them just so obvious what the task actually is.”

“We had huge success with the checklist and I think a little bit of a culture shift will come now in our school. In that staff won’t think that looking at the reliability and validity of your assessment and making sure it is inclusive isn’t a criticism. It’s an important part of education.”

“There has got to be a partnership between teacher and student that you can ask them: did that help and that they can give you an answer and just turn around and say, no.”

“I think a lot of it comes back to... knowing your students and knowing what they are capable of and what their needs are and having those relationships with them so that they know that they can achieve what you want them to achieve and that you are doing the best by them.”

“So they’re responding in all different ways that are quite amazing... I had some going “Can I challenge that? I want to do character types, the saints, the sinners and the ones who reflect the audience?” I'm like, “go for gold. What have we got to lose?””

“They chose all of the characters that they were doing. They chose who and what they were going to say, so I think that made it really helpful. I did split them by - it was all mixed ability, so that it wasn’t just all the students in this group who do really well, and all the students in this group who need a little bit of a hand. It was so that they could actually help each other and make sure that they were getting different ideas.”

“A student emailed around the expert criteria group notes from different groups...Someone I totally didn’t think would have looked at that, did. He put his hand up and he said, “Miss, I actually looked at that and thought where am I not doing well according to those highlighters and that’s what I spent time on over the weekend”. I just about fell over because in the past he has not seemed to want to do those things.”

“There are a couple of kids I know who were achieving threshold Cs and I see them starting to hit Bs in some of their criteria now. Because their understanding of what they’ve got to be talking about or addressing in their story, how that transformation is supposed to mirror that text, is just that little bit closer in their mind.”
4. Are responsive formative pedagogies aligned to the purpose and summative assessment?

With fewer summative assessment tasks in the new senior paradigm, formative assessment has become even more significant as an opportunity to check the meanings that students are making through the learning activities. Teachers are then able to adjust their teaching through observing how students engage in formative tasks and give specific advice to students to improve their work while it is in progress.

Classroom activities are formative when they inform and form how students come to know the expectations for quality performance. Assessment for Learning, or formative assessment approaches are well established as enhancing students as active and agentic learners while providing teachers information to guide their teaching (Cowie & Harrison 2016; Charteris & Thomas, 2017; Willis, 2011). Through dialogue students hear how others think, and through scaffolded activities around key ideas, skills or assessment criteria, students can assume greater responsibility for the quality of their learning.

Formative pedagogies can be time consuming to enact, so their deliberate design into the teaching plan, and their alignment to the big purpose and summative task is an essential step in the argument for validity. It is also important that teachers plan to teach students the skills needed for these formative tasks, and establish a classroom climate that supports conversation and productive collaboration.

Validity inquiry 4- This inquiry occurs both in the planning stage, and throughout the teaching. In the planning stage, the teachers design key or signature activities that will enable the students to self-assess and self-regulate the quality of their work. In this project activities included students giving peer feedback on paragraphs in online forums, or creating an order out of disordered paragraphs before justifying their reasoning. Discussion through interactive games, mock trials or Socratic seminars required students to lead discussions, and draw on evidence to support a point of view. The teachers took short video recordings or photos of some of these key events and these were used as prompts to ask students for feedback on the usefulness of these activities.

Everyday evidence from responsive formative pedagogies—

The teachers highlighted these formative activities in their teaching plan and used colour-coding or track change comments to link the signature activities to the big purpose. This enabled peer reviewers to affirm or make suggestions about these links. Teachers then prioritized these activities when there was time pressure and not all planned activities could occur. The students provided ongoing feedback to their teacher through their discussions, but also more formally through a focus group discussion. A neutral interviewer showed students the artifacts that were generated through the formative activities and asked how these activities may have helped their learning, and what might be done to improve them. This feedback was de-identified and summarized before being given to the teachers for reflection and improvement.

If valid and ethical assessment design includes building on student strengths, and

If engaging in formative feedback with students is necessary for teachers to understand learner progress, and for learners to assume more responsibility for their learning,

then inevitably a curriculum plan will be adjusted as student learning needs and ideas emerge.
Teacher reflections:
It took a little bit longer than I had planned initially to get them to get all their evidence together and start writing up the scripts. But because they had to ask each other questions, and they had to engage, it did mean that they were saying, “well if we say this, what about this”. Then other students were saying, “well I’ve already written this, so can we put that in there”. So it really allowed them to just look at it from - and get other people’s perspectives, as well as those students who were listening as well.

I think they asked a lot more questions, than they have before...sometimes I’d gone in “we’ve got to do this today, and we’ve got to do this, and we’re going to do” - off the plan, and then we’d just throw it out the window, because they’d be asking really good questions. Like, “you know how we were talking yesterday about that thing”, or “about when we do this with our introduction, or when we do this with this quotation, can we talk more about this element, or that element of it?”.

It was a creative way for the students to really think outside of the written element of theme and a shape of something so different to think about...It got them very excited and encouraged them to really think in a different way, once they thought about it in terms of a shape, they found themselves finding a quote so much quicker. I didn’t give them any help on the quotes. Sometimes I would say things like ooh, maybe you want to have a look at a different one. But I never told them which quotes to choose. They so easily responded to the idea of a shape representing a theme. Students that are - that challenged themselves a bit more, chose more abstract shapes. I think that that was really interesting and exciting to see the different - the diversity within that seemingly small activity.

Instead of us, like, well me, just saying, here’s you know, you’re A, essentially, or here’s a - the model - it was more actually, what’s wrong with it because they picked that out anyway. But to have them sort of sit down in groups and work together and cut up the paragraph and then rearrange it, and they all had quite different paragraphs in the end, which was interesting. I sort of got them to explain where they put them to each other, to the different group, and we got them to explain where they put it, why they put it that way... it’s allowed them to be a little bit more successful going into an unseen exam. That they have some sort of structure in the back of their head, that they can actually manipulate it.

This will probably sound like I’m oblivious but I actually had no idea what the criteria meant until it was explained to me this term. Or I knew what it meant from knowing the words but I never had really applied it or had it explained to me how it correlates to my essay. To have her constantly relating it back in class and really just having it up there it really helped me.

Peer assessment just gave me a deeper understanding of the criteria and helped me to aim towards my goal. It also helped me recognise quality, even in other people's work, so I could therefore recognise it in my own; seeing what I’m really good at, what I’m missing, so I can therefore work on it.

That activity ...about criteria 1, which is responding to the task question or something like that, which I'm not good at and I can’t find it in my own work. So that’s finding like point sentences and link sentences and that kind of stuff...it just confused me more to be honest.

After doing this I have actually thought to myself, oh I need to understand my other subjects’ criteria really easily because all the criteria has changed now from junior to senior, kind of. So I’m confident now that I understand what the English criteria is asking of me, it’s because we’ve done these really narrow looking at the criteria and going these, all these activities to say this is what you’ve got to do.

It gives you that bit of confidence if other people are having the same ideas or writing the same way as you.

We know we got better...Because at the start we were all just writing, like it was mainly an explanation. But towards the end we got to see our improvement [in our online discussion forum] and our last ones were a lot more structured than our first ones...and our teacher’s writing just kept getting smaller and smaller. She didn’t have to write as much.

I think if I had a wish I’d ask for more one-on-one conferencing. I know that’s kind of hard, because there is like 25 kids in a class, but there is some - yeah I just find that really helpful, like talk through your idea.

Student Feedback:
As long as the teachers explain how this relates to the unit and how it’s going to help with the type of assessment that’s going to be done I think that’s also good.

Normally in class in previous terms and years we’ve really just been listening. I’m a visual learner so just looking at something I’m able to get - with the paragraphs I was able to really once I’ve seen a paragraph I can break it down and understand it. For some students who really need to do it this in visual and kinesthetic ways it really helped. A few of my friends were saying that as well.
5. How can feedback inform assessment quality?

Intelligent accountability occurs through processes of building trust and strengthening collective responsibility for students to achieve valued outcomes (Sahlberg, 2010). Onora O’Neill, a philosopher who focuses on trust and accountability in public life explains in her TED talk that trust increases when we give each other adequate, useful and simple evidence, drawn from our day to day practice, that we are trustworthy in our professional decisions. It is most effective when it occurs through processes of self-governance within a framework of reporting, when professionals give an account of what they have done to those who “have sufficient time and experience to assess the evidence and report on it” (O’Neill, 2002. p. 58). The accountability system therefore needs to include evidence from everyday practice, opportunities for shared dialogue about that evidence, and a framework for reporting that evidence.

As assessment occurs all through learning and teaching activities, and the design decisions are adjusted and adapted along the way in response to the teacher observing the consequences of design choices, these five inquiries all include an element of intelligent accountability. In each of the five validity inquiries, teachers are invited to articulate their intentions and share the everyday evidence with others. In some cases this is with peers, at other times with students. Teachers were then able to reflect on the feedback that either affirmed or challenged their practice, enabling them to adapt it. In this way the teachers became more aware of the consequences of their design decisions. It takes time for teachers to develop confidence and familiarity and to think critically about the quality of assessment (Black et al. 2010), so this process is an ongoing commitment to professional learning dialogue.

Validity inquiry 5: As well as the cycles of feedback throughout the teaching time, it is important to engage in a critical reflection on the overall outcome. Events such as regular social moderation where peers can discuss the outcomes of their work, and seek feedback on whether their judgments are supported by the evidence, are well-established activities in Queensland secondary schools. It requires a commitment to a shared focus, shared time and coordination between networks. It is not only important to confirm assessment decisions. Teachers need to also inquire about the consequences of the assessment for the students. Who succeeded through these learning activities? What informed their success so it can be built on or repeated for future students? Importantly, who did not succeed, and what might be done in response? Finally you can decide how you might share this account, to continue to build trust in your professional decisions.

Everyday evidence – The teachers in this project found that student feedback on their learning activities gave them immediate feedback on the usefulness of activities. Formal student feedback was highly valued and could be shared with supervisors, parents or students. Teachers often use checklists or feedback forms to give each other advice about the reliability of their judgment making, or their assessment designs. These documents are everyday evidence to review and inform new actions. A simple summary of those ideas then provide the next link in the chain for future learning. Ongoing reflective conversations can also provide evidence for the validity of the assessment design.
Accountability came through reflective conversation – How do I know my assessment designs helped students learn?

- It’s allowed them to be a little bit more successful going into an unseen exam.
- Allowed them to get that wider scope of what was going on.
- Their practice paragraphs got better.
- They asked a lot more questions, than they have before.
- It allowed them to sort of engage with ideas they hadn’t engaged with before,
- They were able to manipulate it [unseen question in exam rather than regurgitate a memorised answer]
- Some really out there quotations
- I have seen some of them come up quite well in their grades, and particularly for an unseen exam
- Some of their extrapolations at the end of their essays to sort of the [modern day] were really good.
- There was not as much of me hearing myself as I’ve had in the past.
- Students having that confidence and that knowledge
- [They were keeping each other] accountable [for the quality of their evidence and argument] - making sure that they weren’t just relying on me.

Feedback from reviewers was really helpful…It was interesting for me thinking about the restrictions that we had, because we went into it feeling - well I went into it feeling free and oh, we’ve got all this time and we can do all these things. Then the term turned out the way it was and then we got this and I thought oh, wait a minute, there's things that we can't change and there's things that we can't do. I think also the audience questions were interesting.

The kids keep themselves really accountable like “no, that's not true, you made that up. You can't say that”. Or they sort of - because they have that opportunity to come back at each other they, ask “you know, but could you consider it in another light?” That allows them to think and to extend their ideas and extend their knowledge in an interesting way I think. I think it really allows them to sort of get outside of like “I have a thesis now, I have to do this”, which gives them that broader context to come into the exam and then use all those analytical essay writing skills. Just harness it.
A summary of project findings

In this research, designing for learner agency and equitable outcomes were identified as first-order priorities to anchor explorations of task validity. Assessment design was seen as an ongoing process where summative assessment tasks, curriculum plans and formative pedagogic tasks were aligned and adjusted in response to peer, researcher and student feedback. Learner agency encapsulated the idea of students as active decision makers in their learning and assessment processes. Equity was seen as the full participation of all learners.

Finding 1: Student-focused planning included responsive decision making
Seeking feedback on the alignment between curriculum plans and summative tasks, and providing evidence of that alignment to reviewers was new work for the teachers who valued the opportunity to refine their thinking. Teachers created evidence of their responsive planning as increasingly the student questions drove learning and discussions highlighted areas for further teaching. Where students found it ‘scary’ to be in charge of discussions, or to experiment with some of the writing ‘recipes’ they had learned in junior school the teachers spent more time developing trust to help students take these risks. All teachers had to decide ‘what needed to be sacrificed’ in their teaching plans as there was not enough time to cover the pre-planned content as well as the responsively planned material.

Finding 2: Increased focus on critical agency enhanced other elements of learner agency
There was evidence of enhanced agency from students through their participation in class activities, and also the quality of their summative results. Through giving peer feedback and other collaborative activities in class, students shared their different perspectives and their thinking behind their assessment work. As students heard and read a variety of quality answers, there were greater opportunities for uncertainty to be shared and ideas to be tested. Flexibility in how to use the same evidence in different ways helped students feel prepared for unseen exam questions, and also for workplace discussions. Students’ summative assessment had evidence of more original ideas and ‘less regurgitation’, and greater skill in analysis and argument, suggesting a relationship between critical and creative agency.

Finding 3: Increased critical agency in one subject area had positive effects in other subjects
For many of the students it was the first time a teacher had explicitly taught them what the assessment criteria looked like using multiple examples. Students independently used skills they had learned to self-regulate their work. Students then transferred the skills they had learned in English about understanding criteria to other subjects, and advocated for supported peer feedback in other classes.

Finding 4: Prioritising accessibility in assessment design lead to enhanced equity
The teachers designed for equity in task scaffolding through including peer support, mixed ability groups, multiple modes of engagement through a mixture of visual, auditory and hands on tasks. Using a checklist about the accessibility of the summative task sheet was a “huge success” leading to a culture shift in one school that “making sure it is inclusive isn't a criticism”. Consultations with students identified areas for ongoing research as different students approached their assessment by drawing on different information in the task sheets and activities.

Finding 5: Feedback loops enhance assessment literacy of teachers
The five inquiries became a chain of decisions that generated evidence for teachers to argue for the validity of their work. Feedback from students, researchers and colleagues was seen as essential information that enabled them to “step outside of the silos” of their work and have their ideas questioned, confirmed and challenged.
Everyday evidence of learner agency

Suggesting alternative ideas for assessment responses.
Transferring the skills independently and clarifying issues with the teacher.
Applying concepts to other subject areas.

Moving easily between individual, peer, small group and whole group learning.
Making connections between learning activities.
Seeing multiple perspectives.
Connecting ideas to their life worlds.

Goal setting from previous feedback.
Different ways to show knowledge – visual, matching, talking.
Choice of tasks.
Mixed ability groups
Multiple opportunities to practice new skills.

Participating and being valued in the class.
Being challenged to take risks with an idea.
Peer review.
Sharing work in progress.
Individual conferencing with teacher.
Fun competition in groups.
Leading discussions

Seeing examples of high quality work and analysing those examples.
Knowing the criteria in depth.
Giving peer feedback based on criteria.

Connecting criteria to learning.
Understanding a range of quality through giving peer feedback.
Being expected to be a knowledgeable peer evaluator.
Recommendations for school managers of a system of local intelligent accountability

Validity relies heavily on human judgment (Crooks, Kane & Cohen, 1996). Therefore the management of this system needs to be based on processes where the trustworthiness of human judgment can be established.

Often systems of accountability seek to achieve fairness or validity by emphasising conformity of practice and auditing through comparisons of large data sets that attempt to control the work of teachers at a distance (Biesta, 2009). As school assessment is increasingly used as a governance tool for measuring quality in education, the system is becoming algorithmic, corporate and opaque, and far removed from the human actors who complete the assessment (Williamson, 2016). Schools are increasingly being compared through centralised information which has led to positive outcomes such as coordinated support, yet it has also led to concerning outcomes where accountability pressures having significant side effects on the experiences of teachers and students (Bloxham et al., 2015; Spina, 2016). Teachers can feel pressured to narrow their teaching to prepare students for assessment with never enough time to do all that they ‘should’, with accompanying feelings of guilt if they pursue values outside of those that are easily assessed (Thompson & Cook, 2017; Ball, 2003). Principals and teachers need to be able to respond to the complexity of local contexts without being hamstrung by an over-reliance on accountability driven only by high stakes assessment data (Gable & Lingard, 2016). The following proposed system reflects some of the emerging ‘soft’ governance principles for complex systems (Burns, 2016) that offer alternatives to top down systems.

A greater focus on processes of dialogue and capacity building can create feedback that informs the system and enables the system to respond to complex and changing events (Synder, 2013). Wilkoszewski and Sundby (2016) identify four key governance principles. There needs to be a) clarity about the roles and responsibilities within the system b) routines for how accountability will be established c) clarity about resources and a d) flow of information to promote transparency. These can help school leaders and managers to enact a local system of accountability.
a) Clarity about roles and responsibilities

When undertaking an inquiry into the validity of assessment, multiple stakeholders can potentially be engaged. It is important to establish what roles and responsibilities are fixed versus negotiable in a local system. Teachers, middle managers, and school leaders may all have capacity to contribute to decisions about elements such as the purpose and shared language of an assessment.

As the system depends on regular opportunities for dialogue, clarity about the various roles that each person may play helps with the coordination and the expectation that there can be a shared responsibility and accountability for quality. These roles would be a focus of ongoing negotiation when the limitations to or pressures on this decision-making emerge, for example, when first-order priorities for inquiry/change are set at the school or system level.

- **Individual teachers** necessarily play the main role as assessment designers. Many of these roles are evident in the Australian Professional Standards for teachers. Some of the activities include;
  - Being open to sharing ideas and receiving peer feedback.
  - Negotiating and ‘buying in’ to the first-order priorities for curriculum change.
  - Annotating a unit plan or creating a theory of action (if-then-so) that reflects the overall design intent. This alignment is often an unspoken understanding by a teacher. By creating the design brief, there is opportunity for collegial feedback and suggestions.
  - Reflecting on the strengths and also the challenges that students have in learning, and designing opportunities to remove barriers to learning so all students can access the requirements.
  - Collecting samples of student work to review with students, and with colleagues.
  - Continuing to link the purpose of learning activities to the big purposes of the unit.
  - Unpacking the assessment criteria in with multiple activities.
  - Setting up opportunities for peer review and discussion so students can learn from others and teachers can hear and see where students are at in their learning.
  - Adjusting curriculum plans in response to where students are at in their learning.
  - Seeking out peer and student review.
  - Sharing the evidence of learning with others, including students and parents to create trust in the assessment design and decisions.

- **Students** play a valuable role in providing feedback via formal and informal methods. In schools where formal opportunities are given to students to respond to and critique assessment design, clear advice can be systematically collected. However, students also provide informal feedback in the course of everyday learning, which can be captured as evidence by teachers for subsequent professional discussion.

More formal involvement in providing feedback to teachers about learning and assessment activities created evidence that was highly valued by teachers. Before participating in a more formal feedback situation students would need to be reassured that their comments would jeopardize their relationships or results. In this project that ethical commitment occurred by seeking parents and student permission, de-identifying all comments, and having an independent person conduct and transcribe the focus groups interviews before giving a summary back to the teachers.
• **Middle managers** – Department-level leadership was recognized in this project as vital in establishing common purpose and shared assessment language in this project. Heads of Department and Teaching and learning coordinators are well positioned to integrate planning and review processes into the existing curriculum planning processes of schools. This can include coordinating:

  o Review of student outcomes to inform new ways of addressing equity barriers,
  o Collaborative curriculum planning time.
  o Requiring evidence of alignment such as annotated plans or checklists to argue for the validity of designs prior to teaching.
  o Reviewing summative task designs for accessibility.
  o External peer review for formative feedback on assessment designs and outcomes.
  o Calibration activities around student evidence such as annotated examples of student scripts.
  o Encouraging cross-disciplinary peer review.
  o Supporting teachers to try new formative assessment ideas.
  o Processes for recording lessons for teacher and student review.
  o Focus groups with students to review experiences.
  o Sharing evidence of student outcomes in terms of agency and equity with teachers, students and parents.
  o Resolving issues and ideas that emerge from practice through an ongoing process of learning and inquiry.

• **Interviewer of students** – If student feedback is to be sought and valued, an independent person would ideally conduct and transcribe the focus groups interviews before giving a de-identified summary back to the teachers. This person would understand their role was to inquire rather than judge. They would need to commit to principles of privacy and confidentiality. In research projects this is managed by the signing of a confidentiality agreement.

• **School leaders** are instrumental in providing opportunities for collegial sharing and reflection, which can take place within the local system environment, or in wider professional learning networks. This can involve

  o Promoting partnerships with professional associations, local universities or employers, and cluster networks for teachers to share ideas between schools.
  o Coordination of common times and a shared focus for local sharing possibly during pupil free planning.
  o Development of local guidelines governing the confidentiality of discussions.
  o Awareness that the established practices around formative moderation or confirmation of teacher judgment that are already successful provide a good basis to build on.

• **Reviewers** – The reviewer’s role is to check whether they can follow the alignment and reasoning of the assessment design, and ask questions and make suggestions to prompt for greater alignment. Reviewers can include colleagues within the same subject, year level, department, school, cluster, system, professional association, or as with this project, your local university. Ideally there is a reciprocal benefit for reviewers, as peer feedback can lead to learning for the reviewer and writer. Often it is helpful to have an outside facilitator to lead a review conversation.
b) Routines for accountability processes

Ideally the processes of moderating or validating assessment decisions through local networks between schools and also seeking feedback from students will become part of the regular curriculum planning routines of the school. This process would not necessarily occur for every assessment task, but at planned stages across year levels and teams. Conditions for success would include coordination of times for professional conversations within departments in schools or between schools, and a culture of supportive peer conversation and feedback.

Harlen (2007) suggests “providing samples of work and showing how certain aspects relate to the criteria of assessment” as one form of quality assurance. Importantly the focus of this practice “is on the process of arriving at that outcome” (p. 78). Annotating assessment samples prior to teaching as a calibration activity (Adie & Willis, 2016) is another form of sharing quality work. This project suggests also that sharing a theory of action, or creating an argument for validity based on the formative pedagogies and student feedback is an additional form of evidence.

The accountability system therefore needs to include evidence from everyday practice, opportunities for shared dialogue about that evidence, and a framework for reporting that evidence. There would need to be agreement about how the evidence would be used as there are different outcomes when evidence remains the property of the teacher for learning purposes, or when it might be considered evidence of performance for supervisory purposes. There are also resource implications for the time to share and receive feedback and act on it.

c) Resources

Videos of classroom formative activities or artifacts of student work are a helpful way to prompt a purposeful review conversation with peers or with students. The videos can be shared, rewatched, and discussed. Resource support for creating and selecting short video snippets would need to be easily accessible. Media release policies for capturing images in classrooms and an understanding of ethical storage and how the videos can be shared would need to be negotiated in advance.

When planning to interview student focus groups, costs for transcription need to be included in a budget alongside small audio recorders. Administrative support to coordinate knowledge sharing and recording ideas arising from evaluation would also be a recommended resource.

Professional learning for teachers to have open to learning conversations may also need to be included in the resources. Most significantly the project teachers noted that time was the greatest resource and lack of time the greatest barrier both to enacting the intended curriculum and in co-planning with colleagues.

We just don’t get that collegial time. We work really well in a space; we’re like, right you do this, I’ll do this, and get it done. But I think our whole faculty should be doing it because I think - yeah, it’s just finding the time where it’s not going to be people feeling like they’re being disrespected for the lack of time they have. It is a really nice idea. We did use it well at the end of last year. So I’m hoping that we just sort of start keeping these little things we want to do for our teacher preparation day, as a little list...It’s just getting other people on board with it pretty much.

Time was definitely a barrier. We lost a lot of time, particularly at the start of the term. Not just with public holidays, but students had sporting commitments and music commitments and those sorts of things. So it was trying to make sure that everyone was on the same page, or around the same page at least. It was quite difficult, so there were some things where we just like, well, we'll just have to cut that, or we'll have to use that in a small part, or we'll just shape that differently. I found a little bit that I had to go back over their knowledge about aesthetic features, a little bit more than I expected to.
**d) Information**

One of the reasons for investing in local systems of intelligent accountability is to continue to build the trust in the assessment designs and decisions that are being made. Communicating evidence about the validity of assessment work, and the outcome of the ongoing inquiries is therefore important. Find ways to share with parents and students the reason why these inquiries are occurring. A key message of the process can be one of shared responsibility for learning - that everyone learns here.

Equally important is the consideration of how that information about validity will be shared. Teachers and students can be involved in the careful selection of evidence, so that trust starts with those whose work is being discussed. If there are regular updates based on student feedback and teacher or school responses to that feedback that were shared through channels like the newsletter or teacher features, this helps to build trust within the school community. Regular evaluations from external partners can provide alternative views.

Sharing everyday evidence between schools would need to be maintained as a commitment to professional learning or a form of soft governance, as top down pressures focusing on a quick fix and measurement in an audit society can erode the teacher professionalism that is required to enact equitable and valid classroom assessment in local contexts (Mills & McGregor, 2016).

**Monitoring and improving**

This recommended process is a formative one that is designed for the purposes of informing improvement. A risk is that it could easily become a system of accountability to an historical ideal of validity rather than be a continuously evolving process of enhancing student equity and student and teacher agency. To address this risk, ongoing formative and formal evaluation of the system is essential and this can be through reflections by the teachers and students at regular intervals.

To ensure consequential validity, there needs to be a regular critical review of the consequences for students to identify how assessment has supported their learning and why, and which students have not experienced success and why, and what might be changed. Student and parent perspectives are essential data in these critical conversations. From these critical reviews, there needs to be feedback into the system and a search for evidence of change.

The system needs to be adaptable for local conditions and disruptive innovations. As digital technologies disrupt how assessment is conceived there also needs to be room for quite radical innovation such as adaptive testing, online social moderation, and increasing use of analytics and auto feedback (Cope & Kalantzis, 2016). This room for innovation can be achieved by investing in small groups of people who are willing to try out new ideas and share them with the wider teaching team.

Sharing knowledge through learning networks requires an investment of energy and time, as knowledge networks are unstable (Schildkamp, et al., 2017). Building the capacity of champions and leaders to help them access alternative ways of thinking is an important part of leading this new work (Cherney & Head, 2011). We suggest that studying for a Masters of Education or a Graduate Certificate of Education is one way to engage in advanced professional learning to extend your understandings of equity and inclusion or learn how strategic thinking can create alignment within a dynamic and complex system. At QUT many of our Master of Education students study part time and externally while they work and we enjoy learning with and from them.
Conclusion

In times of rapid change the sheer number of changes can feel overwhelming for students, parents, teachers and leaders. A lot of the energy and time can be spent on working out what needs to be done, and the ‘why’ it happens that way can get a bit lost in the noise. This resource is designed to support Queensland teachers and schools to keep students at the center of the assessment and learning story. It is a proposal on which you can build your own system of intelligent accountability. In your school there may be more or fewer than five inquiries. In the case studies with schools, often the first three inquiries were simultaneous.

Queensland teachers enjoy a reputation around the world for excellence in designing and evaluating learning through summative assessment that has local validity and authenticity (Darling-Hammond, Wilhoit & Pittenger, 2014). We also know there is more to do to increase the equitable outcomes for all students (Ehrich, et al., 2015; Thurlow & Kopriva, 2015) and that a high degree of trust in teacher judgment is integral not only to day-to-day teaching in the classroom, but also to accountability systems (Heritage, 2013; Stobart, 2015). This research aims to support schools and district leaders to continue to invest in this ongoing learning with teachers and students.

The research team had great pleasure in collaborating with the teachers from the two schools in this study. We enjoyed learning from them and with them and seeing the high degree of trust that was evident between them and their students. This resource is intended to assist the ongoing work of quality teaching and assessing in schools in Queensland. Our hope is expressed in the encouragement by Professor Gordon Stobart, that continuous development in Queensland will lead to ongoing development in teacher assessment practices around the world.

Queensland has something that is really remarkable – a system of high stakes certification in which teachers are at the centre and where school based assessment is trusted as a way of arriving at a reliable decision about student achievement. You will be well aware that elsewhere in the world teachers would not be trusted like that, and that we rely heavily on external exams. In some other countries teacher assessment does play a part but it is a minor part because nobody is sure about standardising the teachers, the quality of the assessment and the like. So Queensland has something remarkable. To maintain this we need a process of continuous development and continuous professional development. So my message is keep up the good work Queensland. There are many of us watching you as a model of what can be done. There will be many teachers across the world who would be amazed that teachers can be given this responsibility. We look to you.

Emeritus Professor Gordon Stobart
Institute of Education, University of London.
September 1, 2015
Introducing a QCAA video for teachers

https://www.youtube.com/watch?v=jV2kbU37Kk
### Glossary

These can be read to complement the definitions from the Queensland Curriculum and Assessment Authority [https://www.qcaa.qld.edu.au/k-12-policies/student-assessment/understanding-assessment](https://www.qcaa.qld.edu.au/k-12-policies/student-assessment/understanding-assessment)

<table>
<thead>
<tr>
<th>Term</th>
<th>Meaning</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>‘conditions of access’ dictate additional, implicit skills known as ‘criteria of access’ that students must apply to successfully complete the task.</td>
<td>Burbules, Lord &amp; Sherman (1982)</td>
</tr>
<tr>
<td>Agency</td>
<td>to have opportunity to take action – drawing on past patterns and habits, practical judgement in the present and an imagined future.</td>
<td>Emirbayer &amp; Mische (2009).</td>
</tr>
<tr>
<td>Creative agency</td>
<td>to create an innovative intervention in the world that draws on one’s interests, experiences and aspirations.</td>
<td>Cope &amp; Kalantzis (2009)</td>
</tr>
<tr>
<td>Critical agency</td>
<td>to analyse oneself and others to interrogate and understand interests, motives and relationships of power.</td>
<td>Cope &amp; Kalantzis (2009)</td>
</tr>
<tr>
<td>Equity</td>
<td>is evident in assessment when it is in the best interests of the student, it is non-discriminatory and enables full participation.</td>
<td>Elwood &amp; Lundy (2010)</td>
</tr>
<tr>
<td>Intelligent accountability</td>
<td>occurs through processes of building trust and strengthening collective responsibility for students to achieve valued outcomes.</td>
<td>Sahlberg (2010)</td>
</tr>
<tr>
<td>Learner agency</td>
<td>is evident when students are active decision makers in their learning and assessment processes.</td>
<td></td>
</tr>
<tr>
<td>Reliability</td>
<td>describes “the extent to which the results can be said to be of acceptable consistency or accuracy for a particular use”.</td>
<td>Harlen, (2013, p. 9)</td>
</tr>
<tr>
<td>Validity</td>
<td>is evident when assessment designs and decisions are trusted, and the assessment fits the purpose for which it was designed.</td>
<td>Stobart (2009)</td>
</tr>
<tr>
<td>Validity - Content</td>
<td>refers to the content that is the focus of the assessment and how aligned and representative the focus of the assessment task is with the required skills and knowledge. This can include whether it represents the big ideas of the discipline, the syllabus requirements, and the alignment with what has been taught to students.</td>
<td></td>
</tr>
<tr>
<td>Validity - Construct</td>
<td>refers to the design of the task and whether it gathers evidence about what it was intended to assess, and enables students to create that evidence within the conditions of the task. This can include attending to the accessibility in the design, alignment of all of the elements, achievability of the scope of the task, and the mode being fit for purpose. It is closely linked to the principles of reliability and equity.</td>
<td></td>
</tr>
<tr>
<td>Validity - Consequential</td>
<td>refers to extent to which the assessment achieves the purposes for which it was intended for all students. It is closely linked to principles of equity, as assessment can inadvertently have different consequences for different groups of students. The inquiry is whether the values and ideologies represented in the task, and the potential and social consequences are equitable (Messick, 1989 p. 20).</td>
<td>Elwood &amp; Lundy (2010) Messick (1989) Stobart (2009)</td>
</tr>
</tbody>
</table>
References


Popham, W. J. (2014). The right test for the wrong reason: The tests we use to evaluate student achievement may well be sound measures of what students know, but, at best, they are faulty indicators of how well they have been taught. *Phi Delta Kappan*, 96(1), 46.


Willis J. & Adie L, (2013) Negotiating the meaning of achievement standards in the Australian curriculum, *Curriculum Perspectives 33*(1), 52-62


