Implementing Programmatic Assessment across the Continuum of Medical Education

Case Studies

2017

Contents

INTRODUCTION

Assessment is integral to education programs across the continuum in medical schools, the prevocational years, specialist training and in the assessment of International Medical Graduates (IMGs).

Why is assessment important?

Assessment is the mechanism by which the education provider determines the ability of individual members to meet specific milestones of the training program and ultimately measures readiness for unsupervised practice.

What are AMC standards on assessment?

The key concepts underpinning AMC standards on assessment for training providers across the continuum support contemporary approaches to assessment. The standards are:

- **Assessment approach:** The assessment program is aligned with learning outcomes, with requirements clearly documented and easily accessible to staff, supervisors and students/trainees/interns.
- **Assessment methods:** The program contains methods that are fit for purpose, has a blueprint to guide assessment through each stage and uses validated methods of standard setting.
- **Assessment feedback:** The provider/program facilitates regular feedback to students/trainees/interns to guide their learning, gives feedback to supervisors on assessment performance and has processes for underperforming students/trainees/interns and implementing remediation.
- **Assessment quality:** The provider regularly reviews its program of assessment to ensure the validity and reliability and scope of its practices, processes and standards is consistent across teaching sites.

Why are approaches to assessment changing?

In recent times, the field of medical education assessment has undergone significant change. This change is linked to the adoption of competency-based approaches to medical education whereby supervisors are required to make decisions about the learner's competence across a range of pre-determined standards (Ten Cate and Scheele F 2007). Supervisors require a large amount of information to support these important decisions about competence and progress.

The emphasis on assessment for learning has highlighted the shortcomings of assessments based solely on high stakes examinations. Such examinations do not provide the nuanced information required to have full confidence in the accuracy of assessment decisions, particularly on the assessment of professionalism and actual real world ability (Rethans J, Norcini J, Báron-Maldonado M, et al. 2002; Creuss et al 2006). This has seen an increased emphasis on work based learning and assessment (Norcini J, Blank LL, Arnold GK, et al. 1995; Govaerts MJB, Van der Vleuten CPM, Schuwirth LWT, et al. 2007). It also features new thinking about how multiple data points from formal exams and regular work-based low stakes assessments can be synthesised as a program of assessment to make progression and high stakes decisions on performance and work readiness (Van der Vleuten CPM, Schuwirth LWT. 2005; Van der Vleuten CPM, Schuwirth LWT, Driessen EW, et al. 2012.)

How can we get assessment to work well on the ground?

Given the scope of these changes there is also a recognition that improved implementation is paramount to the success of assessment innovation. This includes incorporation of change management strategies including co-design, broad consultation, communication and supervisor and assessor training.

Programmatic Assessment, first proposed by leading medical educators Profs Cees van der Vleuten and Lambert Schuwirth is a useful term which encapsulates the key concepts underpinning newer ways of thinking about medical education assessment.

Programmatic assessment is an integral approach to the design of an assessment program with the intent to optimise its learning function, its decision-making function and its curriculum quality-assurance function.

A key component of programmatic assessment is the separation of data from decisions, that is, not all assessment episodes need to be accompanied by a summative decision. Instead, high stakes decisions are made only after a sufficient number of observations of a trainee's performance have been gathered and synthesised.

The following representatives been involved in the implementation of programmatic assessment and have provided a case study focusing on providing insights into innovation in assessment:

1. **Professor Cees van der Vleuten:** Programmatic Assessment in the Netherlands
2. **Professor Lambert Schuwirth:** Programmatic Assessment Pilots at Flinders University
3. **Ms Christine Cook:** Programmatic Assessment in GP training
4. **Professor Elizabeth Molloy:** Building feedback and assessment literacy: A national professional development initiative (Project Funded by the Office of Learning and Teaching, Melbourne, Monash and Deakin Universities)
5. **Professor Kichu Nair / Dr Beth Mulligan:** Programmatic Assessment for IMGs
6. **Professor Tim Wilkinson:** Programmatic Assessment at University of Otago
7. **Dr Marie-Louise Stokes:** Programmatic Assessment in Physician Training.

This series of case studies focus on how a range of national and international medical education providers has drawn on newer thinking about assessment, including programmatic assessment to improve their assessment practices. Focusing on implementation, the case studies provide a brief overview of the context, purpose, design, implementation and evaluation of assessment across the medical continuum.

CASE STUDIES

Case Studies presented at the AMC Workshop, 1 November 2017:

	**Professor Cees van der Vleuten**	Programmatic Assessment in the Netherlands
	**Professor Lambert Schuwirth**	Programmatic Assessment Pilots at Flinders University
	**Ms Christine Cook**	Programmatic Assessment in GP training
	**Professor Elizabeth Molloy**	Building feedback and assessment literacy: A national professional development initiative (Project Funded by the Office of Learning and Teaching, Melbourne, Monash and Deakin Universities)
	**Professor Kichu Nair**	Programmatic Assessment for IMGs
	**Dr Beth Mulligan**	Programmatic Assessment for IMGs

Additional Case Studies:

	**Professor Tim Wilkinson**	Programmatic Assessment at the University of Otago
	**Dr Marie-Louise Stokes**	Programmatic Assessment in Physician Specialist Medical Education

Case Studies Presented at AMC 1 November 2017 Workshop

Programmatic Assessment in the Netherlands

Professor Cees van der Vleuten

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

4-year graduate-entry program where students come in with a relevant bachelor program. On top of their medical degree, they receive a master in clinical research. Clinical research therefore receives abundant attention in the curriculum. The Netherlands has adopted the Canmeds competency framework on a national basis. The curriculum is structured according to this framework. Moreover, the first 2 years is based on the educational philosophy of problem-based learning (PBL). The cohort size is 50 students. In the first, this year consists of more classic PBL using written problems as a starting point for learning. In the second year, real patients seen in outpatient clinics are the starting point for learning. The third year consists of clinical rotations, while the fourth year consists of a long (6 months) elective clinical placement and a long research elective. The Netherlands has no national licensing examination, so each university is responsible for the quality of its own graduates. Six out of the country's eight medical schools use progress tests, which they collaboratively produce and simultaneously administer.

PURPOSE

PBL intends to promote self-directed learning. However, in a classic system of modular units with end-of-unit exams, self-directed learning is hampered. "Teacher control" enters through the back door, by an end-of-module testing approach. To support and stimulate self-directed learning the assessment program needs to be more longitudinal and more informative in that it provides ongoing

and constructive feedback. The uptake of the feedback is promoted through reflections that form the basis for creating a learning dialogue between student and a mentor.

We realised programmatic assessment to a large extent by:

- removing any summative decision-making linked to an individual assessment
- introducing mentors who follow the students during the entire training program
- introducing an electronic portfolio that accumulates all data points and supports the assessment processes
- ensuring proportionality by only making high stake decision-making at the end of the year and intermediate evaluations during the year

DESIGN

There are four key elements in the design:

1. Module-related assessment in different formats (MCQ, essay-type, assignments, orals)
2. Longitudinal assessment: in the cognitive domain, progress tests (4 times per year); in the behavioural domain through periodic self, peer and staff assessments.
3. Mentoring system with regular meetings between mentor and students which are prepared and minuted by the students
4. End-of-year assessment in all competencies by an assessment committee using competency standards fail, pass or distinction (all ECTS credits are handed out with a pass or distinction). Mentor provides a recommendation but has no vote in the decision-making itself.

IMPLEMENTATION

The very first implementation failed because we maintained the summative module-related assessment. Once we removed this, the implementation proceeded much better. Ensuring buy-in from teachers was critical. Ensuring student receive sufficient meaningful feedback in the assessment program remains a challenge. Electronic tools for feedback provision work really well. Mentors find considerable meaning in the strong relationship with students and vice versa; mentors are very highly appreciated by the students. The decision-making is without problems. We have developed a highly efficient procedure for decision-making by deliberating more extensively only for those students who need deliberation.

EVALUATION

There are several PhD students currently studying programmatic assessment. First evaluations have been the following publications:

1. Bok, H. G., Teunissen, P. W., Favier, R. P., Rietbroek, N. J., Theyse, L. F., Brommer, H., ... & Jaarsma, D. A. (2013). Programmatic assessment of competency-based workplace learning: when theory meets practice. *BMC Medical Education*, *13*(1), 123.
2. Heeneman, S., Oudkerk Pool, A., Schuwirth, L. W., Vleuten, C. P., & Driessen, E. W. (2015). The impact of programmatic assessment on student learning: theory versus practice. *Medical education*, *49*(5), 487-498.

These studies reported successful implementations, but also found that students tend to see low-stake assessment as high stake. In a recent follow-up study (not yet published) it turns out student's agency on the assessment was a key determinant for the perception of the stakes and the way

staff interacted with the students. For example, in a setting where grades were removed from the assessment, the feedback culture was much stronger and the stakes much lower.

1. Heeneman, S., & de Grave, W. (2017). Tensions in mentoring medical students toward self-directed and reflective learning in a longitudinal portfolio-based mentoring system–An activity theory analysis. *Medical Teacher*, *39*(4), 368-376.

This study showed that mentoring skills are positively influenced by a longitudinal mentoring relationship and presence of an active mentor community, but contextual factors such as the assessment procedure and how the portfolio is used, can lead to tensions in the mentoring process, and less self-regulation of learning by the mentee.

2. Heeneman, S., Schut, S., Donkers, J., van der Vleuten, C., & Muijtjens, A. (2017). Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Medical Teacher*, *39*(1), 44-52.

This study showed that students in programmatic assessment use the online feedback significantly more than students who are not in such a system. The study also demonstrated large knowledge differences at the end of training in favour of a program using programmatic assessment.

Finally, here are a few quotes that come from the qualitative inquiries from several studies:

> *"If I don't understand something during an oral, I can ask questions straight away like: but what do you mean by this or how exactly does that work? . . .it's more that you get more feedback like: OK, this isn't quite right, this works like this, and that makes you think: all right, OK, or I did it wrong, or yes of course you're right, but then you know where you stand."*
>
> *"Often you think that without the portfolio it would be easier to think: OK good, I've gotten this feedback, fine, we'll tuck it away there. But now you are challenged to link feedback, look for common threads but also for conflicting points and to think about them. I think it's useful, particularly for this purpose."*
>
> *"I have lots of different people talking about my professionalism and so, each one has a less importance, it's okay to get some negative feedback, because you have a lot so you have some negative and some positive, just, you know that's how it tends to balance out."*
>
> *"In undergraduate studies a test would make me nervous and anxious and worried about how well I was going to do. But because these assessments don't have the same consequences [receiving grades], because they are just to help me identify what to study, I don't feel the same nervousness that I did before. It's the good without the bad."*
>
> *"I feel very comfortable looking stupid"*

FUTURE FOCUS

Scaling up to other programs.

1. The master phase in our undergraduate-entry medical course program in medicine (cohort size 300) has recently been remodelled and uses programmatic assessment.
2. In the bachelor phase of the undergraduate entry programme programmatic assessment is being applied to learning clinical skills in the Skillslab Starting a discussion on the utility of grades.

Programmatic Assessment Pilots at Flinders University

Professor Lambert Schuwirth

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

Flinders university medical school a 4-year graduate entry programme leading to an MD. In its curriculum there is a focus on longitudinality. In the advanced studies programme, the research component of the MD curriculum, student conduct or participate in a longitudinal research project, starting in year one and finishing with a capstone event in year four.

Flinders University medical school was also the first in Australia to develop a longitudinal integrated clerkship (LIC) for the rural and remote areas. This very successful pathway was similar to many such developments around the world. Currently the experiences of this are being used for a longitudinal integrated clerkship approach for the clerkships at Flinders Medical Centre, meaning that all medical students do their clerkships now in a LIC format.

PURPOSE

Although the curriculum design has always been innovative with the first graduate entry, one of the first MD programmes and the development of rural LIC and urban versions of an LIC the assessment system did not follow. The assessment programme still relied heavily on traditional barrier examinations and OSCEs. At the same time the assessments in the final year were loosely organised and diverse.

A second perceived problem was the inability of the assessment system to pick up early warning signals for students whose development of competence and professionalism was at risk. Too often individual signs could not followed up or handed over to the coordinators of a next study phase. In some case this led to more serious problems in the more senior years, which the staff felt could have been remediated earlier on.

There was a felt need to modernise the assessment system and use the latest form of evidence about the nature of medical competence and professionalism and how to assess it.

DESIGN

A new assessment system was designed using the principles of programmatic assessment for learning (PAL). Central in this system is progress testing. Progress testing serves as a backbone assessment with its focus on functional and relevant knowledge. In progress testing all students of all year classes sit the same test and four different tests are created and administered four times per year. Because the test is attuned to graduate level knowledge and could in principle contain questions from all domains of medical knowledge students cannot revise for the test. Therefore a decoupling of assessment moments and decision moments takes place. Progress testing is not new, it was developed at Maastricht University and Kansas City University in the 1970s. Yet the concept fits the concept of programmatic assessment for learning well; it is a form of longitudinal assessment and students are given ample feedback so they can use the results and the feedback to steer their own learning and remediate when necessary. In addition, there are numerous smaller tests (we employ a TBL system with frequent individual and group assessments that provide the students with frequent feedback), and there are numerous occasions at which feedback is given.

All this information is collected and collated in a portfolio; the students write self-analyses, which are discussed with a mentor on a regular basis.

IMPLEMENTATION

It has taken very long before the actual implementation could start. Gradually the organisational culture at Flinders University had changed from an innovative/experimental to a more conservative and risk averse one. This has had its influence on the medical school as well. Planning for PAL has therefore taken about 5 years. Visits from international experts (Cees van der Vleuten, Albert Scherpbier, Olle ten Cate) have helped enormously because they were able to tell about successful implementations elsewhere with good results (Cleveland, Maastricht, etc.). A detailed planning was made and an estimate of the necessary investments and returns on investment.

The implementation has just started in year 1 and is now almost 11 months underway. The medical students in years 2-4 are still studying mostly under the traditional assessment regime, although the barrier examinations have been removed and replaced by the progress test for all students.

The implementation has been successful and there is considerable excitement about being innovative in education. Yet, some staff and some students are unhappy about it, they fail to see the advantage and are very vocal about this. Their dissatisfaction sometimes leads to deliberate misinformation. Yet, this is not new and has been encountered in many educational changes – such as PBL for example – in past decades.

EVALUATION

No formal evaluation of the learning effects has taken place yet, but there is early anecdotal evidence of students being flagged earlier on in the curriculum. Those students receive support to help them set up their own remediation plans. Staff indicate that they are relieved to finally have to possibility to voice their concerns with the knowledge that it will lead to concrete actions to support those students in difficulty.

Not all staff are happy. It is clear that this is a change that is huge for some and that it is difficult to change thinking from ‘assessment as a stick to drive unmotivated and lazy students to studying’ to ‘assessment as a companion for students to help them plan and gradually take agency over their own assessment’.

At this moment, it is still too soon to be able to produce any evidence about the usefulness and effectivity of the system.

FUTURE FOCUS

We hope to be able to develop the PAL programme further and extend it to all four years. In addition we hope to be able to have a system in which the responsibility and control over the assessment will gradually be handed over from the medical school to the student, so that the student is optimally prepared to take responsibility for their own lifelong learning after graduation. We would consider the programme ideal if the 'final examination' would be a procedure in which the student presents and defends their portfolio an 'applies' for graduation like for a job or internship. So future developments would focus at having a process of handing over responsibility.

Programmatic Assessment in GP Training

Ms Christine Cook

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

General Practice vocational training consists of a hospital year, followed by 2-3 years of training in community practices. Following successful completion of the training program and successful performance in fellowship examinations, registrars are awarded with fellowship from either the Royal Australian College of General Practitioners (RACGP) or the Australian College of Rural and Remote Medicine (ACRRM). Training is delivered by regional training organisations to the standards of the two colleges.

GP365 is a vocational general practice training program for registrars. It is a programmatic assessment system developed by ModMed (formerly Sturt Fleurieu Education and Training) and The Prideaux Centre, Flinders University. The program is currently delivered by regional training organisations in both South Australia (GPEx) and Western Australia (WAGPET) to over 1000 registrars in urban, rural and remote locations. Registrars are able to undertake the program either full-time or part-time.

PURPOSE

In 2013 the RACGP began transitioning to outcomes-based vocational training standards that required all registrars to be signed off as “exam ready” by their regional training provider, demonstrating that they had successfully completed the training program and were ready to be assessed by the colleges (sit the fellowship exams) as safe independent practitioners. This led to a review of the training program, with the guiding questions ‘how do we know they are safe?’ and ‘how do we know they are developing general practice competencies?’

In 2014, the federal government began a reform process to drive efficiency across Australia, seeking to reduce the number of training organisations and contract funding, while increasing the number of doctors in training. Sturt Fleurieu Education and Training embraced the reform process to review their training delivery model, seeking efficiency and scalability without compromising quality.

With multiple stakeholders involved in the training of registrars, and large amounts of information collected, information often resided in many locations, not easily accessible or updatable and impacting upon transparent and consensus driven decision-making. Additionally, given much of the education and assessment was occurring in multiple training posts, RTOs needed to ensure there was a consistent and minimum standard of teaching and feedback.

RTOs require robust assessment and monitoring systems to ensure that struggling registrars are identified early and appropriately managed. At the same time, RTOs want to ensure continuous learning throughout training with a focus on developing registrars into safe, reflective independent practitioners.

A programmatic assessment program *for* learning (PAL) with multiple data points to make robust high-stakes decisions about exam readiness, integrated with ModMed's long standing education program provided the solution for this issue.

DESIGN

The key elements of GP365 are:

1. It is a longitudinal assessment program which continues throughout the registrar's entire community based vocational training, regardless of duration.
2. The assessment framework is integrated and woven into the educational components and for the most part relates to the individual patients that a particular registrar is seeing in a particular region. This allows it to be relevant for both urban, rural and remote registrars.
3. The program enables closer monitoring of registrar progress / evidence of continuous assessment and provides opportunities for meaningful feedback to registrars.
4. Early identification of registrars at risk of professional or clinical issues.
5. The ability to individualise learning to a registrar's learning needs. Registrars are able to complete many of their education and assessment requirements in any order to ensure they can tailor this to "what walks through the door". They are also able to tailor the time spent on different curriculum areas/competencies to their pre-existing knowledge and strengths and remain in control of their learning.
6. Each assessment is designed to promote learning. After each assessment registrars develop learning goals to work on the areas of weakness identified in the assessment. This promotes the development of reflective life-long learners.
7. There are multiple different assessments (direct observation visits, video reviews, MCQ, multi-source feedback, entrustable professional activities, clinical case analysis, small group work and key clinical assessments) to ensure all areas of the curriculum and college domains and competencies are being assessed and developed.
8. The program is scalable with no limit on the number of registrars being able to undertake the program.
9. The program is efficient as it provides a central location for information relevant to the registrar's learning and progress, automates workflows, and reduces the load on medical educators and administrators.

10. A consistent approach to the critical education topics required for general practice training, a teaching tool for supervisors and learning tool for registrars.
11. The online learning environment supports longitudinal programmatic assessment. The programmatic assessment system brings together multiple data-points, combining words with numbers to create meaningful outcomes, in real-time and with ease.

IMPLEMENTATION

The PAL principles adopted by GP365 were fundamentally different to the traditional approach to education and training.

The implementation of GP365 required a systematic approach to change management, risks, technology, business processes, training, stakeholder engagement, communications, understanding roles, conflict resolution, and well-being.

In this case study, we want to elaborate on the educational change lessons from a successful implementation in a General Practice (GP) training setting.

Lessons learned

Two aspects were essential in the change to PAL.

First, traditional views on assessment run deep with stakeholders and need to be taken seriously. Activities to actively change the organisational culture and to educate all stakeholders were helpful in adapting the stakeholder's intuitive ways of assessing registrars into a more formal structure.

Second, continuous focus on the principles of PAL and their implications for assessment design was needed to ensure that the transfer of PAL theory was correctly translated into practical assessments. Even though there was a considerable buy-in of the concept of PAL, the translation into the actual design and application was found to be a cognitive theory-to-practice translation problem.

Conclusion

The most important barriers with the implementation of PAL in a difficult-to-control context were not the buy-in of the concept or the knowledge about PAL, but the necessary change to the organisational culture and the transfer problem in the theory-to-practice translation.

EVALUATION

Formal evaluation GP365 has commenced and outcomes will be published in 2018.

Some early indicators for GPEx have shown that at a local level, registrars have obtained higher pass rates in the fellowship exams than other RTO registrars. Satisfaction ratings by registrars are very high and the withdrawal rate from the program is about 5%. Supervisors are also very satisfied with the program and no supervisors withdrew during the transitioning period.

A recent survey of GPEx registrars found that:

- 90% of registrars receive meaningful feedback from their supervisor often or always
- 95% of registrars agree to strongly agree that their supervisor is aware of their learning needs
- 96% of registrars agree to strongly agree that their supervisor assists them to address their learning needs.

Feedback from registrars:

"The overall training program your team has put together is so far the best education program I've ever done. And I've completed a few! I particularly love GP365 and the way we use a self-directed

learning approach to access your treasure trove of useful information as we go as it is immediately relevant to a patient we've seen. The effort that has gone into this is phenomenal."

"GP365 is a very useful tool and should be available to all registrars. The reading modules are very informative; information about health assessments and item numbers has been helpful and; the links to current guidelines help me review my practice."

"I am so impressed with GP365 - I actually love it. It is the best learning platform that I have encountered in the past 8 years of my medical education. I feel that I am learning so much and it is laid out in a sensible way. I really wish that med school had something similar, I would be a better doctor right now if they had. I have no doubt that if I keep working through GP365, I will pass my exams with no problem."

Feedback from supervisors:

"The online nature of GP365 has the ability for the registrar, supervisor and medical educator to asynchronously input data independently – and all at the touch of a button on my phone."

"I have found it easy to navigate through the modules and assignments. The notifications when a task has been graded and comments provided are useful, and the calendar is a plus too."

FUTURE FOCUS

Our future focus is on optimising GP365, building capacity of the delivery teams to master feedback for learning, supervision, and meaning making.

We will continue to innovate with our digital thinking and custom online learning environment to have a fully optimised programmatic assessment system that is relevant to the continuum of lifelong learning.

Building feedback and assessment literacy: A national professional development initiative (Project Funded by the Office of Learning and Teaching, Melbourne, Monash and Deakin Universities)

Professor Liz Molloy

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

Feedback is one of the most problematic aspects of the student experience in higher education worldwide (Boud & Molloy 2013, Hattie & Timperley, 2007). Despite a substantial body of literature exploring feedback as part of assessment, and countless models and strategies aimed at improving practice, feedback remains a consistent source of dissatisfaction for both students and educators. However, data also suggest that there are pockets of excellence in feedback across Australian higher education. What are these teaching teams doing that is so good, and what are the conditions enabling success?

This 2 year project, funded by the Office of Learning and Teaching (OLT), employed a large scale survey of both learners and educators across 2 institutions along with focus groups and interviews. From these data, 7 case studies were selected to illuminate characteristics of effective feedback designs. A resultant framework for 'feedback for learning' was produced with 14 conditions across 4 categories (including concrete strategies to achieve each condition).

The project outputs include a website with all resources, and a national roadshow, which engaged educators across faculties and intuitional leaders in discussions about how to better engage learners in productive feedback.

PURPOSE

Feedback is critical for promoting student learning in medical education (Molloy and Boud 2013). There is consensus in the literature that feedback is inseparable from the learning process, with notable researchers arguing that feedback is the most powerful single influence on student achievement (Hattie & Timperley 2007; Hattie 2009; Orrell, 2006; Orsmond & Merry, 2011).

Without feedback students are limited in how they can make judgements about their progress, and how they can change their future performance. However, despite the potential of feedback to improve learning, feedback is under-utilised and often misunderstood by both students and educators. Educators are concerned that students will be upset by feedback comments, and are frustrated that students often do not take their feedback comments into consideration. Feedback regularly receives the lowest course satisfaction ratings in student surveys, with students reporting that feedback is meaningless, confusing, and too late to be useful.

This research-based initiative is designed to improve student learning (and experience) through improving institutional, educator, and student capacity to generate and use assessment feedback. A set of resources have been developed, and a national roadshow was implemented (October 2017) to engage educators and university leaders in discussions about how to improve feedback in their own context.

DESIGN

We surveyed approx. 4,500 students and 400 staff from two universities to identify units of study that demonstrated effective feedback. From these we chose seven cases both staff and students thought were effective. We spoke with 20 students and 14 staff across these cases about their feedback practices and looked at the course designs and materials. We analysed these cases in terms of what was enabling effective feedback, and identified 14 conditions for success, grouping them into four categories. Our result definition of feedback is:

"Feedback is a process in which learners make sense of information about their performance and use it to enhance the quality of their work or learning strategies."

Feedback is successful when…

Capacity for feedback

1. Learners and educators understand and value the purpose of feedback
2. Learners are active in the feedback process
3. Educators adopt an inquiry mindset
4. It is understood to involve a complex set of practices improved over time

Designs of feedback

5. It is aligned with the learning outcomes
6. The performance information is usable and learners know how to use it
7. It is adapted to meet the needs of diverse learners
8. It uses a variety of sources and modalities
9. It involves multiple instances over time for any given outcome

Culture of feedback

10. It is a valued institutional enterprise
11. There are processes in place to ensure consistency and quality
12. There is continuity of vision and commitment from leaders and educators

Resources for feedback

13. Educators have flexibility to deploy labour to best effect
14. It is supported by space and technology

The 7 case studies are available on the webpage under the tab “Feedback Possibilities”.

IMPLEMENTATION

1. Half day educator workshop

Description: The half-day workshop focused on how to improve feedback practices in higher education. The project team shared findings from the multi-institutional study examining ‘good’ feedback practices and designs, including an empirically-derived framework that may help to guide the development of programmatic feedback. Participants had opportunities to consider how the guidelines compare with current practices, and we discussed ideas relating to implementation across different practice contexts.

Who attended: People in charge of feedback in a unit (unit coordinators/chief examiners)**,** educational designers/academic developers**,** course leaders**,** program leaders**,** Associate Deans/Associate Heads of School (Teaching and Learning).

Learning Outcomes:

Change conception of feedback and its importance

- Describe a more productive approach to feedback
- Identify and understand current problems and tensions in feedback practice within their context

Awareness of any extended range of feedback practices, beyond present repertoire

- Focus on common approaches to feedback in their context
- Identify any key tensions and strengths in their own institutions that may influence feedback practices
- Identify and plan next steps in terms of improvements in their own feedback contexts

Awareness of key issues in improved feedback practice and how to influence others

- Identify institutional/departmental factors that might improve learner and teacher engagement in feedback
- Share project resources with colleagues and use to improve practice

2. Roundtable Discussion with Institutional Leaders

After the workshop, this 45 minute discussion was held between senior leaders of the Institution (eg. PVC, DVCE, Faculty Deans, Faculty Associate Dean L & T or equivalent) and members of the project team. Key challenges and opportunities facing assessment feedback in the current higher education climate were discussed, and the key results and practice implications arising from the large-scale, multi-institutional study between Deakin, Monash and Melbourne universities. As facilitators we shared examples of effective feedback designs currently in use in Australian universities, examined the conditions that can assist effective feedback practices to develop and flourish, and discussed possible policy directions at an institutional level.

EVALUATION

This project adopted a participatory design approach as a way to increase impact and quality. In this approach, up to 300 educators, academic developers, instructional designers as well as leaders in learning and teaching from around Australia have actively negotiated how the project outcomes might work within their institution and courses.

Workshop materials have been collected (participants' group notes on feedback and assessment priorities in their programs/institutions, as well as team members' notes from roundtable meetings with senior leaders).

Workshop participants were provided with an online survey immediately post workshop to gauge level of engagement and perceived utility of the framework and accompanying resources. Participants will also be surveyed in 3 months to evaluate the degree to which they have implemented the goals they identified during the workshop.

These data will be used to inform the refinement of the Feedback for Learning Framework (feedbackforlearning.org). The final shape of this framework will be guided by the data collection and through negotiation with stakeholders (e.g. participants, leaders, project group, and project evaluator). The framework is conceived as a guide for institutions and educators to assist in decisions and planning of programmatic feedback. The framework will include a definition of feedback, a series of challenges to traditional feedback practice, 14 conditions for success (each with examples of strategies to achieve the condition). The 7 illustrative case studies are available on the website with videos, visual representations, and descriptions to illuminate facets of practice that have been seen to be effective.

The project has benefited from the ongoing input from evaluator Professor David Carless, Associate Dean (Learning & Teaching) of the Faculty of Education at the University of Hong Kong.

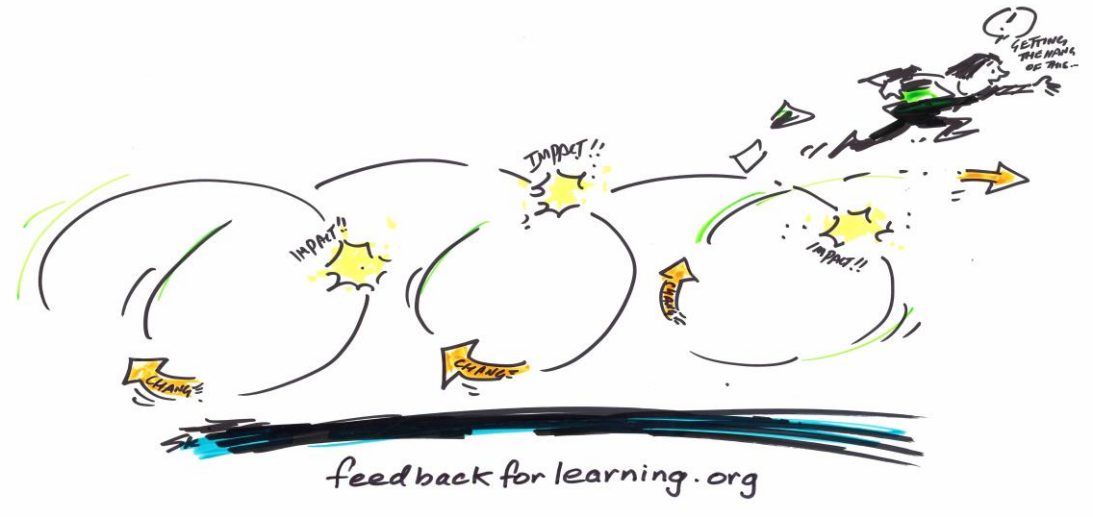
FUTURE FOCUS

We hope that the project will result in highly refined local action plans, as well as improved project outcomes given the framework and resources will be adapted to include new perspectives from participants.

All resources are freely available on the website, and we hope this accessibility to materials will generate motivation for sustainable faculty development, and improved feedback literacy of educators and learners.

Professional learning resources including PowerPoints, lesson plan, worksheets, synopses and links to specific examples within the rich cases of feedback designs, prepared specifically for academic developers and workshop facilitators, will be available on the site to promote self-guided professional learning.

Many of the workshops were oversubscribed (eg n=70 waitlist for Melbourne-based workshop). For this reason, the project team will run a webinar later in the 2017 to increase reach. We hope that feedback practices will shift in accordance with the images below (the latter image reflecting feedback as a process enacted over time, where the information has an effect on learner performance).



References

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education, 38*(6), 698-712. doi:10.1080/02602938.2012.691462

Hattie, J. (2009). The black box of tertiary assessment: An impending revolution. In L. H. Meyer, S. Davidson, H. Anderson, R. Fletcher, P. M. Johnston & M. Rees (Eds.), *Tertiary Assessment & Higher Education Student Outcomes: Policy, Practice & Research*. Wellington, New Zealand: Ako Aotearoa.

Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*(1), 81-112.

Molloy, E., & Boud, D. (2013). Seeking a different angle on feedback in clinical education: the learner as seeker, judge and user of performance information. Medical Education, 47(3), 227-229. doi:10.1111/medu.12116

Orrell, J. (2006). Feedback on learning achievement: rhetoric and reality. Teaching in Higher Education, 11(4), 441-456.

Feedback for Learning: Closing the Assessment Loop (OLT 2016-2017) Team Members: Michael Henderson, Elizabeth Molloy, David Boud, Phillip Dawson, Michael Phillips, Tracii Ryan, Paige Mahoney, David Carless. *Feedbackforlearning.org*

Programmatic Assessment for IMGs

Professor Kichu Nair

Dr Beth Mulligan

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

Approximately 20-30 % of the Australian medical workforce is made up of International Medical Graduates (IMGs). Many of them work as Junior Medical Officers (JMOs) while waiting for the AMC clinical examination, after passing the English and MCQ assessments. Performance issues have been reported amongst IMGs even after passing the AMC examination.

We introduced the workplace-based assessment (WBA) for IMGs in 2010 to assess their performance while working as JMOs waiting to do the AMC examination.

PURPOSE

- To test alternate way of assessing IMGs while waiting to do the OSCE assessments
- To test various tools of WBA , including mini-cex, CBDs, MSF ,DOPS in a longitudinal program
- To test the feasibility of WBA in an Australian context

DESIGN

We devised a program where IMGs are tested on their performance using the above tools in various combinations, over a period of 6-12 months.

The candidates and assessors were calibrated before the assessments to make sure that they were familiar with the tools and criteria

All assessments were blue printed and scheduled from a central office.

IMPLEMENTATION

This was challenging since this was a major change in assessment philosophy and the tools were not familiar for all concerned.

The hospital and health system had to be brought in line with the innovation.

In Newcastle we have assessed 180 IMGs and in Launceston 110 candidates until now.

EVALUATION

Reliability studies, economic evaluation and qualitative studies were done.

FUTURE FOCUS

Streamlining the assessment and measuring composite reliability.

Longitudinal evaluation of the WBA candidates against the OSCE candidates.

Electronic database and portfolios for the assessors and candidates.

Feasibility and testing whether WBA can be introduced for other assessments.

Additional Case Studies

Programmatic Assessment at University of Otago

Professor Tim Wilkinson Assessment at University of Otago

This case study focuses on providing insights into the follow key questions about innovation in assessment:

1. **Context:** Provide a brief background about your institution, your learners and assessors.
2. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
3. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
4. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
5. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
6. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

MBChB programme, focusing on years 4-6 in Dunedin, Christchurch and Wellington, and focusing particularly on the workplace, or in-course, assessments.

PURPOSE

Problem definition

Previously, each of our in-course assessments also required a summative judgment – did the student pass it or not? Where there was uncertainty, an assessor was more likely to pass than to fail. At the end of the year, we then had a series of 'passes' but some assessors were uncomfortable about the standard of some students. Yet, the students of concern hadn't failed anything. We tried including a 'borderline' category, but this just meant we had some students at the end of the year with some borderline assessments, yet we still didn't; know enough about them to make decisions, and these students didn't know what they needed to do to improve.

Solution

We drew on principles of programmatic assessment. In particular:

1. We removed the need for summative judgements to be made on all assessments,
2. We related progress to criteria rather than just teacher judgements and
3. We made decisions only after looking at a student's progress longitudinally.

DESIGN

There were three main components

1. Standards based assessment. That is, criterion referenced, rather than norm referenced and a move from graded passes to pass/fail/distinction
2. Regular student progress committee meetings
3. Use of conditional pass.

Of these, conditional pass was the most novel as many schools have progress meetings and use criterion referenced assessments. For us, the conditional pass system allowed raters to defer judgments on assessments, instead focusing on the learning aspects. It also allowed raters to alleviate their anxiety when uncertain if a student was truly above or below a required standard. The impact therefore on failure to fail was very positive.

IMPLEMENTATION

The student progress committee meetings served not only to make robust decisions on students, but also acted as important professional development for staff – the discussion around the table drew out many of the assumptions of teachers around assessment and allowed for open discussion around these. Over time, there has been increasing alignment among teachers towards principles of programmatic assessment. Secondly, the conditions of the conditional pass require scrutiny. They need to be specific enough to drive learning and to inform summative end-of-year decisions, sometimes these conditions were not specific enough but addressing these deficits contributed to group learning by staff.

EVALUATION

The new system of conditional pass helped in failing students on the basis of professionalism. [1]

The change to standards based assessment improved intrinsic motivation of students. [2]

1. Wilkinson TJ, Tweed M, Egan T, Ali A, McKenzie J, Moore M, et al. Joining the dots: Conditional pass and programmatic assessment enhances recognition of problems with professionalism and factors hampering student progress. BMC Medical Education. 2011; 11(1): 29.
2. Wilkinson TJ, Wells JE, Bushnell JA. What is the educational impact of standards-based assessment in a medical degree? Medical Education. 2007; 41(6): 565-72.

FUTURE FOCUS

This is one way to overlay principles of programmatic assessment onto an existing more traditional assessment system.

We need to evaluate more clearly the impact on learning of these student who get a condition pass – while it clearly leads to tem mostly meeting the conditions, what does it do for motivation?

Programmatic Assessment in Physician Medical Education

Dr Marie-Louise Stokes Royal Australasian College of Physicians (RACP)

This case study focuses on providing insights into the follow key questions about innovation in assessment:

7. **Context:** Provide a brief background about your institution, your learners and assessors.
8. **Purpose:** Explain the rationale for your innovation in assessment – what problems or challenges were you attempting to address through your approach?
9. **Design:** What are some smart design principles you used to maximise the success of your assessment innovation? (Consider any international guidelines i.e. AMC standards)
10. **Implementation:** What challenges did you face in implementing the new approach? How did you maximise a smooth transition and take up of the new approach?
11. **Evaluation:** How did you evaluate the success of your assessment innovation? On reflection, what if anything would you do differently if you knew what you now know?
12. **Future focus:** What are your next steps and what are you working on now to further improve your system of assessment?

CONTEXT

The Royal Australasian College of Physicians (RACP) provides accredited specialist training to trainee doctors who have completed their medical degree and wish to further specialise as physicians in Australia or New Zealand. The RACP offers 61 training pathways in 33 diverse medical specialties as a part of Division, Chapter, or Faculty training programs. The RACP also offers six joint training programs with other Australasian colleges.

Trainees typically complete Basic Training, during which time they gain 36 months of experience in a range of internal medicine or paediatric disciplines. Following successful completion of Basic Training, trainees typically complete three or more years of Advanced Training in one of 33 medical specialties.

Advanced Training in Community Child Health is typically a three-year training program, which is completed by trainees under the supervision of a practising Community Child Health specialist at an accredited training site. Fellowship of the RACP (FRACP) is awarded upon completion of the program, and Community Child Health specialists generally work across the areas of developmental-behavioural paediatrics, child protection, and child population health.

PURPOSE

The RACP conducted a pilot of Entrustable Professional Activities (EPAs) in the Community Child Health Advanced Training program from January to August 2015. The purpose of the pilot was to explore the usefulness of EPAs in terms of both curricula design and workplace application for a RACP training program.

The pilot aimed to trial four key changes:

1. Focus learning and assessment on the core activities trainees need to be able to do independently by the completion of Advanced Training in Community Child Health.
2. Streamline the work-based learning and assessment process by introducing a single Learning Plan and Supervisor's Report form, in place of the existing Learning Needs Analysis and Supervisor's Report (separate forms).
3. Improve the collection of evidence about trainee progression by focusing existing work-based assessments on EPA tasks.
4. Encourage longitudinal assessment of trainees by requiring the setting of goals at the start of the training rotation, a mid-point review of goals and progress, and a final review of progress against learning goals.

DESIGN

The RACP used two main approaches to maximise the success of the pilot: an **organising framework** to structure the EPAs; and **training and support** for pilot participants.

An organising framework is an important part of curricula design, used to divide the content into logical units, often called 'themes' or 'domains'. The organising framework for Community Child Health EPA development was used as a basis for determining the important tasks that trainees need to be guided and assessed on across the breadth of the specialty:

1. Developmental-behavioural paediatrics
2. Child protection
3. Child population health
4. Community Child Health professional skills

The framework was developed by practising Community Child Health specialists to ensure context appropriateness.

23 teams, comprising 23 trainees and 33 supervisors, enrolled to be part of the pilot. The pilot was conducted across 19 different training sites in Australia.

Prior to the start of the pilot, participants were invited to attend an induction workshop. During this workshop, participants were introduced to the concept of EPAs and discussed the objectives, timelines, processes, and support available for the pilot. Questions were encouraged during the workshop to develop participants' understanding of the pilot concepts and process. Throughout the duration of the pilot, a dedicated RACP staff member provided guidance and regular updates to participants.

IMPLEMENTATION

Two challenges to implementation were highlighted during evaluation of the pilot: an **increase in the time commitment** required by supervisors; and **difficulties** associated with **using paper-based forms**.

In post-pilot evaluations, supervisors reported that they were required to sacrifice clinical and administrative work in order to spend more time observing trainee performance. Trainees reported that this was a positive aspect of the pilot and they appreciated having dedicated time to spend with their supervisor discussing their performance and being observed for EPA-related tasks.

This feedback suggests that successful implementation would require a cultural shift towards an increased focus on teaching and learning. An increase in the number of supervisors may also

alleviate the pressure placed on individual supervisors to observe and provide feedback to trainees, and could encourage a cultural shift towards increased valuing of trainee education, versus service provision.

A key aspect of the pilot was the use of a combined learning plan and progress report form, formatted in Microsoft Word. Pilot participants had difficulty using the form, reporting that there was insufficient room to enter learning goals, that the length of the form made navigation difficult, and that information required was repetitive, for example pre-, mid- and post-rotation EPA levels.

To support implementation, an online form that allowed more space for free text, easier navigation, and pre-population of repetitive information, such as EPA levels, would be very useful for trainees and supervisors.

EVALUATION

24 key exploratory research questions were established that directly related to the objectives of the pilot. These questions explored the utility, feasibility and content of the Community Child Health EPAs and were sorted according to Kirkpatrick's Hierarchy of Evaluation[2] .

The pilot was evaluated at the start, mid-point and at the conclusion of the pilot. Evaluations were conducted using four methods:

- Semi-structured phone interviews
- Post-pilot online surveys
- Post-pilot focus groups
- Content analysis of the Learning Plan and Supervisor Report forms

Participants were invited to take part in the first three evaluations through pilot bulletins emailed to them each month. Prior to the commencement of these evaluations, participants were also sent regular personalised emails, encouraging them to participate.

Results

- **Desirability.** 100% of the trainees and supervisors surveyed after the pilot agreed that the changes proposed in the pilot are desirable to implement in Advanced Training in Community Child Health and will increase the quality of the program.
- **Feasibility.** 87% of trainees and 85% of supervisors surveyed after the pilot agreed that it would be feasible to implement the proposed pilot changes in the Community Child Health Advanced Training program.

The two key areas which impacted participants' ability to adapt the pilot's changes to their daily practice were:

1. Time spent on program activities
2. Suitability of the workplace context.

2 Hutchinson L. Evaluating and researching the effectiveness of educational interventions. BMJ: British Medical Journal. 1999;318(7193):1267.

- **Impact.** The majority of participants agreed that the pilot impacted most on trainees' performance as learners. The EPAs were highly favoured by pilot participants because of their practical nature, offering greater clarity on what a trainee should be able to do, and supervisors commended the EPA structure for improving their ability to make objective, evidence-based decisions on their trainee's competence in the workplace. Participants largely agreed that the EPA system would help to identify trainees in difficulty

A number of recommendations were drawn from the results of this study:

1. Implement EPAs in College training programs.
2. Develop online learning environments and other key resources.
3. Develop a process for tracking trainee progression.
4. Educate and inform trainees and supervisors on their roles and responsibilities.
5. Consult on implementation in the workplace.

The full report of the Community Child Health EPA pilot is available on the RACP website.

FUTURE FOCUS

The RACP is currently engaged in a program of education renewal, designed to enhance the quality education services that the RACP delivers, to transform the business processes that support these services, and ultimately to better serve the health of patients, carers, communities, and populations.

The program comprises a series of interrelated projects that will be delivered over several years from 2018, and will transform the trainee journey, educator experience, and physician workforce of the future (see Figure 1).

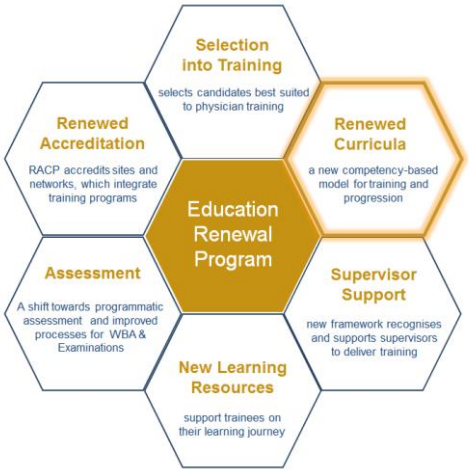


Figure 1. Education renewal graphic

A key component of the education renewal program is the renewal of curricula. The RACP is moving towards a new hybrid time- and competency-based model for training and progression, by redesigning, developing, and implementing 40 curricula across the College's Basic and Advanced Training programs. Revised curricula will be structured according to the RACP curriculum model (see Figure 2), which is structured around the concept of what trainees need to be, do, and know. Curricula standards will be supported by learning and teaching programs and assessment programs.

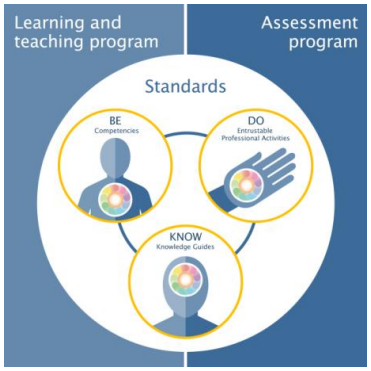


Figure 2. RACP Curriculum Model

Further information about RACP curricula renewal is available on the RACP website.

Notes

CONCLUSION

This series of case studies shares evolving practice of a range of leading medical educators who have adopted various approaches to innovation in assessment by drawing on programmatic assessment concepts.

The AMC intends to further develop its suite of case studies on good practice in assessment. If you would like to show case innovations in assessment or other aspects on medical education related to AMC standards please email the AMC at accreditation@amc.org.au