

**Achievement Matters:
External peer review of accounting learning standards**

An innovation and development project sponsored by

The Office for Learning and Teaching, Australian
Government, and the Australian Business Deans Council

Final Report of the External Evaluation

*Patrick Boyle MAES,
Consultant and Director, Q Associates
& Visiting Fellow, UNSW*

External Evaluator

December 2014

Contents

| | |
|--|----|
| Caveat and special notes | 3 |
| Acronyms used | 4 |
| Executive Summary | 5 |
| Acknowledgments | 13 |
| Description and context | 14 |
| Evaluation for the Project | 22 |
| Conclusions referenced to evaluation questions | 25 |
| Evaluative opinion in summary | 35 |
| References | 36 |

Caveat

External summative evaluation of a project can only provide one evaluative narrative and this can never capture the full story and nuances of results, merits and processes. This limitation is exacerbated when external evaluation personnel are engaged more at the periphery of a project than at the core or there are other factors that limit the richness of engagement (e.g. geographical dispersion of parties; limited budgets). In any case, if done comprehensively and openly, the reports provided by a project's leadership and team do a better job of illuminating its merits and any limitations.

Special Notes

1) The *Executive Summary* in this report is exceptionally long. This is because due to some unusual circumstances near the end of the Project I provided the Project Leaders with an early stand-alone version of this summary.

2) Understandably, readers of evaluation reports often go straight to sections concerned explicitly with "evaluation", such as conclusions, findings, etc. Often, contextual factors (e.g. special circumstances, project complexity) are highly relevant to a project's evaluation. This is certainly the case with the Achievement Matters Project, so I suggest that Section 1 of this report will have relevance for readers with a deeper interest in the Project's richness and merits.

Acronyms used

- **AACSB**
Association to Advance Collegiate Schools of Business
- **ABDC**
Australian Business Deans Council
- **AFAANZ**
Accounting and Finance Association of Australia and New Zealand
- **AM Project**
The Achievement Matters Project (short form)
- **BE**
Business Education
- **CAANZ**
Chartered Accountants Australia and New Zealand
- **CPAA**
Certified Practising Accountants Australia
- **EQs**
Evaluation Questions
- **HEPs**
Higher Education Providers
- **HESP**
Higher Education Standards Panel
- **OLT**
Office for Learning and Teaching in Higher Education, Australian Government
- **PT**
Project Team
- **TEQSA**
Tertiary Education Quality and Standards Agency

Executive Summary

This report concerns the commissioned external evaluation of the project *Achievement Matters: External peer review of accounting learning standards* (the AM Project).

Overarching comments

The AM Project has been an extraordinary success and there is no doubt that it is an exemplar for its kind. While its focus has been on Business Education (BE) in Australia, its effects beyond this field are expanding rapidly and it is receiving international recognition (e.g. Rust (2014), Times Higher Education, UK). I use the term “exemplar” very rarely so a little later I provide a brief description of my rationale for this judgement. Before doing that I provide a brief overview of what the Project was about and then list in general terms some of its stand-out merits.

The Project developed a widely commended model, based on researched good practice, and a set of aligned operational processes for peer review-based assessment and validation of student learning outcomes in higher education (HE) referenced to nationally-agreed learning standards. The processes incorporate the continuing refinement of educators’ and practitioners’ understandings of standards and the quality of their assessment judgments through iterative critical reflection and learning. This has never been done before in Australia nor so thoroughly anywhere else.

The merits and value of the Project are difficult to summarize briefly so at this top level I offer the following highlights.

1. The AM Project sought to effect substantial change and faced a number of complexities, so it had a very high degree of difficulty. In light of this, it was a very courageous initiative.
2. The Project achieved sophisticated and highly valued results. These spanned thought leadership in relation to standards-referenced assessment, real change in practices at national level and commended contributions to discourse and scholarship.
3. Overall, the Project achieved results well beyond its objectives.
4. Three process characteristics warrant highlighting as critical success factors: outstanding leadership, a very effective overall project implementation

strategy and a comprehensive and strategic approach to stakeholder engagement and change.

It is worth explaining the broad basis for my “exemplary” call, in part because of the positive coincidence that the Project is about *standards*. My evaluation, elaborated on in this report, is grounded in clear evidence of the merits of the Project on several dimensions. This basis is analogous to criteria-referenced assessment of student learning. However, my judgement is also underpinned by comparative or ‘norm’-referenced evaluation. Over the last seven years I have been the external evaluator for several OLT-sponsored projects and I have a good knowledge of many more. As well, for more than twenty five years I have been involved internationally and in Australia with more than one hundred other educational development and change projects that are comparable in scope and context to the AM Project. This experience-based reference frame is helpful for project evaluation work. In my view only about five of these 100+ projects warrant the judgment of *exemplary*.

Working as the external evaluator for the AM Project provided me with a surprising first experience. Part of my overall philosophy on evaluation is the belief that, in the interest of maximizing the utility value of findings for different people, it is important to be balanced in what is reported. While asymmetry in favor of positive aspects of a project is always pleasing, there is a need to provide fair and useful critique as part of an evaluative narrative. For the first time in my project evaluation work, with the AM Project I could only find strengths, mostly exceptional ones, and at the summative level there was simply nothing of significance that warranted critique.

Evaluation overview

Broad context and drivers

My reading of the Project Leaders’ motivation and rationale for the Project was that essentially there were two components, intrinsic and extrinsic. Together, these motivations point to the great importance of the Project’s mission. The intrinsic motivation relates to the leaders’ ethical concern for good practice in higher education, specifically, in this case, the need for improved enabling practices for and greater confidence in the validity of high-stakes summative assessment of student achievement, particularly at the level of their capabilities when graduating. The extrinsic motivation derives from the return to the foreground of a national debate and related activities concerning standards in HE and the need identified to be actively engaged in this. This most recent round gathered steam around 2009 and has the usual political, philosophical and technical dimensions, as well as several important engaged stakeholders. Leaving aside conjecture about deeper ideological aspects,

both Liberal and Labor federal governments continue to have concerns related to international reputation and the export value of higher education (currently > AUD 10 Billion). At the policy and operational levels the Higher Education Standards Panel and the Tertiary Education Quality Standards Agency (TEQSA) continue to grapple with how to improve the means by which HE providers will be accountable for learning standards, particularly at the graduating level. Clearly, universities and their schools and disciplines have a stake in the standards debate and the policy, accountability, guidelines and practice changes that will flow from this. Professional associations such as CAANZ and CPAA (for Accounting) are another key stakeholder group. Of course, academics and students in HE are the people affected most by assessment practices and the aligned learning and teaching that should underpin them. They are very important stakeholders, although their level of engagement overall with the issues around standards has been limited.

Having accurate and informative student achievement results that make sense to people interested in them, and which can be compared meaningfully across institutions, is self-evidently important for students, employers and ultimately the wider community. However, historically there has been very little evidence at a comprehensive level to show that processes exist that satisfy this aspiration. One of the major reasons for this is the complexity of the field of assessment of learning. My own experience and that of several expert colleagues suggest that it can be a minefield of challenges when seeking to effect change. In short, for assessment practices to be of high quality overall, many process and practice elements, which are often complex in themselves, need to line up and be right. This is the case, regardless of whether the particular difficulties associated with accurate *measurement* are present (as part of assessment methods). If even one of the key elements involved is low in quality (e.g. the definition of a capability being assessed; flaws in assessment instruments that reduce reliability of resulting judgments) then the validity of whole processes and results will suffer. This general requirement is well documented in the vast literature on assessment and evaluation (see for example Balla and Boyle, 1994; Baume et al, 2004; Price et al, 2008; Sadler, 2009; 2014).

Purpose and challenges

At the highest level, the AM Project was about effecting real change in embedded university practices, where the need for change as perceived by powerful stakeholders has historically been low and their resistance to it high. The literature on change makes it clear that in complex practice environments it is extremely difficult to know with confidence how to effect change in beliefs, attitudes and practices. This circumstance is the first key indicator of the high degree of difficulty of the Project and the courage behind its mission.

At the next level, the domain of practice where change is being addressed is assessment of student achievement. Within this, the Project's focus was on critical review and improvement of key assessment processes so that achievement results can be accurately derived, referenced to and understood in terms of explicit capability standards (e.g. expected standards for *knowledge* or *communication* at the end of an undergraduate degree program).

Along with these two major challenges, like many OLT funded projects the AM Project had a limited budget and was implemented by a part-time team of very busy people.

Success in the light of challenges and difficulty

Because of the challenges it faced I formed the view that the Project as an extremely courageous one. It is worth noting that much of the literature on change facilitation and innovation stresses the importance of courage as a critical success factor.

On the ground, the Project's direct work to date has been in BE, particularly accounting. It has focused on the development of improved understandings and practices in standards-referenced assessment in business schools in HEPs across Australia. It has achieved much on this front and in light of the complexity and challenges faced, this is an extraordinary result. Overall, the Project achieved way beyond its objectives and in some areas effects are accruing that could not have been reasonably expected (e.g. the breadth and richness of professional learning and development for educators; the extent of penetration achieved into the national and international discourses on standards-based assessment in HE).

The leadership of the Project has been outstanding. Mark Freeman (USyd) and Phil Hancock (UWA) demonstrated constant passion for its cause, strategic creativity, great ability to facilitate collaborative work with educators and other important stakeholders, and exceptional work ethics. Underpinned by this leadership, the Project Team (PT) dealt extremely well with the challenges and complexities faced and as a result successes continue to be achieved that otherwise would not have been possible. The team's collective intellectual capability was very high and members brought a diverse range of talents and areas of expertise. The members of the PT, including the people who worked in the important project management-support roles were:

Associate Professor Mark Freeman, USyd, Project Leader

Professor Phil Hancock, UWA, Project Leader

Associate Professor Anne Abraham, UWS

Associate Professor Bryan Howieson, University of Adelaide

Professor Paul De Lange, Curtin
Professor Brendan O’Connell, RMIT
Professor Kim Watty, Deakin
Ms Coralie Bishop, UWA
Ms Catherine Vogel, UWA

The AM Project has delivered excellent value for funds and time invested. Because of the deep stakeholder engagement and wide buy-in it has established, there is high potential for more and wider positive effects to be added over the next few years.

Results highlights in brief

The Project has achieved several very valuable results, particularly when a finer grained view is taken. These will be described comprehensively in the reports submitted by the Project’s leadership to the Office for Learning and Teaching (OLT). The Achievement Matters website also provides a good summary of the Project’s work (<http://achievementmatters.com.au/>). For me, at the highest level there are three stand-out achievements. Here, I will simply list these and comment briefly on the first of them. Elaboration on all three is provided in the body of this report.

1. Development and implementation of a conceptually sound, culturally appropriate and workable “model” and a set of processes for standards-referenced assessment of student achievement based on calibrated external (academic) peer review.
2. Highly valued learning and development for academics and other educators, with positive compounding effects for assessment design and practices (and curriculum development more generally) in business schools within Australian higher education providers. *See special note below.*
3. Substantial enhancement of the discourse and provision of thought and practice leadership and development in Australia and contributions to international scholarship pertaining to standards-referenced assessment in higher education. *See special note below.*

Special note concerning achievements 2 and 3 above. *As a result of a sophisticated and comprehensive dissemination-engagement-change strategy, a wide range of key stakeholders in BE (and HE more generally) remain actively engaged in learning from and applying outcomes of the AM Project. These include, the HESP, TEQSA, the OLT, business schools and their academic program teams (in universities and other HEPs), and several professional bodies, including CPAA and CAANZ.*

The first of these major achievements underpins the others. From the early days of the AM Project I held the view that the “model” being developed, while suitably pragmatic, was genuinely leading–edge in nature. The most powerful and distinctive feature of the model is its emphasis on effective (real) calibration of external/independent assessors of student performance. This is one of the necessary conditions for being able to demonstrate good validity for high stakes summative assessment results. While regarding the model very highly in conceptual–technical terms, I had concerns about the challenges that would be faced to get it widely accepted and working well in practice.

Evidence of the high quality of the AM model continues to grow. This includes expert opinion that it is clearly superior to current ‘next best’ approaches for validating summative assessment in HE (e.g. Rust, 2014; Times Higher Education, UK). The approach often referred to is the traditional *external examiner process*, such as that used in the UK and sometimes in Australia. Historically it has been argued that this enables accurate validation of assessment results and provides a sound basis for comparing standards across institutions. Details of the theoretical basis for and workings of the model were recently published in the peer–reviewed international journal *Assessment and Evaluation in Higher Education* (Watty et al, 2014). The model is conceptually and technically sound. Importantly, it has been shown to work well in practice, mainly because it is well grounded in existing academic cultural norms (e.g. affinity between discipline–based peers) and its operational processes have been well designed, tested, implemented and improved based on stakeholders’ critique.

While modest financial cost is a factor in the proper implementation of the model, there is an extremely strong argument that the value added would be well worth the cost. The principal component of this value would be real and much increased confidence that students’ university results could be understood by reference to meaningful capability standards and that results from different universities (at discipline level, such as *accounting* or *finance*) could be meaningfully compared. The opportunity costs of prospective decrease in the reputation and standing of Australian HE, if this kind of model is not adopted, need also to be factored in.

Concluding comments and the future

Expanding stakeholder buy–in and embedding change on a wider scale, particularly in BE, are likely to be the highest priorities for the people who are enthusiastic about sustaining the effects and momentum of the AM Project.

For addressing these related priorities, the model and processes developed by the Project are great strengths. They are receiving national and international commendations, principally because they are technically sound and have good efficacy for enabling authentic standards-referenced summative assessment and external validation of standards being achieved. They provide a credible linchpin and clear guidance for facilitating important improvements in three critical components of high-stakes summative assessment.

1) How to accurately assess (judge) levels of student achievement or performance by reference to explicit discipline-level learning standards;

2) Provision of an effective means for the continuing calibration of external assessors, which is acceptable to educators, and which is a necessary condition for achieving suitable accuracy in high-stakes assessment.

3) As a result of 1) and 2), provision of a more sensible basis for comparison of achievement standards across institutions or programs along with consequential learning about how improvements in curriculum and teaching can be made.

In addition, if the use of processes similar to those developed by the AM Project became the norm across the higher education sector, Australia would have a much stronger basis for its claims about taking standards seriously and having the means in place and evidence to demonstrate this.

I suspect that 'politics', not the quality of the AM model, will be the biggest longer-term challenge to achieving wider buy-in and action. The politics I refer to relate to the formal government levels and the more circumscribed domains of universities and key stakeholders such as academics. In brief, arguments against significant change will ostensibly hinge on cost (for perceived return), aversion to increasing complexity, more important priorities, the belief that current practices are good enough, workload pressures, and philosophical-ideological strands such as the sanctity of the university's and the academic's roles and judgements.

In the shorter term BE in Australia has a great opportunity to be a trail blazer. Some disciplines in HE already have quite rigorous means for assurance of graduate capability and performance standards (e.g. Medicine). However, from a technically sound evidence-based perspective, many disciplines have weak approaches. There is clearly a chance to make a mark nationally and internationally by adopting the highly credible and more rigorous approach to assessment developed by the Project.

The critical mass of people who are currently actively engaged in the Project or with its outcomes bring an enormous body of knowledge and 'how to do' capital. What is needed over the next two or three years to expand and embed effects is high level championing and support by leading entities, particularly the ABDC, professional bodies like CAANZ and CPAA, and the OLT. Enabling financial resources will also be required, but in my opinion, the amounts needed would be modest particularly when the potential for return on investment is factored in.

My commendation of the AM Project could not be higher and I advocate strongly that interested parties provide as much material support as possible for well-conceived initiatives or entities that will sustain its momentum.

Acknowledgements

From the beginning of the AM Project I have been very grateful for the opportunity to be associated with it. The experience has been extremely stimulating and I continue to value it highly. For this I must thank the Project's leaders and team for having confidence in me.

As the Project Leaders, it's hard to overstate the level of passion, leadership and facilitation talents of Professor Phil Hancock (UWA) and Associate Professor Mark Freeman (USyd). In my career, I have been fortunate to have worked with a number of inspirational leaders and I certainly count them in this small group. Importantly, in the context of my role as external evaluator, Phil and Mark were always open to my advice, showed respect for it and often factored it into their decision making, particularly where decisions or actions related to project evaluation and related matters such as strategy or planning for project success.

In addition to getting the required work done on such a demanding initiative, the Project Team was a lively bunch. Members brought a diverse mix of talents, high level collective intellect and good humour. This made my role easier and more enjoyable. On the operational management front, my special thanks go to Catherine Vogel and Coralie Bishop (during the first half of the Project) for their valuable skills and willing assistance whenever I needed them.

The Project would not have been successful without the continuing active engagement of the university academics, other educators and professional practitioners. As well as thanking them for their contributions overall, I am especially grateful to the people who agreed to talk with me, either formally or informally. Their observations and judgments about the merits and value of the Project were extremely useful for helping me with the external evaluation.

I would also like to express my gratitude for the many valuable contributions, of different kinds, made by internationally respected experts associated with or in the space of the Project. For my evaluation work, some of their written materials, presentations and personal communications were very insightful and helpful. I am particularly grateful to Dr Berry O'Donovan (UK), Professor Royce Sadler (Australia) Professor Alan Robson (Australia) and Professor Chris Rust (UK).

1. Description and context of the AM Project

Meaningful evaluation requires at least an adequate description of what is being evaluated, that is, the *object* of evaluation. When it comes to project evaluation, usually the documented purpose and objectives of the initiative provide the core of such a description. However, and this is quite uncommon for OLT sponsored projects, in my view the AM Project is not sufficiently defined by its original articulated purpose and objectives. This is simply an observation, not a criticism. Because they are important for a richer description of the Project and for realistic evaluation, I am highlighting an objective and a contextual factor that are not emphasized explicitly in its foundation objectives.

First, in a broader sense the Project aspired to effect *real and substantial change* in a key area of academic practice, summative assessment of student achievement. It is well known that academic practice is one context where change is often very difficult to achieve. Second, the Project had considerable technical complexity. For example, in seeking to improve assessment (evaluation) of *student learning*, which in itself is complex, it would need to focus on several elements and sub-processes, including the mental models and judgement processes of academics, their attitudes on different aspects of assessment, pedagogy, accountability, etc. and the need for them to engage in iterative learning. Taken together, this change objective and the technical complexity factor meant that in prospect the Project had a high degree of difficulty.

In this section, I provide a summary of my reading of the Project's intent, a brief account of its context and some elaboration on complexity and difficulty. All of these matters are important for framing and reporting on a realistic summative evaluation.

1.1 Purpose and objectives

The AM Project set out to build on earlier initiatives concerned with HE standards (e.g. the OLT funded project Learning and Teaching Academic Standards (Accounting), 2010). In broad terms its initial aim was to develop and implement a *model* for benchmarking assessment and moderation against agreed student learning standards. One of the key features of the model was that it would serve both quality enhancement and quality assurance purposes in relation to assessment processes and results. The Project's overall process was to be highly collaborative and the model itself would hinge on a discipline community-based consensus moderation approach to standards-referenced assessment. In addition the model and its operational practices would be applicable to other disciplines within BE and beyond. In building

on previous work the Project would be highly inclusive by, for example, inviting the participation of people from a diverse range of HEPs.

Objectives to expand on this purpose were expressed as intended outcomes. The following is a version of these articulated in 2011, early in the Project's life.

1. External peer-reviewed evidence of accounting academic standards in all HEP types, benchmarked against the LTAS accounting threshold learning outcomes
2. A model process for obtaining and using blind external peer review evidence that captures inputs (i.e. assessment specifications) as well as outputs (i.e. completed student work)
3. Professional learning and capacity building

Even collectively, these intended outcomes understate the technical complexity of the proposed work and say almost nothing explicit about the fact that the Project was concerned with effecting real change in thinking and embedded practices. To do justice to a *description of the Project*, it is important to elaborate a little on its complexity and underscore the change facilitation challenge. I will return to complexity in section 1.3.

To provide a brief picture of complexity, I describe in my terms what I understood the aspirations of the Project were. To be "brief" I limited myself to four short statements of intent. As these reflect my understandings, the set will be imperfect. That doesn't matter much, because I am confident that I am not too far off the mark – and it is the complexity I want to illuminate, not the precise details of intent.

What the Project was setting out to achieve through my eyes

1. Through a facilitated consensus reaching process, for the first time, enable HE Accounting educators (initially) from across Australia to develop clear and shared understandings of established accounting learning standards for a set of capabilities (e.g. knowledge; communication and teamwork).
2. Enable educators to develop high level capabilities in relation to the derivation of categorical assessment results (i.e. 'meets a standard'/'does not meet a standard') based on their judgements of student performance (quality) on a range of assessment tasks by reference to the learning standards.
3. As part of the Project process, critically review a range of student assessment tasks and other inputs (e.g. unit of study outlines) and facilitate academic

learning and development to support the improvement of assessment design and practice at program and unit of study levels.

4. Enhance the quality (reliability and validity) of educators' assessment judgements through iterative processes involving review of data sets, feedback loops, reflective discussions and learning.

One extraordinary strand of the Project's overall aspiration, which has received little attention, is that it would be seeking to change quite deep and sensitive aspects of educators' ways of thinking and doing. This isn't the place to go into detail on this, suffice to note that the Project would be disrupting established mental models and personal equations in relation to student assessment and then aiming to facilitate refinement of them through critical reflection, learning and practice.

Aside from this deeply individual type of change, it is also important to acknowledge the challenge brought by the broader *change facilitation* aspect of the Project's purpose. Historically, the need for the kind of change in practices being sought, in the eyes of many influential stakeholders, has not been recognized. More generally, the literature makes it clear that in complex practice environments it is extremely difficult to know with confidence how to go about effecting change in embedded beliefs, attitudes and practices.

The Project's change aspiration, along with its technical complexity, showed me that it had a *high degree of difficulty*. In turn, this led me to the view that it was a very courageous initiative.

1.2 Strategic context and drivers

The Project Leaders were the driving force behind the idea and realization of the AM Project. Of course, there is a wider strategic context, but I have chosen to summarize this in terms of the leaders' motivations.

My reading of the leaders' motivation was that there were two related components, intrinsic and extrinsic. Considered together, these provide a good indication of the roots and importance of the Project's mission. The intrinsic motivation relates to a clear ethical concern for good practice in higher education, specifically, in this case, the need for improved practices for ensuring and greater confidence in the validity of high-stakes summative assessment of student achievement, particularly at the level of their graduate-level capabilities.

Much of the extrinsic motivation derives from the return to the foreground of a national debate and related activities concerning standards in HE and the need they

identified to be actively engaged in this. This most recent cycle of dynamism on standards emerged from around 2009 and has the usual political, philosophical and technical dimensions, as well as several important stakeholders. Leaving aside conjecture about ideological threads, both Liberal and Labor federal governments continue to express concerns related to international reputation and the export value of higher education (currently > AUD 10 Billion). At the policy and operational levels the HESP and TEQSA continue to grapple with how to improve the means by which HEPs will be accountable for *Threshold Standards*. The AM Project work is focused specifically on the element in the HE Standards Framework concerned with the design and achievement of course learning outcomes. Clearly, HEPs and their schools and disciplines have a stake in the standards debate and the policy, accountability, guidelines and practice changes that will flow from this. Professional associations such as CAANZ and CPAA (for Accounting) are another key stakeholder group. Of course, academics and students in higher education are the people affected most by assessment practices and the aligned learning and teaching that should underpin them. They are very important stakeholders, although historically their level of engagement overall with the issues around standards has been limited.

1.3 A high difficulty index as a special contextual factor

As noted earlier, the AM Project was complex in a number of respects. In this part of the report I discuss complexity a bit further in order to enrich the Project's description. The fact that the leadership and team dealt very effectively with this complexity is a major point of merit which has helped shape my evaluation opinion (more on this later). My comments below elaborate only on the two key aspects of complexity that I identified earlier.

For me, the most powerful factor contributing to complexity was that the Project domain was *educational measurement and evaluation* and incorporated *standards*. While for some it seems easy to use the term "assessment" and imply that 'there's nothing too complicated about that', the fact is that this field is a complex one, particularly if a scholarship-based good practice orientation is adopted. Before moving on to the transdisciplinary field of evaluation, I spent more than a decade immersed in the technical study and applications of measurement, physical and psychosocial, and I specialized in educational measurement. My foundations in this discipline were shaped in large part by being taught by some of the world's leading scholars in the field at the University of Western Australia. So my remarks under this heading are made with my (purer) educational measurement hat on.

1.3.1 Evaluation of student learning: a potential minefield

Late in the Project's life I interviewed Dr Berry O'Donovan one of the ABDC's 2014 visiting international experts. As it aligns well with her research and development work in the UK, she had a strong interest in the work of the Project. While talking about the merits of the model being developed and how it might be received more widely, she made the following remark.

Well, the problem is that there are so many parts (to assessment) and you need to get them all right

For some time I had been reflecting on the difficulties and complexity of the Project and this general comment from Dr O'Donovan, along with others she provided, reinforced some of my own thoughts and helped me to figure out how I could write about them. To keep it relatively brief, I provide some bullet points below. These could all be elaborated on but my aim here is to indicate the range of things (the "parts" referred to above by O'Donovan) that need to be considered and can present challenges for efforts to develop improved processes in *evaluation of student learning*. I use this term because it encompasses: assessment, as we use the term in Australia and the UK, evaluation (because, in fact, that is usually the end-game when we 'assess' student achievement for summative purposes), and standards, which is a particular focus of the AM Project.

- To be considered technically sound (e.g. accurate and precise enough for the particular use intended), assessment requires both processes and results that possess a number of properties. A few of these are: clear definition of the variable (e.g. capability) being assessed; definition of meaningful units and/or levels of measurement (where measurement is involved); evidence of the nature of the dimensionality of variables being assessed; the existence of suitable assessment instruments (e.g. accurate, precise enough); at least sufficient observations to enable reasonable judgements; and knowledge of and the capability to allow for errors. It is widely accepted that when assessing in the psychosocial domain it is often very difficult, if possible at all, to satisfy some of these technical requirements.
- Many learning variables that we assess are complex (e.g. ethical orientation; critical judgement; understanding of difficult concepts; communication capability across different contexts). Sometimes they are latent, that is, intrinsic to the individual and not observable directly, they might have

multiple facets, and efforts to assess them are often influenced by variation in contexts, conditions, etc.

- Assuming valid observations or other information (e.g. measurements; material evidence) concerning a learning variable have been obtained, the perennial challenge faced by educators is how to evaluate results effectively. How good is this result? What score or performance level is satisfactory? How will excellence be identified? Which criteria or standards are used to reference and report achievement and how are the observations/other evidence used to help do this? How is variation in observations *and* evaluation dealt with?
- In general, having credible standards is a big help for evaluation of the quality of any phenomenon. Historically in HE, the setting of (real) standards by which student achievement will be evaluated has proved to be difficult. Even when supposedly good standards exist, dealing with variations in interpretations remains one of the challenges. It is also often difficult to get academics and other educators to actively engage with explicit learning standards, particularly if they have been developed by others.

The points I touch on above and many others are elaborated on in the vast literature on assessment (see for example: Balla and Boyle, 1994; Baume et al, 2004; Price et al, 2008; Sadler, 2009; 2014). My hope is that this limited identification of the challenges that can be encountered when doing innovation and development work in educational evaluation provides an indication of the high level of complexity of the AM Project's domain.

1.3.2 Getting academics to actively engage with substantial change

The general difficulty of facilitating change in embedded practices warrants further emphasis. There are bodies of well-regarded research and literature on the difficulties of facilitating change in behaviors and cultures, including in universities. It is not appropriate to go into this here, but it is worth noting a few general points.

- A common reaction when people are asked to actively engage with change is – things seem to be working well enough, why change anything?
- Educators are extremely busy and many feel increasingly pressured, particularly about their performance in research – so time is precious.

- Many academics believe strongly that the sanctity of their own judgments, as closed groups and individuals, must be protected – ideas, processes, practices and particularly scrutiny (from ‘outside’) are not needed or wanted.

While academics more or less accept critical review of what they do (in research), historically they have been far less comfortable when it comes to matters relating to teaching and learning, including assessment.

1.4 Project strategy

Having a good idea of the *overall strategy* adopted for achieving a project’s objectives can be important for both formative and summative evaluation. For summative evaluation, it is sometimes expected that the effectiveness of project strategy be systematically assessed. Even when the efficacy of strategy is not questioned formally as part of evaluation, as was the case with the AM Project, identification of key elements can help to provide a more complete picture of the project. Of particular interest are some of the strategic choices made about ways of working and methods. These provide reference points for judgements made later about project success factors (causes) as well as factors that might have inhibited success. Such judgements contribute to summative evaluation.

In addition, description of project strategy, along with subsequent evaluative findings about elements that worked well, can be valuable for people working on other projects, particularly when they are in the planning and design stage.

As with some other aspects of the Project, I have left detailed descriptions of strategy to the Project Leaders via their report to the OLT. For the current purpose, I provide a synopsis of what I perceived to be the leaders’ overall strategic perspective for implementing the Project and then I list and comment very briefly on some of the main elements of strategy. I express these as *strategic choices*, made by the leaders and team that in my view were largely evident from the beginning. In the main they concern characteristics of the intended way of implementing the Project.

My perception of the Project Leaders’ overall perspective on strategy*

** More or less at day 1 in the life of the Project*

The Project is about effecting substantial change in thinking and practice in assessment – this won’t be easy. There will need to be a clear, methodical and

energetic approach to achieving the objectives. At the same time, improvisation and adaptability will be necessary (and a good thing). While direction and structure will be important, the best approach will be one that gets all direct participants, including the leadership, working in a flat structure and as a genuinely collaborative peer group. Strategic and comprehensive dissemination to actively engage all key stakeholders will be critical for success. It will also be important to show people that the processes being advocated are workable. This will need to include efficiency, unobtrusiveness and respect for sensitivities.

Main elements of strategy: expressed as strategic choices

1) Ensure 'cultural fit' for BE academics

Work with a tight discipline-based peer group (e.g. accounting academics). Respect for and association with their discipline along with comfort working with peers are well known features of academic culture. A distinctive feature guiding the Project is the notion of emulating accepted cultural practice for assuring research outcomes for publication, namely double blind peer review by two external reviewers.

2) Demonstrate the case for the work/change being embarked on

Focus on the emergence of stronger external expectations (e.g. government through TEQSA; accreditation bodies such as AACSB), our (educators') own concern for standards, threats to 'Brand Australia' in relation to BE, and having better evidence of standards and improvements (e.g. increased confidence in assessment judgements and results).

3) Ensure a collaborative and non-threatening group disposition

Learning together experientially in a constructive and non-threatening manner is the only likely path to success. As leaders and a team we need to be peers and facilitators, not holders of authority. No 'culture of criticism'. It will be necessary to develop individuals' confidence and comfort to speak about sensitive matters, encourage active listening and mutual respect, and generate group learning. Confidentiality and anonymity in some aspects of the work will be necessary.

4) Learn through doing, reflecting and feedback

Effective processes will be used to enable induction and experiential peer group learning. Time needs to be allowed and a safe and supportive ethos created to ensure

that people will express themselves openly, listen actively and respectfully, respond to constructive feedback and learn continually.

5) Provide a credible technical foundation (scholarship of assessment)

It will be important to provide a clear and sound basis for the work, principally from assessment theory and good practice (literature-based).

6) Use suitable paradigms to help achieve goals (e.g. social moderation)

Challenges will exist for achieving learning (change) goals, some of which involve sensitivities and entrenched beliefs. Evidence suggests that adopting a *social moderation* approach will improve the likelihood of success in getting people to seriously question their mental models, developing shared understandings of learning standards and engaging actively in improving their standards-referenced assessment judgements.

7) Ensure stimulating and effective activities/tasks for participants

Use real student assessment tasks and performances (responses) from participating universities. Have clear processes and guidelines in workshops as well as very effective facilitation. Ensure the mix of ICT is appropriate.

8) Demonstrate the workability of processes

The processes being demonstrated will need to be seen by educators and senior people in organizations as workable. This means they need to be efficient, relatively unobtrusive, cost-effective and non-threatening.

9) Integrate strategically effective dissemination

Dissemination must address different strategic goals and emphasize achievement of active stakeholder engagement. Expectations and dissemination methods need to be matched to what is practical and the needs of different interested parties.

2. Evaluation for the Project

The Project Leaders brought an open and critically reflective disposition to project evaluation. They needed little encouragement to integrate evaluation into implementation of core work. They were supportive of or proactive in facilitating in-

progress review of strategy, activities and results. Importantly, they expected and modeled a culture in the PT (including behaviors) that ensured effective monitoring of effectiveness and quality assurance.

2.1 Evaluation approach and methods

Given resource and logistical constraints, a pragmatic evaluation approach (strategy) was necessary. In broad terms, it can be characterized as follows.

- Type/Focus: Primary – outcomes/impact; secondary – learning about and improving the efficacy of processes
- Purposes: Formative, learning and summative
- Strands: Merit (extent and quality of outcomes); perceived value
- Methodology: Integrative, utility-oriented, mixed-methods, pragmatic

Essentially, formative evaluation at the operational level was left in the hands of the PT. As external evaluator I provided general strategic advice, facilitated a broad critical review of the Project's logic (strategy for success) in early 2012 and had regular discussions with the Project Leaders together with periodic ones with the PT.

Formative evaluation was taken seriously. Gathering and acting on relevant feedback on key aspects of the Project, including critical success factors (e.g. effectiveness of assessor calibration workshops; means for communication about the Project), was an integral part of its implementation, as were genuine critical reflection at PT level and a willingness to continually improve processes and materials.

2.2 Evaluation questions (EQs) for summative evaluation

Summative evaluation was framed by a set of *Evaluation Questions (EQs)*. In 2012, I facilitated a consultative process with the PT to agree a set of EQs. In early 2013 these were refined slightly to reflect the evolution of the Project.

As is often the case in professional evaluation, the summative EQs were set at a relatively high level. Their main purposes were to make clear the broad criteria by which the Project would be evaluated and guide the review of available data and evidence to draw conclusions. They have another valuable purpose, which is to help generate more precise questions for data gathering processes and cues for review of existing data/evidence. For the AM Project they helped shape the final Project Internal

Evaluation Summary (PIES) (e.g. questions asked; information requested), interviews conducted and desk-based reviews of information.

Evaluation Questions

1. What are the major achievements/results of the Project and how well aligned are these with its main objectives?
2. What are the principal merits of 'the Model' developed through the Project for enabling the assessment and validation of student learning standards?
3. How effective has the Project been in terms of dissemination?
4. In relation to its principal outcomes and objectives, to what extent has the Project succeeded in generating positive engagement, buy-in and/or support amongst key stakeholders, particularly:
 - (a) the Australian higher education community in Accounting; and
 - (b) the two key Australian professional bodies in Accounting, ICAA (CAANZ) and CPAA?
5. To what extent has the Project enabled valuable learning/professional development for those engaged with its activities?

2.3 Methods and information used to help answer the EQs

It was intended that most of the evidence (e.g. data, material resources) needed to help me derive answers to the EQs would be accumulated as a result of the production work of the Project (e.g. documentation of outcomes, including publications; implementation of demonstrably successful workshops) and the formative evaluation processes facilitated mainly by the PT (e.g. critically reflective and evaluation-focused discussions in PT meetings; surveys used to monitor and improve the effectiveness of the assessor calibration workshops).

I also facilitated or engaged in processes that were intended, in part, to yield evaluative information. The following were particularly valuable.

- Development and refinement of an intended project logic (IPL) in 2012
- Reflective discussions with the Project Leaders, several times each year
- Discussion around and agreement on the evaluation questions in 2012/13

- Interviews and conversations with project participants and other relevant people, including experts, particularly in 2013 and 2014
- Participation in a number of PT meetings over three years
- Observation of three assessor calibration workshops
- Submission by the Project Leaders in October 2014 of a structured Project Internal Evaluation Summary based on a template I provided

These mechanisms generated material evidence (e.g. documents), perceptions and judgements of well-placed people, my own observations and impressions, and other data that have helped me to draw conclusions aligned with the EQs.

The Project's comprehensive dissemination-engagement-change strategy, which included a large number of activities for different interested parties, resulted in numerous public media items that referred to its work, particularly during 2014. These items, along with publications and other written works, have provided useful information for summative evaluation.

3. Conclusions referenced to the Evaluation Questions

In this section I provide conclusions I was able to draw about the achievements and value of the Project aligned with the EQs. I will present or refer to examples of the evidence which underpin my conclusions, however to avoid repetitive references, I note that all conclusions are based on evidence resulting from the processes identified in section 2.3.

Understandably, the EQs overlap to some extent, so while some matters (e.g. dissemination) are relevant to more than one question I aim to limit my comments under each question to its main focus.

Evaluation Question 1: What are the major achievements/results of the Project and how well aligned are these with its main objectives?

The achievements of the AM Project are many and varied and the PT's final report to the OLT will describe these in detail. The Achievement Matters Website also provides a good summary of the Project's results and work (<http://achievementmatters.com.au/>). Here, I have chosen to focus on three high-level results including two which are at the wider *impact level*.

In highlighting selectively, I am not implying that other more specific or concrete outcomes are not important. In addition, some of these more specific outcomes are subsumed by the broader achievements I do highlight here. For example, achievements 1 and 2 (see below) both encompass the development of a well-prepared group of assessors (i.e. educators who have learned and been initially 'calibrated' to implement standards-referenced assessment in accounting). This in itself is clearly an important and valuable result. Achievement 3 subsumes the success of a very well-conceived and implemented dissemination and engagement strategy which resulted in, amongst others, the following positive results.

- Reference to the value of the Project's work by respected experts such as Professor Alan Robson, Chair of the HESP (Australia) and Professor Chris Rust, a highly respected scholar and academic leader (UK)
- The PT winning the 2012 BEHERT Award for Excellence in Accounting Teaching Collaboration
- Active positive engagement with the Project by key HE entities such as TEQSA and the OLT
- Citations in public and HE-focused media sources
- Publication of scholarly journal articles and other works

Typically, it is difficult for short-timeframe projects in areas like health and education to achieve substantial and/or widespread impacts, particularly if intended impacts involve change in embedded beliefs and practices. The fact that the AM Project has already achieved substantial and widespread impacts is an indication of its high merit.

Some Achievement Highlights

1. Development and implementation of a conceptually sound, culturally appropriate and workable "model" and a set of processes for standards-referenced assessment of student achievement based on calibrated external (academic) peer review. The merits of the model will be elaborated on under EQ2 below.
2. Highly valued learning and development for educators, with positive compounding effects for assessment design and practices (and curriculum development more generally) in business schools within Australian universities and other HEPs. *See special note below.*
3. Substantial enhancement of the discourse and provision of thought and practice leadership and development in Australia pertaining to standards-

referenced assessment in higher education. Contributions to international scholarship in this domain is a related outcome. *See special note below.*

Special note concerning achievement highlights 2 and 3 identified above. *The PT devised and implemented a strategic and comprehensive dissemination–engagement–change strategy. I will report on this under EQ 3 below, however, it is important to emphasize here that this strategy was a key success factor for these two impact–level achievements.*

These achievements, along with other valuable outcomes, align very well with the objectives for which the Project was funded. In fact, they go beyond what could have reasonably been expected, particularly in terms of dissemination effects such as the scale of active engagement with outcomes achieved.

Evaluation Question 2: What are the principal merits of ‘the Model’ developed through the Project for enabling the assessment and validation of student learning standards?

From here, when I use the term “the model” I am referring to the conceptual design and frame along with the operational processes that collectively constitute the approach developed by the Project for enabling demonstrably valid standards–referenced assessment.

My main focus here is on the *merits of the model*, not a detailed description of it, which will be provided in the PT’s final report on the Project to the OLT (and see Watty et al, 2014). However, before I return to the evaluative focus, there is a need for me to provide an overview of my understanding of the model.

To me, the model provides a conceptual framework, based on researched good practice, and a set of aligned operational processes for (HE discipline level) peer review–based assessment and validation of student learning outcomes referenced to nationally–agreed learning standards. It also enables the continuing refinement of educators’ and practitioners’ understandings of standards and the quality of assessment judgements through iterative critical reflection and learning.

At the Project level, the learning standards refer to expected levels of student achievement or performance in relation to particular capabilities (e.g. communication and teamwork; judgement; knowledge). Levels of achievement, indicated by students’

responses on assessment tasks, were differentiated as either 'meeting' or 'not meeting' particular standards.

Principal merits of the model

It is difficult to briefly summarize the merits of the model, so again I am selective. To begin with, three general characteristics warrant emphasis and commendation. First and foremost, the model has been shown to 'work', not just in an operational sense, but in terms of its suitability for academic culture. For reasons relating to project difficulty and complexity discussed earlier, this is a remarkable achievement. Keys to this aspect of success included: adoption of a very clear discipline peer and community based approach; use of a social moderation paradigm and practices aligned with this (particularly for facilitating development of shared understandings of standards and learning about and enhancing the quality of judgements); and the credibility of and exceptional facilitation provided by the Project Leaders. Second, because it is grounded in credible educational theory and research, the model is seen by educators to be conceptually and technically sound. Though more specific, a third merit warrants highlighting at this top level. The model has accommodated the critical factor of the *calibration of assessors*. This is a necessary, but almost always missing condition, for being able to demonstrate the overall validity of summative assessment results, including the suitability of results for comparison of achievement levels across programs and HEPs.

Other strengths and the value of the model in practice can be identified in more precise terms. The following are examples worth noting.

- The model provides clear evidence that valid standards-referenced assessment can be implemented in higher education
- In doing this it provides clear and credible means for taking care of issues associated with the calibration of assessors, such as the need for reduction to an acceptable level of the variance in inter-assessor judgements of student performance against standards
- Conditional on good quality facilitation, the model provides an excellent vehicle for important and valued learning and development for educators, particularly in relation to aspects of assessment design and practice
- Related to the above, in prospect the model provides a linchpin for a wider community of practice for BE (and other academic discipline groupings) around authentic standards-referenced assessment and related program quality matters
- The model also provides knowledge and processes for helping to address the external quality assurance expectations of TEQSA and the needs of

accreditation and professional bodies, which are increasingly focused on having credible evidence that assessment results can be referenced to explicit capability standards

Finally and importantly, the AM Project, and particularly the merits of the model it has developed, are increasingly being recognized by international experts. In a high profile critique of the situation in the UK, after identifying some of the positive characteristics of the AM Project model, Rust (2014) offers the following prompt.

Why is it that this issue is being seriously, and apparently successfully, addressed in this way in Australia, yet despite all the evidence and repeated criticism of our current systems, five years now since the Select Committee Report, there is no indication of HEFCE, the Government, UUK or the QAA wanting to seriously address the standards issue.

O'Donovan (Interview with the author, 2014) pointed to the critical importance of the quality of and coherence between the main elements of assessment if good levels of overall validity are to be achieved. Paraphrased slightly, the elements she identified were: the learning outcomes intended; standards for defining different levels of achievement; assessment tasks; and assessor quality, which incorporates the need for effective calibration. In her words:

The difficulty and goodness of it (assessment and the AM Project's model) is that you have to get all of these (elements) right, otherwise it falls down.

In my view, the model and associated processes developed by the AM Project gets many things right in the overall assessment process and in terms of international norms provides an enhanced and workable approach for standards-referenced assessment in HE.

Evaluation Question 3: How effective has the Project been in terms of dissemination?

The overall dissemination and engagement strategy implemented in the AM Project was extremely successful. I have never seen or heard of a more effective one in an OLT sponsored project. In macro terms, the main evidence of the strategy's success lies in the following results.

- The very broad reach of the Project including direct impact in 17 Australian HEPs, the growing of an interested parties network of 800+, active positive

engagement of key entities including government agencies, professional associations and employer groups, and increasing interest from respected scholars in the field of assessment

- Real academic development and curriculum change effects in an increasing number of accounting programs and domains (and in other disciplines)
- Explicit and willing commitments from several business schools and important professional associations to maintain the momentum of the development and change work of the Project

Despite good practice guidelines being provided by the OLT, historically it has often been the case that OLT project leaders and teams have conceived dissemination in quite simplistic terms. Essentially, the most common perspective adopted has been that dissemination equals *passing on of information*. Often, lots of information is 'passed on', with variable effectiveness to a number of audiences using some variety of means (e.g. conferences, documents, publications, web-sites). Nevertheless, the goal seems to be limited to the spread of undifferentiated information. Other valuable strategic goals, such as the seeding of action or the mobilizing of deeper engagement are usually neglected.

The Project was guided by a sophisticated approach to dissemination. First, it was seen as an integral part of implementation and methodology and as a critical success factor. Second, dissemination, engagement and the achievement of change (impacts on beliefs, learning and practice) were viewed as being inextricably linked. In reality, the Project had a *dissemination-engagement-change strategy* at its core, not simply a dissemination strategy. Third, the achievement of active stakeholder engagement in the Project or with its agenda, in whatever ways were sensible and possible, was an overarching goal for dissemination.

There are a number operational aspects of this strategy that warrant highlighting.

1) Getting accounting academics from 17 HEPs nation-wide and other key people (e.g. officers from professional bodies; leaders and practitioners in organizations) to deeply engage with the learning, values and outcomes of the Project. With encouragement, this enabled knock-on effects, for example, people returning to their organizations and engaging others.

2) Differentiating stakeholders (e.g. business school academics, university senior leaders, professional associations and peak bodies) and ensuring that information and activities provided by the Project matched the needs of these parties and the potential for them to influence the advancement of its cause.

3) A very active campaign to build a network of interested parties and keep them informed and engaged in work and/or agenda of the Project (at last count more than 800 people or entities remain actively engaged or interested).

4) Numerous appropriate activities and modes were implemented to achieve comprehensiveness of reach and buy-in. These included: the calibration workshops for deep learning by educators; more than 20 activities with peak bodies (e.g. The ABDC, CAANZ Board); around 200 presentations or organized discussions across universities, conferences and other domains (e.g. AACSB Assessment Conference, AFAANZ Conference); multiple active engagements with related projects (by invitation) for members of the PT or Reference Group; and a continuing line of publications including scholarly works in international journals.

There is no doubt that the AM Project's dissemination-engagement-change strategy has been very successful for engaging relevant people with its cause and outcomes. Given the challenges of the Project, this achievement is substantial evidence of its merit and value.

Evaluation Question 4: In relation to its principal outcomes and objectives, to what extent has the Project succeeded in generating positive engagement, buy-in and/or support amongst key stakeholders, particularly:

- (a) the Australian higher education community in Accounting; and
- (b) the two key Australian professional bodies in Accounting, ICAA (now CAANZ) and CPAA?

It is not unusual for there to be overlaps between the foci of summative evaluation questions. All are usually concerned with related aspects of project success. In this case EQ3 (above) and EQ4 also have a clear logical connection. EQ3's focus had both process and outcome aspects, whereas EQ4 is primarily outcome oriented. In both cases the outcome domain of interest is engagement and impact. In most change-oriented project contexts, the more effective dissemination is, the more likely it is that engagement and buy-in goals will be achieved well. As a result of the success of the Project's dissemination work, my answer to this question was easy to arrive at.

Because of the wide and constructive reach of the Project there are numerous indications that the Australian HE community in accounting is positively engaged with its agenda and stage outcomes. For example, accounting program communities in most of the 17 participating HEPs are continuing to review and refine their assessment practices and the ABDC is maintaining its peak body support. Importantly, the two key professional bodies, CAANZ and ICAA are actively supportive and continue to draw on the results of the Project to foster greater alignment between

HEPs' programs and their accreditation and certification requirements. Remarks made as early as 2011 by Professor James Guthrie, then Head of Academic Relations for CAANZ, indicated the commitment of that professional body.

For the accounting profession it is critical as the project offers an objective unbiased way to get feedback on the education standards being achieved across the learning outcomes instead of relying more on self-reporting by providers indicating they are covering the various generic skills in various units. The benefits are graduates from accredited accounting programs will have the minimum levels of skills both technical and non-technical to make them employable in the industry. (Campus Review, August 2011).

It is also noteworthy that the Project has succeeded in constructively engaging academic communities beyond accounting (e.g. finance; marketing) and several other HE stakeholders that are important for facilitating improvement in practice in relation to standards-referenced assessment (e.g. the HESP; OLT; TEQSA). Indications of the quality of engagement and buy-in achieved can be gleaned from responses provided by educators involved in the Project (in interviews with the author) and those offered by other parties. The following are a few examples.

It (the Project) has provided me with a license to have good hard conversations with people at my university....and vindication for changes we need to make. It has been proactive for helping to deal with external demands (AQF and the like). I'm impressed. (Senior academic, university business school)

For us it's great to have a methodical approach to learning outcomes and standards and how to deal with subjectivity...This has been practice changing and a great platform for curriculum design. ..The developments it is enabling are very important.....I hope we can keep it going. (Senior academic leader, private non-university HEP).

The community of learning is great - there's a common goal to improve student experiences and outcomes ...and it reduces the sense of isolation felt in our own places. I go back to my university with greater legitimacy for advocacy and the development and improvement role I have....It has provided a great scaffolding. (Senior academic, university business school).

The increasing reach of the Project beyond BE is indicated by the following remarks.

Your account of how the Business Deans have gone about articulating what the TLOs mean and how they might be applied was inspirational to us scientists.

Your methodology could easily become a national model. It was clear from Alan Robson's opening address that your approach aligned well with his thinking, which was another feature attractive to us. (Professor John Rice, Executive Director, Australian Council of Science Deans, February 2013)

One of the more powerful general indications of the level and quality of engagement facilitated by the Project was reflected in some remarks made to me at a post-workshop dinner in July 2014. At the time there were about forty people present, mainly accounting educators from across Australia.

Five years ago I would have known to some extent maybe ten people in the room. Because of the project, now I really know just about everyone ... and we're all learning from each other and getting good things done. (Senior academic, university business school).

Evaluation Question 5: To what extent has the Project enabled valuable learning/professional development for those engaged with its activities?

Learning and capacity building was one of the explicit objectives of the Project. However, I suspect it was never anticipated that such an extraordinary range and depth of learning would be achieved. Before citing some more concrete markers of the learning enabled by the Project, I offer some summary comments.

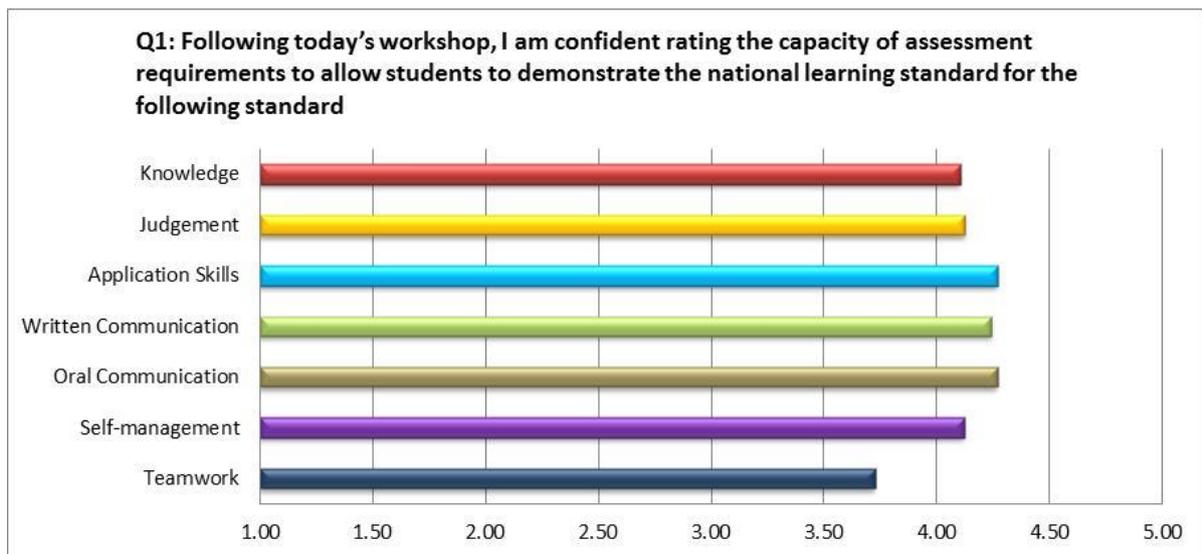
Having been in and around learning environments and activities for more than half of my already long life, like others with similar experience I have developed a sense that seems to detect when substantial in-situ learning is happening and to some extent what kind of learning. I had limited exposure to the workshops, team meetings and other communal activities, however this did run to more than fifty hours of observing and listening. To summarize my impression of the learning that was occurring, the best I can do is borrow the cliché – “it was palpable”. Many good signs of a real learning community were nearly always evident: the ‘hum’; intense focus and concentration; regular questioning; fierce debate; full participation; active problem solving; contributions that indicated critical reflection and creativity; sharing by individuals of changes in practices linked to their earlier learning in the Project; and explicit acknowledgments by individuals of shifts in their understandings, attitudes and/or practices.

Learning and development effects were most evident in the following areas.

- Understanding of assessment concepts and principles and application of these to the enhancement of policies, practices and assessment task development

- As an important sub-set of the above, much enhanced knowledge and capability in relation to the establishment and implementation of authentic standards-referenced assessment practices
- Evaluation, design and improvement of student learning experiences and approaches to teaching at unit of study (subject) and program levels
- Improved self-understanding, particularly in relation to mental models, personal philosophies and assumptions, and how these influence practices

Feedback from the educators who participated in the calibration and review workshops was very positive overall and in relation to the value of the learning they achieved. The following table, which provides a synthesis across several workshops of participants' feedback on one aspect of their experience, is indicative.



Workshop participants that I interviewed in 2014 were clear that valuable learning was a take-out for them. The following are typical of the views expressed.

I have gained hugely from it (the Project). It is transforming assessment processes where I am, particularly at course level. ...I'm a better academic and teacher for it...I'm on fire about it. (Senior educator, non-university HEP)

The workshops have created a community of learning with a common goal to improve students' experiences and outcomes...The learning for me has been very valuable...In our department we are redeveloping two postgraduate programs influenced by the Project. (Senior academic, university business school)

4. Evaluative opinion in summary

My evaluative opinion of the AM Project could not be higher and I commend it unreservedly to all interested parties.

The Project has achieved outcomes aligned with and beyond its objectives, including substantial impacts at a national level on thinking, understandings and practice in HE assessment. In addition, its value is being increasingly recognized internationally and remarkably this has all been achieved in about three years and in the face of several very substantial difficulties.

The AM Project provides an exemplary case on what can be achieved by modestly resourced educational development initiatives (and valuable guidance for people working on projects). Of course, *what is possible* depends on wise determination of strategy, particularly critical success factors, and consistently implementing project work in ways that align with these. The AM Project did this extremely well. Its success resulted mainly from outstanding leadership, insightful strategy, a very good project team, and a highly strategic and well implemented approach to dissemination and engagement.

Implementing processes of the kind developed by the Project on a wider basis, which is what is needed, will require the maintenance of change-oriented momentum over the next few years and the investment of some resources. If these occur, the pay-off for key HE stakeholders could be high. Leaders and educators in universities and other HEPs will gain a sound and workable methodology for dealing authentically with many of the challenges of standards-referenced assessment. The knowledge and practice capital generated by the Project is already proving to be valuable for the HESP, TEQSA, and important professional bodies such as CAANZ and ICAA.

At a more overarching level, if the use of processes similar to those developed by the AM Project became the norm across the HE sector, Australia would have a much stronger basis for its claims about taking standards seriously and having the means

in place and evidence to demonstrate this. The potential benefit for Australia's HE 'brand' reputation and those of individual universities and disciplines is high. The possible costs of a prospective decrease in reputations, if these kinds of processes are not adopted, need also to be factored into decision making.

I advocate strongly that interested parties provide as much material support as possible for well-conceived initiatives or entities that will sustain the development and change momentum generated by the AM Project.

References

Balla, J. & Boyle, P. (1994). Evaluation of student performance: a framework for improving practice. *Assessment and Evaluation in Higher Education*, 19 (1), 17–28.

Baume, D., Yorke, M. and Coffey, M. (2004). What is happening when we assess, and how can we use our understanding of this to improve assessment? *Assessment and Evaluation in Higher Education*, 29 (4), 451–477.

Price, M., O'Donovan, B., Rust, C. and Carroll, J. (2008). Assessment standards: A manifesto for change. *Brookes eJournal of Learning & Teaching*, 2(3).
http://bejlt.brookes.ac.uk/article/assessment_standards_a_manifesto_for_change/

Rust, C. (2014). Are UK degree standards comparable? *Times Higher Education*.
<http://www.timeshighereducation.co.uk/features/feature-are-uk-degree-standards-comparable/2016838.fullarticle>

Sadler, D.R. (2009). Indeterminacy in the use of preset criteria for assessment and grading in higher education. *Assessment and Evaluation in Higher Education*, 34, 159–179.

Sadler, D.R. (2014). The futility of attempting to codify achievement standards. *Higher Education*, 67(3), 273–288, DOI 10.1007/s10734-013-9649-1.

Watty, K., Freeman, M., Howieson, B., Hancock, P., O'Connell, B., de Lange P. and Abraham, A. (2014). Social moderation, assessment and assuring standards for accounting graduates, *Assessment & Evaluation in Higher Education*, 39 (4), 461–478.