

Assessment in a global context: An international perspective on dental education

Patel, Upen

DOI:
[10.1111/eje.12343](https://doi.org/10.1111/eje.12343)

License:
None: All rights reserved

Document Version
Peer reviewed version

Citation for published version (Harvard):
Patel, U 2018, 'Assessment in a global context: An international perspective on dental education', *European Journal of Dental Education*. <https://doi.org/10.1111/eje.12343>

[Link to publication on Research at Birmingham portal](#)

Publisher Rights Statement:

This is the peer reviewed version of the following article:
Patel US, Tonni I, Gadbury-Amyot C, Van der Vleuten CPM, Escudier M. Shaping the future of dental education: Assessment in a global context: An international perspective on dental education. *Eur J Dent Educ*. 2018;22(Suppl. 1):21-27.
which has been published in final form at <https://doi.org/10.1111/eje.12343>.
This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

Shaping the Future of Dental Education

Assessment in a Global Context: An International Perspective on Dental Education

Authors

Upen S Patel, BDS, Ph.D., Clinical Lecturer in Restorative Dentistry and Assessment Lead, School of Dentistry, University of Birmingham, UK.

Ingrid Tonni, Ph.D., MOrth., Researcher in Orthodontics, University of Brescia, Italy

Cynthia Gadbury-Amyot, Ph.D., Associate Dean and Professor of Instructional Technology and Faculty Development, the University of Missouri-Kansas City (UMKC), Kansas City, Missouri, USA.

Cees PM Van der Vleuten, Ph.D., Professor of Education at Maastricht University, Department of Educational Development and Research in the Faculty of Health, Medicine and Life Sciences. The Netherlands.

Michael Escudier FDS RCS, FFGDP, Clinical Senior Lecturer in Oral Medicine, Department of Clinical and Diagnostic Sciences, and Deputy Director of Education (Student Assessment), King's College London Dental Institute, UK.

Abstract

Assessments are widely used in dental education to record the academic progress of students and ultimately determine if they are ready to begin independent dental practice. Whilst some would consider this a 'rite-of-passage' *of* learning, the concept of assessments in education are being challenged to allow the evolution of 'assessment *for* learning'. This serves as an economical use of learning resources whilst allowing our learners to prove their knowledge and skills and demonstrating competence. The Association for Dental Education in Europe and the American Dental Education Association held a joint international meeting in London in May 2017 allowing experts in dental education to come together for the purposes of Shaping the Future of Dental

Education. Assessment in a Global Context was one topic in which international leaders could discuss different methods of assessment, identifying the positives, the pitfalls and critiquing the method of implementation to determine the optimum assessment for a learner studying to be a healthcare professional. A post workshop survey identified that educators were thinking differently about assessment, instead of working as individuals providing isolated assessments, the general consensus was that a longitudinally orientated systematic and programmatic approach to assessment provide greater reliability and improved the ability to demonstrate learning.

Keywords

Assessment, Dental, Dentistry, Education, ePortfolio, Feedback, Global, Programmatic, Quality Assurance, Reproducibility, Validity, Workplace Based Assessment (WBA)

Introduction

Assessments are widely used in dental education to record the academic progress of students and ultimately determine if they are ready to begin independent dental practice. Utilising assessments solely for this purpose can be considered a traditionalist and ‘old-fashioned’ approach to education with the potential to be a wasted resource for learning. Identifying the student as the learner, the delivery of education has become more student-centred, and so must their assessments. Moving away from the paradigm of ‘assessment *drives* learning’ and towards a constructivist learning theory that utilises ‘assessment *for* learning’, we can improve the quality of education delivered to our students. Assessments must be suitably designed to allow for the effective evaluation of each student against pre-determined learning outcomes that require the testing of various different skills. Unfortunately, determining the competency of students cannot be carried out by a single assessment alone. The task of planning and delivering appropriate assessments for dental curricula, that are both reliable and valid, is challenging. Attempting to deploy assessments in isolation can lead to over assessment and the inadequate assessment of competencies that span across multiple subject areas in dentistry. Global assessment refers to the use and interpretation of multiple assessments over multiple time-points, and is the ethos at the heart of the programmatic approach to assessment. It is accepted that no assessment is perfect and striving for perfection can limit both the reliability and learning value of this activity. This paper reports on the discussions and findings of the ‘Assessment in a Global Context’ workshop of the ‘Shaping the Future of Dental Education’ meeting in London in May 2017. Three main subject areas were presented and discussed; (1) a programmatic approach to assessment, (2) ePortfolios and workplace based assessments as a global measure for capturing dental student competency and (3) quality assurance of assessment. The aim of this paper is to identify good practice and recommendations for the implementation of assessment strategies in dental curricula.

Pre-workshop online registrant survey results & reading list

Sixty-four delegates registered to attend the ‘Assessment in a Global Context’ workshop representing 20 countries and twenty-eight of those completed the online registrant survey prior to attending the meeting in London. The global distribution of delegates completing the pre-workshop survey was; seven from Asia, fifteen from Europe and six from North & South America. The majority of the delegates’ that completed the pre-workshop survey identified themselves as expert (25%) or proficient (36%) in their level of knowledge/experience with assessment. Twenty eight percent reported they were learning, and eleven percent enquiring. The Single Best Answer (SBA) (82%) and Objective Structure Clinical Examination (OSCE) (75%) were identified as being used by the majority of delegates when assessing students. Half of delegates reported that less than twenty-five percent of all assessments in their schools were standard set, however, where standard setting was utilised, the most commonly used method was Angoff (43%). Direct Observation of Procedural Skills (DOPS) (64%) was recognised to be the most used method of Workplace Based Assessment (WBA) amongst delegates and Assessment of Audit (AoA) (7%) was found to be the least used. Professionalism was reported as being assessed in the majority of schools from which delegates attended (71%) indicating the use of WBA’s as the most popular method of assessing this skill. Leadership and Management skills were less likely to be assessed, with only 29% of delegates indicating that these skills were assessed in their school. The majority of those delegates that did assess these skills indicated the use of Multi-Source Feedback (MSF). The pre-workshop data indicated that further guidance would be beneficial for the participants of the assessment workshop. It was also instructive to know that a majority of the pre-workshop survey participants identified themselves as expert/proficient, so bringing that expertise to the discussion would be critical. To aid discussion during the workshop, a reading list of relevant references were distributed to all registered participants. The references were categorised under the following subject headings; Selection of appropriate assessment tools (1-4). Selection of appropriate standard setting methods

(5-9). Workplace-based assessment (10-14). How best to assess “soft-skills” such as professionalism, leadership and management (4, 15-18).

Current issues with assessments

At the start of the workshop, all delegates in the room were asked to identify problems that they have been faced with when planning and implementing assessments as dental educators. The group identified the following:

- Inter & Intra examiner reliability
- Validity of only one method of assessment
- Inequality with teaching/teachers - assessor variability
- How to combine assessment to promote learning - assessment driving learning and still capturing the ‘bad apples’
- Can an assessment be objective?
- Perspective on feedback
- Lack of time to give feedback
- Learning ourselves?
- How to assess new assessment tools?
- Level of assessment literacy for faculty
- Long rooted traditional assessment beliefs - how to change this - is it about teaching subject vs enhancing learning - getting faculty to change - behavioural change

Important issues were raised, particularly around examiner reliability and the need to change the current methods of assessment. The feedback from the delegates indicated that there was an awareness that the traditional model of assessment was no longer adequate and there appeared to be a desire to move towards an ‘assessment *for* learning’ strategy.

Assessment tools

The type of assessment to be used must be considered carefully and should be aligned with the specific learning outcome being tested. The characteristics of an assessment tool are validity, reliability and educational impact. The four platforms of Miller's Pyramid (1) indicate how assessment should be planned:

- *Knows*; fact-orientated assessment e.g. multiple choice question (MCQ), essay, oral interview
- *Knows how*; Scenario or case-based assessment e.g. MCQ, essay, oral interview
- *Shows how*; Performance assessment (simulation) e.g. OSCE
- *Does*; Performance assessment (live) e.g. clinical performance assessment, MSF, WBA, ePortfolios

Validity refers to the specifics of what is being assessed. Traditionally, curriculums have been input orientated, where the teacher decided what was to be assessed. However, now with the advent of an era with competency frameworks, where the focus is on what the graduate should be able to do on completion of the course, has resulted in an outcome based curriculum and shifted us to providing an outcome based education. Assessments must target these competencies and demonstrate that a specific skill is being appropriately tested. Considering the different competency frameworks from around the world, every country has its own outcomes, set either partially or fully by their national governing or accreditation body; however there are consistencies when comparing general outcomes and domains e.g. clinical, professionalism, management, leadership. Clinical outcomes tend to be straight-forward when planning and implementing assessment however the more challenging ones, such as professionalism, management and leadership, pose a problem when trying to apply the perfect assessment. These are complex skills which are difficult to assess and cannot be delivered with a comparison to a simple checklist. These skills are behaviour led and learnt by experience over time. As a behavioural element, the preferred assessment should come from the

highest platform of Miller's Pyramid and be assessed by observational assessment tools. These tools are newly emerging and being tentatively adopted as they are unstandardised assessments compared to the established 'considered-safe' standardised assessments that are used for the lower three platforms of Miller's Pyramid. This is the biggest challenge for assessment in a health care professional curriculum. Research in assessment validity tells us that there is no magic bullet and that a mixture of methods is needed to cover the competency pyramid. Standardised assessments require quality control around test development and administration of the assessment tool to be vital. Unstandardised assessments consider the educators to be the vital component. The tool is not so important anymore. The educator needs to be able to accurately record, reflect, feedback to the learner and feedforward for progression and high-stakes decisions.

The reliability of an assessment can be determined by a number of factors. The best form of assessment is one where the person being assessed doesn't know that they are being assessed. Traditionally, objective tests have been considered more reliable than subjective tests, however, it has been shown that the length of assessment and number of assessments influences reliability. Research has shown that the longer the duration of the assessment, the better its reliability (19). Mini case-based examination (CEX) was found to be the most reliable assessment at each time point (1hr, 0.73; 2h, 0.84; 4h, 0.92; 8h 0.96) (20) however after 8 hours of testing, all assessments analysed in the study had a similar level of reliability (19). Moonen et al. (21) considered reliability as a function of sample size and identified that different assessments needed to reach a threshold number of times the student was assessed (sample size) for it to be considered reliable. The study also showed that the number of samples needed to achieve reliability reduced when assessments were considered as part of a composite of assessments. Ultimately, acceptable reliability is only achieved with large samples of multiple assessment elements. No assessment method is inherently better than any other as they must be used in combination to provide a global

assessment. Whilst subjective assessments are traditionally considered to have poor reliability, many subjective judgements, from different assessors, allow us to assess the complex competencies.

The relationship between learning and assessment is complex (22, 23). The traditional education paradigm of ‘assessment *drives* learning’ leads to poor learning styles as a result of grade hunting and competitiveness. This can result in grade inflation and students only learning so as to pass the assessment rather than the retention of knowledge or improving understanding. Anecdotally, it is claimed that students lose 50% of what they have learned for an assessment after one week. Whilst this is a sweeping statement, various studies have shown that knowledge is lost by various magnitudes depending on the type of content assessed (24-28). Narrative feedback from assessments is considered to be more beneficial, proving a greater impact on complex skills, than a simple grade or score (29-31). All assessments should provide students with meaningful feedback however a single episode of feedback is not sufficient and a dialogue should be made available. This may require the use of mentors who are able to provide global and longitudinal feedback throughout the training programme. Students will learn from assessments that are aligned with curriculum goals. If the learner understands why they must demonstrate particular knowledge, it is more likely to be learnt and remembered, than simply learnt for assessment. Longitudinal assessment is a key requirement for promoting ‘assessment *for* learning’ as the learner has the opportunity to demonstrate improvement over time rather than a single ‘snap-shot’ assessment at the end of the course. Longitudinal assessment also helps educators to identify learners who are not performing as expected and provides time to implement alternate learning strategies for these select few before high-stakes decisions have to be made.

Programmatic assessment

Based on the discussions thus far, the consideration for an ideal assessment in a health care professional education programme depends on its characteristics. Validity of the assessment would require the implementation of a multitude of assessment methods. Reliability is assured with the gathering and combining of information from multiple assessments. To maximise the impact of learning from assessments requires the provision of a longitudinal and meaningful feedback dialogue. Together these three elements (validity, reliability and educational impact) form the backbone of the programmatic approach to assessment (2). Dijkstra et al. (32) published 73 generic guidelines to support the design of a programmatic approach to assessment and these have been adopted by ASPIRE, a programme supported by the Association of Medical Education in Europe (AMEE) and recognises international excellence in medical, dental and veterinary schools.

A key element of the programmatic approach to assessment is the understanding that each and every assessment is a data point. Each of which must be optimised for learning. Optimisation involves gathering not only quantitative data, but also qualitative meaningful feedback. Qualitative feedback, as a data point, differs from the other conventional data points measured at each stage of the training programme and thus provides variation to the assessment process. Students in summative only assessment systems tend to ignore feedback. A programmatic approach utilises both summative and formative assessments, replacing them with a 'continuum of stakes' that are considered together, to impact on decisions of progression. The number of data points is proportionally related to the stakes of the decision to be taken. The higher the stakes, the more points are considered. Whilst this may appear as over assessment, it is not thought of as such by the student body, as each assessment is meaningful and part of their learning programme. Educators must take the approach of trusting students, rather than controlling them, by creating a constructivist learning environment where students are challenged to do their own learning rather

than being forced to learn. Meaningful aggregation of the data for decision making relies on the results of multiple assessments of different competencies using different methods (Figure 1). This ensures both the reliability and validity of the outcomes of the assessment for that competency.

















		COMPETENCY			
		A	B	C	D
METHOD OF ASSESSMENT	1				
	2				
	3				
	4				
		AGGREGATION FOR DECISION MAKING			

Figure 1 Aggregation for decision making. Table shows how different methods of assessments are used to assess the same competency with aggregation prior to decision making as implemented in a programmatic approach to assessment.

Workplace based assessment

Professional competence, i.e. the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, judgment, emotions, values and reflection in day-to-day practice (33), is best assessed with a WBA approach. It is necessary to assess clinical practice in the unstandardized real world and is best accomplished using observation and judgement. The assessment should ideally include different types of WBA's, across different contexts, subject matters and with multiple assessors that all feed in to an ePortfolio. Obtaining professional high

quality feedback is important when using WBA's. There are three methods of assessment used in the WBA approach; DOPS, MSF, ePortfolio. ePortfolios and their use in programmatic assessment will be discussed in this article.

ePortfolio to assist in programmatic assessment

ePortfolio assessment offers a competency-based assessment strategy that provides integration across competencies and disciplines, sampling from multiple contexts and multiple sources, triangulation of information, and the training of dental students to be reflective practitioners (15). The implementation of an ePortfolio into a curricula that engages with programmatic assessment, challenges students to employ critical thinking and problem solving skills as they construct meaning out of their educational experience. It is widely accepted that students learn differently and this can be based on the different experiences throughout their life. However, when considered as an institution, all students are within the same academic setting, being exposed to the same information. The versatility of ePortfolios allows students to reflect and demonstrate learning in their own individual method and a pace that is appropriate to them. When students start a course of learning they are considered to be at the level of a novice and during the program they should progress to become competent (34). Competence is not an achievement but rather a habit of lifelong learning; assessment plays an integral role in helping students identify and respond to their own learning needs. Ideally, the assessment of competence (what a student is able to do) should provide insight into actual performance (what he or she does habitually when not observed), as well as the capacity to adapt to change, find and generate new knowledge, and improve overall performance (35). Learning is not a spectator sport, learning is active and students should engage to learn (self-directed learning). Research shows that it is in active learning that students move information from short-term or working memory to long-term memory (36). The shift from traditional testing of knowledge towards 'assessment *for* learning' depends on the context; an ePortfolio has the capacity

to adapt to different situations and is context specific. Eraut (37) said that professional competence is more than a demonstration of isolated competencies. When the whole is considered, its parts are seen differently than when seeing them in isolation. An ePortfolio can show the competencies as a whole and provides an opportunity for different members of the faculty to communicate with each other. There can be resistance from faculty, some of them did not learn through Competency-Based Education (CBE) and are unaware of how an ePortfolio can be used for assessment. A significant amount of faculty development is needed. There is resistance from students and those involved in dental educational programmes, with the latter questioning the validity and reliability of an ePortfolio. The literature clearly shows that measurements over time by multiple evaluators and multiple sources provide good validity. The ePortfolio allows for this information to be collated and analysed.

Quality assurance of assessment

Losing control and focus when assessing is sometimes required to assess complex skills. However an element of control must still remain to assist in quality assurance (QA) of the assessment. The quality or 'fitness for purpose' of an assessment must evaluate the assessment characteristics; educational impact, reliability and validity. The different reasons for assessment (i.e. it drives learning, it informs stakeholders, it provides feedback for education) determine the different criteria to use for QA. The purpose of the assessment is also important in QA (38). There is a big debate in the literature between summative and formative functions and whether they exist separately. If there is an assessment *of* learning, the summative aspect is predominant and there should be feedback both to students and teachers. If the focus is more towards assessment *for* learning, then the formative assessment is much more important and the quality of teachers' feedback should be organised. Constructive alignment represents the most important aspect of quality assurance. The objectives of the curriculum should be aligned with the tasks that are organised for students to reach

these objectives (39). The assessment provides students the opportunity to show if they can achieve the objectives.

When considering the aims of QA, there is a choice between ‘control learning’ and ‘improve learning’. It is important to clarify the aim because it determines the type of information that needs to be collected. If the aim is to monitor the program, then detective witnesses are needed to provide flag signals to direct attention if something happens. If the aim is to improve the system, much richer data are needed to know what is happening, why is it happening and how things can be improved. QA is a cyclic process and Figure 2 shows a model that can be applied to assessment. Measurements are initially set and judged and thereafter the information is collected and interpreted for improvement. The cycle is a constant process and within this method, there is nothing as consistent as change.

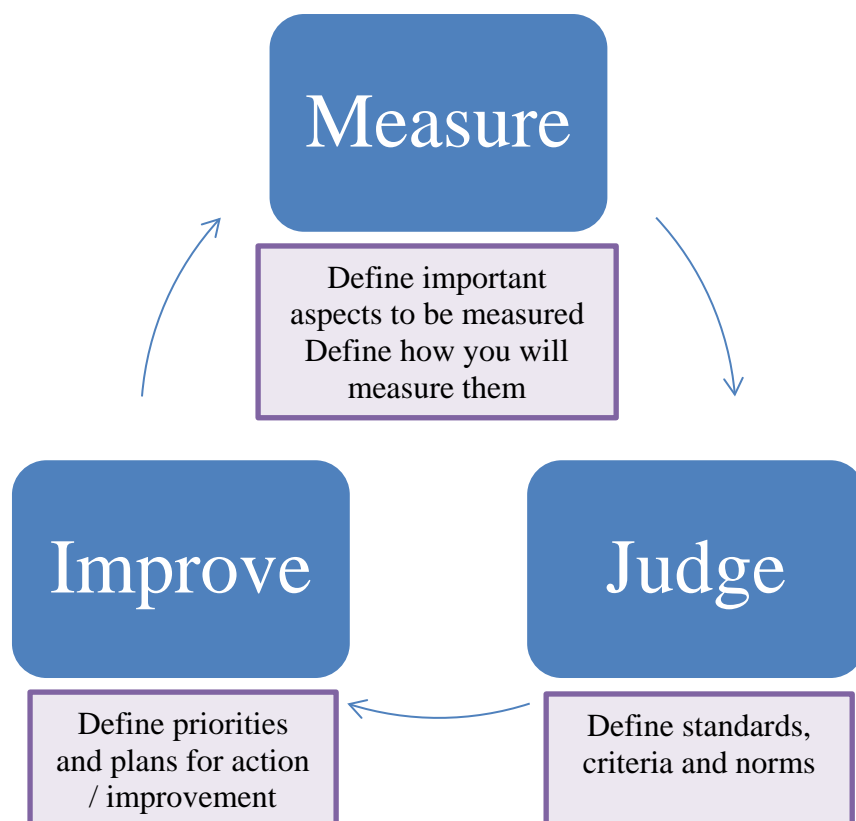


Figure 2 QA is a cyclic process. Diagram shows the different steps of the QA process for assessment. This model is used at Maastricht University.

In the first step of the cycle, important aspects to be measured and how to measure them are defined. In the second step, criteria and norms should be defined. For instance gathering feedback from students is an important data source to see how they are doing. If their satisfaction level is below three in a five points scale a flag is registered. This is not a sign of a real problem with the assessment system; it is just a signal to check why students are so unsatisfied. In the third step, priorities are defined because the results of QA depend on cost and time. Choices should always be made in order to stay aligned with the stakeholders and prioritize their problems, because otherwise people get frustrated. A method may be not to control all the assessment tasks every year, for example, if there is a new course it could be monitored every three years. It is important to decide where energy should be spent to go in depth in some of the assessment tasks and really improve these aspects instead of doing everything superficially.

Post-workshop participant audience response poll results

At the close of the workshop, all participants were asked to electronically respond to a series of questions. At the time, the information was collated and reported back to the participants immediately by way of a word cloud. Further interpretation of the information has now taken place and the information is presented here. A total of 257 responses were received from the four questions asked, of these, 55% (141/257) represented unique participant responses. For the interpretation of the polling questions below, the frequency is shown in parenthesis and only those words with the highest frequency are reported.

Question 1: "Name one or two things that are "take-home" messages from this two-day workshop on assessment."

Excluding the word “assessment”, participants responded with the following words; feedback (10), reliability (8), quality (6), subjective (6), longitudinal (5), validity (5). Some participants responded with the following statements; “Assessment *for* learning far more important than for grading”, “Many observation points and multiple evaluators”, “Narrative feedback from clinical experts is essential”. “Subjective is okay! Longitudinal assessment is better”, “Subjective assessment based on professional experience is necessary”.

Question 2: “What do you perceive are the greatest opportunities for dental education when it comes to assessment?”

Excluding the word “assessment”, participants responded with the following words; learning (12), students (6), portfolio (4). Some participants responded with the following statements; “We can learn a lot from other professions and we should do that more”, “Being better able to assess if patient-centred ‘soft’ skills are being learned”, “Ensure student learning. Improve faculty capacity to teach”.

Question 3: “What do you perceive as the greatest challenges for dental education when it comes to assessment?”

Excluding the word “assessment”, participants responded with the following words; faculty (13), change (7), culture (6), time (5), engagement (4). Some participants responded with the following statements; “Wider recognition of the importance of assessment”, “Longitudinal multiple assessment”, “Teaching the teachers”, “Time for faculty development”, “Low assessment literacy”.

Question 4: “A position paper will be developed based on what we learned over the past couple of days around the topic of assessment. What do you believe is essential content for that position paper?”

Excluding the word “assessment”, participants responded with the following words; feedback (7), importance (6), quality (5), learning (4). Some participants responded with the following statements; “The value of longitudinal assessments”, “Learning is complex so is assessment”, “To also emphasize assessment for instruction; assessment for accountability (quality assurance)”, “Assessment as an opportunity to improve”, “making assessment a learning activity”.

Conclusions

There has been an obsession in education for obtaining reliability and objectivity with assessments; the OSCE serves as one example. However we should now consider ourselves in a post-reliability world where validity is more important when assessing complex skills. Professional judgement should be used and the limitations of objective assessment should be known. Reliability and validity are inversely related (classical or G-theory) and a balance between these elements can be achieved by combining different assessment tools. Subjectivity is dealt with through sampling and procedural bias reduction methods (not with standardization or objectification). Implementation may require a change in management as these ideas are paradigm shifting and a cultural change. Investment is needed to improve learning, both monetary and for faculty buy-in. WBA is potentially the best way of assessing professional competence and is considered best practice for a clinical based education programs. Development of programmatic student portfolios challenge students to employ critical thinking and problem solving skills as they construct meaning out of their educational experience and convey that meaning to others. Quality assurance must take into account the complexity of the competency being assessed and students should be involved in this process to support the

development of appropriate assessments. We have to stop thinking in terms of individual assessment methods, and instead think in terms of a longitudinally orientated systematic and programmatic approach to assessment. Assessments do not need to be resource intensive when they are considered part of learning (assessment *for* learning). The programmatic approach to assessment optimizes both the learning function (through information richness) and the pass/fail decision function (through the combination of rich information).

References

1. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65(suppl 9):63-7.
2. van der Vleuten CPM, Schuwirth LWT, Driessen EW, Dijkstra J, Tigelaar D, Baartman LKJ, et al. A model for programmatic assessment fit for purpose. *Med Teach*. 2012;34(3):205-4.
3. van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. *Adv Health Sci Educ Theory Pract* 1996;1(1):41-67.
4. van der Vleuten CPM, Schuwirth LW, Scheele F, Driessen EW, Hodges BD. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol*. 2010;24(6):703-19.
5. Barman A. Standard setting in student assessment: is a defensible method yet to come? *Ann Acad Med Singapore*. 2008;37(11):957-63.
6. Ben-David MF. AMEE Guide No. 18: Standard setting in student assessment. *Med Teach*. 2000;22(2):120-30.
7. Cohen-Schotanus J, van der Vleuten CPM. A standard setting method with the best performing students as point of reference: practical and affordable. *Med Teach*. 2010;32(2):154-60.
8. Norcini JJ. Setting standards on educational tests. *Med Educ*. 2003;37(5):464-9.
9. Purver J, O'Sullivan D. An introduction to standard setting methods in dentistry. *Br Dent J*. 2015;219(7):355-8.
10. Al-Kadri HM, van der Vleuten CPM. Workplace-based assessment and students' approaches to learning: a qualitative inquiry. *Med Teach*. 2013;35(suppl 1):31-8.
11. Brown JM, Lowe K, Fillingham J, Murphy PN, Bamforth M, Shaw NJ. An investigation into the use of multi-source feedback (MSF) as a work-based assessment tool. *Med Teach*. 2014;36(11):997-1004.
12. Crossley J, Jolly B. Making sense of work-based assessment: ask the right questions, in the right way, about the right things, of the right people. *Med Educ*. 2012;46(1):28-37.
13. Govaerts M, van der Vleuten CPM. Validity in work - based assessment: expanding our horizons. *Med Educ*. 2013;47(12):1164-74.
14. Norcini JJ. Work based assessment. *BMJ*. 2003;326:753-55.
15. Gadbury-Amyot CC, McCracken MS, Woldt JL, Brennan R. Implementation of Portfolio Assessment of Student Competence in Two Dental School Populations. *J Dent Educ*. 2012;76(12):1559-71.
16. Gonzalez MA, Abu Kasim NH, Naimie Z. Soft skills and dental education. *Eur J Dent Educ*. 2013;17:2.
17. Howe A, Barrett A, Leinster S. How medical students demonstrate their professionalism when reflecting on experience. *Med Educ*. 2009;43:942-51.
18. Kalenderian E, Skoulas A, Timoth   P, Friedland B. Integrating leadership into a practice management curriculum for dental students. *J Dent Educ*. 2010;74(5):464-71.
19. van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-17.
20. Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med*. 138(6):476-81.
21. Moonen-van Loon JM, Overeem K, Donkers HH, van der Vleuten CP, Driessen EW. Composite reliability of a workplace-based assessment toolbox for postgraduate medical education. *Adv Health Sci Educ Theory Pract*. 2013;18(5):1087-102.
22. Cilliers FJ, Schuwirth LW, van der Vleuten CPM. A model of the pre-assessment learning effects of assessment is operational in an undergraduate clinical context. *BMC Med Educ*. 2012;16(12):9.
23. Cilliers FJ, Schuwirth LW, Herman N, Adendorff HJ, van der Vleuten CPM. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract*. 2012;17(1):39-53.

24. Harrison A. Using knowledge decrement to compare medical students' long term retention of self-study and lecture materials. *Assess and Eval in Higher Educ.* 1995;20: 149-159.
25. Krebs R, Hofer R, Bloch R, Guibert JJ. Conversation et oubli des connaissances en biologie acquises pour le premier examen propédeutique de médecine. *MEDUCS Bulletin de l'Association Suisse d'Education Medicale.* 1994;4:10-5.
26. Sissons JC, Swartz RD, Wolf FM. Learning, retention and recall of clinical information. *Med Education.* 1992;26:454-61.
27. Swanson DB, Case SM, Luecht RM, Dillon GF. Retention of basic science information by fourth year medical students. *Acad Med.* 1996;71(suppl):80-2.
28. Watt ME. Retention of preclinical knowledge by clinical students. *Med Educ.* 1987;21:119-24.
29. Black P, Wiliam D. Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice.* 1998;5(1):7-68.
30. Butler R. Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology* 1987;79:474-82.
31. Butler R, Nisan M. Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance. *Journal of Educational Psychology.* 1986;78:210-6.
32. Dijkstra J, Galbraith R, Hodges BD, McAvoy PA, McCrorie P, Southgate LJ, et al. Expert validation of fit-for-purpose guidelines for designing programmes of assessment. *BMC Med Educ.* 2012;17:20.
33. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA.* 2002;287(2):226-35.
34. Hendricson WD, Cohen PA. Oral health care in the 21st century: implications for dental and medical education. *Acad Med.* 2001;76(12):181-206.
35. Epstein RM. *Assessment in Medical Education.* N Engl J Med. 2007;356:387-96.
36. Medina J. *Brain Rules.* Seattle, WA: Pear Press, 2014.
37. Eraut E. *Developing Professional Knowledge and Competence.* Washington D.C.: The Falmer press, 1994.
38. Bennett RE. Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice.* 2011;18(1):5-25.
39. Biggs J. Enhancing teaching through constructive alignment. *Higher Education.* 1996;32:347-64.