I’m going to start with a famous phrase that you probably remember hearing ‘Be alert but not alarmed’. Well after hearing John this morning I think we should be alert and be a bit alarmed as well. I think that we really need to think carefully about the way TEQSA could go and we have to think of ways of managing ourselves with our colleagues in all other universities to try to make sure that what we are on about which are the things that are on that board over there are not undone by processes which systematically make it difficult to achieve those ends but for other ends which are somewhat spurious and I know the government’s interest is in terms of trying to guarantee the quality of our universities and I just want to explain the position that I’ve taken in the last few years.

One of the things that really has bothered me for decades actually is the situation where there is some degree of uncertainty in some social system such as the education system or the higher education system and the government says we have to do something and it grasps at something which looks attractive and perhaps affordable and that it implements that only to find out 10 years later that it didn’t achieve the end you wanted it to. And I suspect that’s what would happen if there’s any kind of attempt at comprehensive testing. And the reason I say that it’s not because I fear it simply on its own account but nowhere where standardised testing has been tried in the world has it lead to school improvement. Improvement in the student’s performance. It just does not do that and we’ve got the same sort of thing happening in various sectors of education. So that’s one thing to be wary of. The other thing to be wary of is some kind of atomistic approach to the programs and courses that we teach which try to decompose everything into minute detail and then say this is what’s got to happen for that to be the case of higher education. The kinds of things that we have on that board are not the kinds of things that can be written down in minute detail, have a checklist formed and say if you can get a tick in all the boxes you can do these things. They’re not like that. They’re categorically not like that but neither is a whole lot of other professional training generally.

At the moment I am a little bit involved with the Australian Medical Council and they are facing a situation where their courses in medicine used to be six years long and there was some sort of gravity attached or gravitas attached to the six year length of it and it sort of gave doctors a bit of an image that’s saying, well we’re the highest trained professionals there are. We do six long years and then we do an internship. So we have a natural kind of, the right to say things with patients. Now, that’s all changed in many situations to four years. That’s one thing that’s changed and the other thing that’s changed is that there are people saying time serving in terms of medical education doesn’t guarantee the outcome. So then there’s a movement to try and force medical education into competency based medical training. Now the Australian Medical Council said, well we’re not sure this is the way we want to go especially if it looks like that competent doctors are being defined as those that can exhibit 1000 competencies, or
2000 or 3000 as the case may be. So the committee that I was on, and I’m on a different working party now, said okay, how can we get our heads around what really medical education is about? I won’t go into the detail of that but we ended up producing a discussion paper. It took us a year with a very productive committee. I must say it was a wonderful committee to work with because the people were committed to trying to sort out in their brains what it is really about and one of the words that we did champion was ‘competence’ but we were very clear that competence did not equal the sum of 1000 competencies. So what is the extra? And then we have a big discussion about that and we produced a consultation paper which has been sent to a number of professions and which has now got interest from other professions to say, that’s what we’re on about too. That’s what we’re on about. My view is that’s what higher education is on about. It’s producing people who can think, who can do complex things, above all they can make good qualitative judgements in circumstances that they have not encountered before. They are educated to a higher level. It might be in a profession. It might be in a discipline. It might be in all sorts of areas. To think that that could be undone by some kind of template that tries to classify things minutely is to me, seems to be a mistake of the first order. The problem with it is it would be expensive to do, probably not very effective in a short time term and wholly ineffective in the long term. So I say be alert and be somewhat alarmed.

The address that I gave last year at the first of these symposiums had a pretty well similar title to this one. I want to talk today about the fuller picture. I want to indicate to you the purpose of trying to talk this through. I write articles as many of you know. Most of the things which are seen as innovations in terms of teaching and assessment are based on something that somebody dreams up and tries and seems to make a difference. I have no problem with that, except that some of the things which are dreamed up and tried and seem to make a difference are actually counterproductive in other ways that people don’t realise. So my question is, can we in fact have our feet on the ground, make real decisions, real designs for quality assurance systems for grading that are theoretically sound, practically sound, practically doable but also put the control and leave the control with the academics who make the decisions. So my interest has been decision making on grading is distributed throughout the universities. Every course lecturer makes decisions about grades, about marking student work, about the quality of student work. Everyone does that. What I say is that’s the site at which the quality assurance should be applied. Basically, a point where those decisions are made. We need to be able to find ways to make good decisions. After those decisions are made it’s too late. We can put all sorts of layers of things over the top – it won’t change those decisions. So I’ve been in the process of trying to think through these issues and I must say that I have wrestled with a whole heap of the issues that are involved and as I go around different universities and talk to people I’m really attracted by their questions because the questions are usually searching questions and I had a couple this morning which have made me think. They’re really important questions to deal with and I want to know, so what is their answer to this one. I don’t want to make up something. I don’t want to fabricate stuff. I want to go right back to the roots and say, what is it we’re on about? One of the things that the whole discussion about academic standards is lacking is a clear focus on what the gain is. What is the problem to be solved? Academic standards. When the Carrick Institute was first set up they had a board that commissioned me to write a paper on academic standards and the person who was interim Chair of the board, she knew the position I came from because she had apparently been in contact with some of the work that I’d done and she was in Canberra, and I had visited Canberra for that purpose. And she said, ‘We want actually to get your take on this’. And so I focussed on achievement standards. It went to the Board and the Board said ‘No, this is not the way we want to go’. You know why? ‘That would be too controversial.’ So we looked at standards for inputs, room spaces, floor area, entry requirements, library facilities, IT facilities, all the things which are important inputs to people learning but not the standards, the outcomes. So what happened to my article? I was supposed to get feedback on it in the February meeting, from the February meeting of the board and then work and then produce the final version in June. That got squashed at the end of February. And the reason I found out what happened at the board was that two people who were concerned came to me and said ‘For goodness sake, keep pursuing the directions you’ve got in that paper because eventually that’s the direction we have to go’, and these were people who had been basically, I mean they didn’t win the vote, put that way in the board. And that’s a story out of town. But it’s interesting that there’s been this resistance to focussing on achievement standards and I think that’s the thing that matters. So what about achievement standards? I think it would be a good thing if we could put our hands on our hearts and say the grades that we award to students in their courses and that we write on their transcripts mean what they say. A high distinction means really excellent work. A pass means a pass, and not everything left over after a credit. And a credit means in between and a distinction means in between that and the other one, okay?
So they actually grade it. But that's what we want to have, those grades mean something. Do they mean anything now? I visited some universities where I didn't hear the word achievement. I tried to prime sentences with that. They didn't take it up at all. What I did hear is 'passing', passing, passing. We have to get our students to pass. The one case there and Linda was talking about students who have their grades sort of built up by things which are non-achievements and when I talked about that in one university they said 'If we stripped all of that out, our students wouldn't pass'. I said 'Why?' 'Because they wouldn't accumulate enough marks to get over the threshold to get a pass'. And I said 'So in your heart of hearts are you saying these people do not deserve a pass?' And they said 'Well', and it got such a long pause that I changed the subject. You see this is what they're saying, we have to do this because this is how we get our students to pass. In the event of a query we can show this chain of marks, no question about whether they're built up from things which are non-achievements but whether they're built up or not, if they're over the threshold... That's the sort of nonsense that's around. And I want to say how can we change that? So I've been trying to write articles which are as scholarly as I can make them. I don't say they're perfect but as scholarly as I can make them and trying to deal with this issue. And I coined the term 'grade integrity'. 'Grade Integrity' the article is called, And the Representation of Academic Achievement. Because I think that grades should have integrity, and there's a companion paper to that with Fidelity in the title. Fidelity as a Prerequisite of Grade Integrity in Grading Student Achievement. Now those two papers are available to anybody who wants them. But I must say they're as well argued as I was capable of putting it out because if we have practices which engage us in certain ways which eventually will run counter to TEQSA and I think we should, not because I want to be contrary but because I think the directions in which they're going, about which I'll say more in a moment, those directions are not going to be helpful. If we've got another way, we have to have the arguments ready. We can't go in there half baked and say, 'Just believe us. We do something differently and we want ours to prevail'. It will not cut. What I'm also aware of is that there's no real argument for the other sides. I've looked. So I want to actually say there's ways of doing this, with integrity, that we as academics keep control of the agenda. I know the workload is an issue. I think there are ways to manage that but I think first of all we have to figure out what is the problem to be solved. And I think the integrity of the grades we award in courses is the problem to be solved. And I find that some people agree with me and some people don't disagree with me but they think up other things of what the problem really is.

Now, in the paper that was distributed last year, it was called Assuring Academic Standards in Griffith University, there were three levels identified. This is what my thinking has been. There are some things that we do already in universities which are fairly common, not universal but fairly common. One of those is something that we can build on and that is consensus moderation of student responses to a single assessment item in a course which has lots of students and some markers. So typically what happens is this: you get a sample of student work from different students. The different markers mark them. They come together, they confer, have a discussion saying 'Oh I think that is too high', all these things of what the problem really is. The different cut points are different and we want ours to prevail'. It will not cut. What I'm also aware of is that there's no real argument for the other sides. I've looked. So I want to actually say there's ways of doing this, with integrity, that we as academics keep control of the agenda. I know the workload is an issue. I think there are ways to manage that but I think first of all we have to figure out what is the problem to be solved. When a discussion takes place I think they are agreeing on what quality consists of for that particular work. That sort of task in the student response is the thing. This is what quality means. They haven't got this in. Well you have got this in, and they have a discussion about, but they come to an agreement on what quality is. That might sound absolutely trivial but I think it's fundamental. The second thing this is that a quality is typically distributed across some continuum and the next decision is where do we chop that continuum so that we can call these by a certain name, these by another name, these by another name. Because if you have different cut points from what I do, suppose all of yours are on the high end then you're going to give fewer high marks and high grades than what I do because we both might recognise quality but we haven't yet settled on where the cut points should be. So those two things go on. Now what I'm saying is that that practice is well known to many academics. If they're not doing it in their current courses maybe they've done it in previous courses where they've been part of a teaching team. That's one thing. I think it is common but not universal. The second thing is most academics think it's a good thing to do that. They're not offended by it. They don't see it as a denial of their academic autonomy to confer with others. That's normal so let's celebrate it. Now it's that process which I think can be generalised to other stages, the process involves those two stages generally speaking. I think the process can be
generalised, can be scaled up and there’s one other thing that Nick Buys mentioned before. I must say that the foundation that Tom laid for us this morning, it was just first class in terms of what I want to say and certainly what Nick said. I mean this is really, we have to get our heads around this as a big, we don’t have to all understand the nitty gritty of it but I think you have to sort of realise that whatever we do has to have a sound basis that’s actually sound whichever way it’s looked from. Philosophical, practical, any which way. Look at it and say yes, that makes sense. It is an argument with integrity. Because I can’t find arguments at all with or without integrity or the alternatives.

Now let me talk about the first scaling up. One of the things we don’t do very often is review the work that students have produced from different assessment tasks in the one course. Suppose there’s three of us sit down, we are part of a teaching team. We don’t actually say, so what is this evidence telling us about the grade that should be awarded in this course? Now that really is the key decision we should make. What we do instead is we add up marks and we say, if the marks are over this and under that, then this is the grade that gets attached, right? So the marks are a proxy for the evidence. The moment we code things in terms, when we award marks, we actually lose sight of the evidence and we can’t actually recover the evidence from the marks. The question really is, what level of achievement is this student demonstrating? That’s really the issue. Now why do I have these reservations about marks? Because marks are on pretty well any old sort of scale you like to think of. In Norway the mark boundaries between the different grades are legislated by parliament for all universities and this is a way of getting quality assurance across the universities. And I know people who are in one university in a certain city and they have looked at the work in other universities and they do use the same scale with the same cut points but they say theirs isn’t a patch on ours. So there’s 100 reasons why marks don’t tell the full story. They can’t tell a full story. But once you code stuff you’ve lost the evidence and what we typically do with the evidence is discard it or give it back to students. I won’t thrash that one anymore but the thing is we need to look at that.

Now if we’re looking at all the student’s work, what is the variable we’re looking at? Well let’s take one step back and say when we’re looking at consensus moderation within a course for one assessment item with many markers what we are saying is, what is the quality of these responses? The quality of these responses. Okay? When we come to look at the evidence from different sources in a course we’re saying, what is the level of achievement? Different variable. What is the level of achievement that this data is telling me should be awarded? What grade should be awarded? Now that’s the fundamental judgement that I think we as academics ought to protect. It can’t be substituted by omnibus tests and because we are capable of doing that then we can also say that is the place where we have to attend to the quality assurance agenda. And then we take it one step up. You say, well that happens in my course but what about other courses? Now one of the things that happens, Nick has been a Chair of assessment boards for 10 years he said and I was a Chair of assessment boards for quite a few and what we did was, we tried to get some kind of comparability across courses by trying to get the distributions of grades broadly equivalent. So if someone rocks in with all of their students being given high distinctions we say nup, no go, and we have a look. We say, have you got the data to prove it, to demonstrate it? Ah nup, well go on back. So it’s the same pass. So then you have a sort of discussion, sometimes they get pretty tricky but you try to actually negotiate the distribution of marks. Now what we’re really saying there is it’s the distribution of marks which can take over our role of the judgements and you can have a course which is very poorly taught and the students actually achieve very little but you can distribute the marks in a way which gets passed the Assessment Board, no problem. When I was Dean I two cases where that happened and I had to go back and retrieve information to try to get to the roots of it but until I had the evidence which was the student work which I had to recall, we couldn’t get a handle on the problem. So, all sorts of things can get sort of muddled over by looking at the distribution of grades. So that won’t be a solution either because you’ve got to come back to the integrity of the decisions that we make. So that’s why I put a question mark over grades. I’m not the only one who does that. If you want to look at some of the work of Alastair Pollock he’s got some interesting history on how marks came to be. I’m not going to go there.

Now let’s have a look at cross courses. Now what the rough equalisation of grade distributions does, it gives you the semblance of comparability across courses because you make all of the course grade distributions look something the same. So it gives you the idea, oh well our courses are comparable. Well they’re not necessarily. Maybe they are, maybe they aren’t. So I say, can we look across course boundaries to see the kinds of things that other people call high quality. And people will say, well what do I know about this, what do I know about that? Nick and I have gone around a lot of schools and we
raised this very issue. The idea of crossing course boundaries and saying if we were to look at work in a field that is cognate with ours, and that means it’s broadly similar and it might be something that you’ve studied yourself for example. But do we know enough, is there something that’s kind of core enough in academic standards for us to say that is high standard work. I know it’s in that course which is not the one I teach but yeah, that is high quality stuff and I’d call that low stuff. Now when you actually come to the practice of doing it people may initially feel reluctant to engage in those conversations but when they do they find they can do more than they thought they could do. And when Nick and I have gone around various schools, we’ve found some schools where people have been eager to do that because they’ve felt isolated in their own silos and they said, I’d really like to see what happens in other courses. And that is a very fruitful place for discussions about what quality and what achievement looks like in different courses. It is doable. It doesn’t have to be done every semester with all the student work at all. That’s nonsense. It would drive us all nuts. But it is doable and what’s more, people who have tried that enjoy doing it. They think we have actually learned a lot about this. And we had a little bit of an inkling of that this morning from the policy graduate certificate. So that’s the second level of scaling up.

Now in each case we go back to the primary evidence because that’s really what counts. Now the limitations of a system based on just those three levels which I called in last year’s paper, one, two and three level. One is the consensus moderation we’re familiar with. Level two is the one that looks at achievement from different sources of evidence across a course in a course. And then the third level is across courses. Now all this is set out in the paper that you’ve got today. So what are limitations about that? Well they assume that the stimulus for the evidence was high quality. In other words they assume that the assessment tasks were well constructed. They’re not always that. I’ll give you an example in a minute. The next thing is that they don’t necessarily touch base with what the professions and the employers mean. We could be simply having a little virtuous circle within Griffith University and that won’t work either. So I’m suggesting we introduce two new brackets. In this paper I introduce those and because I’ve already called the others one, two and three, the prior one is called Level-0 and the other one is called Level-4. Easy. You start counting at zero instead of one.

Now, students respond to assessment tasks, that’s the evidence that we have. These tasks may be poorly designed and poorly specified so that the evidence that poorly designed tasks produce is defective. Not through anything that’s the fault of the students but because the tasks are bad. Let me give you an example. This is a real examination question from a real examination paper but not from Australia. I chose one not from Australia, and it says this: Comment on any three of the following in about 200 words each. Easy enough stem. These are the things: Games theory for the study of international politics and its limitations, sustainable development, basic issues in movements for gender justice and relevance of non-alignment in a post cold war era. Now can you guess what the course would be about? Guess, go on.

Current affairs. It’s in politics okay? But you see the stem. The stem says, comment on any three of the following in about 200 words each and it gives you four. Now why I picked that example because I came across, I’ve come across many university exam papers from some prestigious universities as well that are just like that. And there’s one particular professor that I challenged about it and when I’m going to say ‘question’, that’s me asking a question. When I say ‘answer’ that’s his response. My question was this: This assessment task literally just asks students to comment on some topics. It doesn’t set up any problem to be solved, issue to be addressed or question to be answered. How do students know what to do? Answer: Well, they can put different interpretations on it as they wish. There’s always a range of quality. You can’t stop that. I think it’s fine the way it is. It sorts out the sheep from the goats so to speak.

This is as close as I can remember for the real conversation I had. Question: How do the poor performers typically respond? Answer: These students just write down a lot of facts, everything they know about the topic. Question: How do the best performers respond? They analyse the issue, they probe, they evaluate and they draw conclusions ideally through reasoning. Question: Why don’t you explain in the task specifications that that is what you are looking for? Answer: Why? The best students don’t need that, they know what to do. The weaker students, they just regurgitate. In any case there’s always a good spread of marks so I can’t see any problem with this exam question.

Now I’m not joking. That’s a real exam question and those comments follow a very similar structure. Now you see, it may well be some of those students who regurgitate all the facts who could do
something different if only they’d known. My question is, how do they know what’s in your mind? Now I’ve asked some people, how do students know what’s in your mind? Another answer that I got was: Well I don’t always know until I see it in the answers. I find it very illuminating to see how they’ve handled it but I recognise that when I see it, I said that’s what I want. That’s what I want. Then I see three of them in a row and I say that’s what I really want, hey. My heart’s for the weaker students who don’t know what to do. Can’t read people’s mind and sometimes this comes because they start our programs with a lack of cultural capital that doesn’t let them access that and what’s more it’s unconscionable to say, give me this old stuff but I’m looking for something over here. That’s just not on. So that’s what I mean by the quality of the task, how it’s constructed, how it’s devised and how it is specified. It’s crucial. Now last week I was at a college in Sydney, a theology college actually, so I got some theology question papers off the web, some that I could find because a lot of them are secure, and I put these out on a sheet and I said ‘Have a look at these and say if I were a student would I know what to do?’ And they could have spent two hours. I stopped them after an hour but they just got stuck into that. You know when they were simply put into that in the cold and looked at these other questions they were able to see the folly of some of these things. It didn’t set up any problem, it didn’t create an issue, it didn’t do anything. They were just stuff, you know, inviting the students to write stuff and there are different classes, different sort of courses in theology because I don’t really know what they’re all called but I picked different ones out with different course names. I thought they’re probably different sorts of parts of theology or church history or something so I put them out. But see they didn’t have any problem looking at them but they had problems when they tried to construct them themselves because we construct them very often looking at the content we’re trying to cover without looking at what we want students to do. And if we don’t want students to do deep things and require students to do deep things and teach them what deep things are, we get nowhere near these things. Absolutely nowhere near them. And one of the things, I mean you really need to do, and I mentioned it there was, we need to teach students not how to interpret words that aren’t there and how to interpret what was in our mind and guess what will impress us, tell us what was in the lecturer’s mind when he said the thing, we have to teach students how to do it. Here’s an example: if I asked students for a critique what I would hope to get back are critiques of various quality so I can assess their quality as critiques. I do not expect to get a travel log or some other kind of thing. I expect to get a critique. Now if half the students don’t do what I ask them to do but simply write a whole lot of stuff, a lot of tendency is to say these students have done a lot of work therefore I will pass them. Mistake. What that does is reinforce that it’s okay not to attend to the order of things. So we have to teach students to attend to the higher order of things in circumstances where the outcome is not critical, ie, it is strictly formative with no summative implications. So we make certain things high stakes for learning, zero stakes for the grade. We teach differently. You might say ‘The students wouldn’t do it’. Well, I know one lecturer who says to the students ‘You can take that risk if you want but it is a risk because I’m trying to teach you something here. Let me show you those that took the risk last year. These are the pass grades they got and these are the ones who did take the risk and learn from the exercises. You see they did them and here are the pass rates for those. Radically different pass rates’. He said ‘You’re all big people, you take the risk, you work it out. I’m not here to tell what risk to take’. And he said since I’ve done that it’s changed the attitude of students. Now there’s all sorts of things we can do to get our tasks right. Duncan gave us a little questionnaire about things we want to do. But you see how important it is to get that basic task right. A good design. A good design asks students to exhibit those things. Now here’s another thing. We don’t want to create tasks where we coach the students through the answer. I say this: When does a student learn something? This is my definition. When they know or can do something they couldn’t do before. When they can do that sort of thing. That’s what assessment should be about. We’re actually looking at how capable they are, how competent they are, how proficient they are. Use whatever words suit your discipline. We’re trying to see how good they are through the learning that’s taken place. We’re not trying to coach them into one particular way of doing certain things. Okay, so that’s why we’ve got level zero.

We need to review both aspects of design and specifications as a prior stage to be a contentious thing otherwise we cannot in fact rely on the evidence because it’s poor quality evidence. Now what sort of questions do test those? Well I’ve got a few things here which I can give to anyone who wants them afterwards. Does it set up a hypothetical situation? Does it require extrapolation or the application of something to a new context? Does it require analysis, identifying components or something, working out...
how they're related? Require integration, putting things together they'd never put together before. I've asked students to do that sort of thing and they tell me 'Oh it's just great because I never really thought that those things were associated that way and now I see it'. Now how do they know that? Because I gave them a task that required them to do it. Giving them practice in doing those kinds of things. Now if you don't know what you're looking for, here's a suggestion. Bit of risk okay? Write one which is as good as you can find, that you can make up, give it to the students in a non sort of like summative way and then have a look at those which are really, really good. And then say, so what is it about those that are really, really good and maybe incorporate some of those things into your specifications next time. Now note I'm not saying, I'm telling the students the kind of answer, the category of answer that I want. I want a critique about this and this. Don't give me a critique about other things, I want a critique about these things. I do not give the students a recipe for doing it. That's their problem. I tell them what the task is, what the problem is, what the issue is. I tell them what that is so they can see what kind of response I need but I don't give them a recipe for doing it. That's what they've got to show themselves. Now that's level zero.

The next level is level four which is external calibration. We really have to go one step up and look at how would things compare across institutions. Now some people feel very threatened by that and some people in certain institutions feel very threatened by that. I think it's not something that we should be too worried about. If we were to provide certifications of achievement where the student transcript meant what they said for all universities in this country, it's hypothetical but if we're able to achieve that what it would mean is that the reputations of the universities that gave those transcripts would not matter much. Say if you get one from the University of Bulia and it is in our system, then those grades mean what they say. And if you're from the University of Melbourne they mean what they say. University of Queensland, Griffith University, any university you like. They mean what they say. If we could come to that then that is really the ideal and we can only do that by sharing work amongst universities. Now this is not a pipe dream. It is already going on. Kerri-Lee is involved in a project that's been run out of University of Western Sydney. When I was given the kind of outline for what was going on I had some suggestions. One of them was 'Why isn't there a GO8 university in there?' The response was 'There wouldn't be a GO8 who'd play ball with us on this.' So I said, 'Is there no Dean of Business in one of these universities who knows any Dean of Business in a GO8 very well'. They said 'We don't know, go find out'. If you ask Vice Chancellors they might say 'Oh no, no, we don't do that' but if you find Business Deans in this group will know some Business Deans in there, for crying out loud there's probably an Australian Association of Business Deans, is there? Yeah, there you go. They must meet. They must have drinks and tell jokes and all that. They're there. So ask them. See. So we found one. Come in. No worries, no worries. Now what are the implications of doing this? Well there's a really serious implication, it is this. That some universities will have lower proportions of high grades than other universities for a start. I say that's something we have to wear. I tell you what. There are some universities who do excellent teaching and who are not Sandstones and they are saddled with reputations that would be fixed by another 100 years of public funding as Gavin Moody says. There's nothing that they lack except history but they're good at teaching. We've really got to get over this hang up whereby the, I talked to people in GO8's about this and they're happy with that. Some of them will have lesser proportions of high grades that their reputation would lead you to think they'd have because frankly their teaching is not that good and the performance of some of their students is not that good. What's more a number of them will tell you that they have very big differences from one faculty to another. They're not all the same but they do equalise them all. And they say, we need to get over that because that's being inequitable to our own students. So, all these things are there. People are thinking about these things. On a person to person level you can discuss them. If you did it all officially through their Vice Chancellors you may not get anywhere. I think really there's scope for us as academics to be more outgoing in terms of the agenda that we have to follow, to know what it is we're on about. I don't think the practice of it is nearly as time consuming or expensive as a lot of people think but I think we have to be sure that what we're doing makes good sense. I'm not going to talk anymore about the grade integrity across institutions but what I do reject is the premise that each university should be entitled to award the same number of high distinctions as the other universities. Now I have found that argument put up to me. They're saying, but if I did that I wouldn't be able to award my students high distinctions. I would say, well even when they don't deserve them? They say, 'Well yes, because you know, well you have to actually, well, they've got to have a chance. I mean this is what equity's all about.' I said 'No it isn't, not if you want to trust, that will doom your university to always being second rate. So that means your transcripts can't be trusted. You want the transcripts to be trusted.' Now you might think I'm labouring that point but I really believe that's one way in which universities actual outcomes can be sorted out much, much better than some
kind of blanket test that comes across universities. There’s meant to be omnibus across all courses. Look, we have to preserve the variety that we have in our universities in terms of our courses and in terms of our programs. We have to know what comparability looks like. It does not mean sameness. You take the simple case of housing, I talked about this last year. You can have six houses that sell for $500,000 and they’re all different. All different. You say, are they the same? No, but they’re of comparable value. What we want to say is these outcomes are of comparable value. That will come with practice but that’s the problem that we have to solve.

Now what I’m saying to you is there is like what I believe to be a valid way of thinking this issue through. At the very pointy end we come down to how we’re going to introduce this to Griffith University. What I suggest to you is this: We don’t all have to be, get our heads across all of the scholarship that’s surrounding this issue but I think a core of us really have to, but I think the balance of us have to be careful that we don’t let any procedures run away with our labour. I think we have to find ways of doing intelligent but light sampling of certain things. I think we have to be about as rigorous in our procedures as procedures are for peer review of journal manuscripts for example. We have to be as rigorous. I’m talking about double blind. I’m talking about not knowing in advance what the range of, was talking to a person at the University of Sydney yesterday. He said to me ‘So if you’re talking about’ like threshold standards, he was talking about the ALTC project, ‘Should we have you know, only passing grades? You know, the range of passing grades?’ I said ‘No. If you’re talking about threshold that is just enough to pass. You’re getting a fix on that. What you need is a bunch of things at that borderline because that’s the point of maximum certainty. That’s where you have to resolve it. Forget about the high distinction, that’ll be obvious to everybody. Forget about the absolute abominable fails, they too will be obvious. The thing to look at is where is that line to be drawn between a pass and a fail? And that’s the threshold standards. So it’s not so much, that’s a different purpose, the threshold. We’re looking at grades integrity.’ So what I’m suggesting is this: Consensus moderation is well understood by us. We can scale it up and we can scale it back a bit. We can get used to the idea of cognate courses not being frightening to us and we can look across institutions, external calibration, they can take their time. What I’m suggesting to you is this is, I think, a credible way, a useful way, an academically respectable way, a way that respects our integrity but it means that we have to start being less siloist in the way we grade. Less protective of our little patch.

I’ve just written an article, a manuscript and article on academic freedom, achievement standards and professional identity. In this article, as I’ve been writing it, it becomes clearer and clearer to me that in the terms of grading that’s the one bit of academic life that we seem to want to protect. We seem to be quite happy about peer review of all sorts of things. We seem to be quite happy about construction of programs collaboratively, of looking at course construction collaboratively. Even looking at assessment tasks and assessment plans, we’re happy to do that. When it comes to the grading that’s the people at which people get excited and I think why is grading the only thing at which people are protective. It should be part of our professional identity as educators. We are educators. If you work in a university teaching students, you are an educator. When I say the lecturers, who are you, they often say, I’m a chemist and I’m an archaeologist, I’m a this, I’m a that or something else. Hardly ever did they say I’m an educator. If you’re in a university and you’re teaching students half of your professional identity at least is as an educator. And it also then is incumbent upon us as educators, which I am too, educators, to know something about education and to grasp the standards nettle by the throat, wring it out and try to make it work for us so that no-one makes things imposed upon us. So that’s what I wanted to say. You’ll find most of that in the handout that you’ve got in there except that I’ve actually laboured the bit about across institutional comparability because I really think that is something which universities, whether they be currently high status or low status need not fear. And for those that are able to produce the goods with high quality teaching, and the other thing too is they may need differential resource allocations if they are to take on students from backgrounds which are disadvantaged. They’ll need increased resources to be able to get those students up to the level. But I think that’s really something that we’ve got to take on board as an agenda which we as academics argue because I think the grounds on the other side are pretty weak if not non-existent.

Thank you.