

2016-07-11

Progress testing in undergraduate dental education: the Peninsula experience and future opportunities

Ali, K

<http://hdl.handle.net/10026.1/5162>

10.1111/eje.12149

European Journal of Dental Education

Wiley

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Please cite only the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.

“This is a final author’s draft of a paper submitted for publication in European Journal of Dental Education 2016 DOI:
<http://dx.doi.org/10.1111/eje.12156>”.

Title:

**Progress Testing in Undergraduate Dental Education: The Peninsula
Experience and Future Opportunities**

Ali K; Coombes L; Kay EJ; Tredwin C; Jones G; Ricketts C; Bennett JH

Plymouth University Peninsula Schools of Medicine and Dentistry (PU PSMD),
Plymouth PL4 8AA, United Kingdom

Correspondence Author: Kamran Ali

Plymouth University Peninsula Schools of Medicine and Dentistry (PU PSMD),
Plymouth PL4 8AA, United Kingdom

C402 Portland Square

University of Plymouth

Drake Circus, Plymouth

PL4 8AA

UK

Tel: (+44) (0) 1752586768

E-mail: kamran.ali@plymouth.ac.uk

Abstract

Background: Progress testing is well established as a longitudinal form of assessment in undergraduate medical programmes to measure growth in knowledge. Peninsula Dental School is the first school to use progress testing and remains the only one to do so

Aims: To share the experience of developing progress testing in an undergraduate dental programme as a major summative assessment tool at a newly established dental school in the United Kingdom.

Methods: Data were collected for progress tests conducted from 2007-14. The tests were formative in the first two years of the programme and summative in subsequent years. Each test was based on 100 single best answer multiple choice items with an appropriate vignette. The students chose their answer from 5 options. A score 1 mark is awarded for each correct answer; minus 0.25 for an incorrect answer and 0 for 'Don't Know' (DK). The standard setting for each sitting was carried out using Angoff and Hofstee methods.

Results: There were two tests per year with each cohort undertaking 8 tests in their four years of study providing a total 14 test occasions. The reliability of each test for each student cohort tests was measured using Cronbach's alpha. The average reliability over 42 test/cohort combinations was 0.753 (\pm SD0.08). Data analyses shows growth in knowledge of dental students across successive years with the largest increase in knowledge observed between tests 1 and 5 and concomitant reduction in DK responses.

Conclusion: This is the first study to report the establishment and use of progress testing as the principle form of written summative testing in an undergraduate dental curriculum. Progress testing is a valid and reliable tool to assess growth in knowledge longitudinally over the duration of a dental programme. Although a labour intensive process, progress testing merits more widespread use in dental programmes.

Introduction

Progress testing is a form of longitudinal, feedback oriented assessment of the development and sustainability of cognitive knowledge at regular intervals over the course of an educational programme. It was pioneered by the University of Missouri-Kansas City School of Medicine and Maastricht University in the Netherlands to assess the knowledge of undergraduate medical students (1, 2) and later adopted by other schools, notably, McMaster University (3), and Peninsula College of Medicine and Dentistry (4). The key principle of progress testing is longitudinal assessment of growth of knowledge based on a sequence of equivalent, yet different, tests. The results of these are combined to determine the growth of knowledge for each student, enabling more reliable and valid decision making about promotion to a next study phase (5).

Progress testing offers several advantages compared to traditional yearly assessments. Given that assessment drives learning, testing at regular intervals over the course of an educational programme helps monitor the progress of students. Traditional methods involving assessment at the end of each module may promote rote learning and short-term memorisation of facts. Progress testing encourages the students to acquire information and develop understanding breaking the link between learning and revision (1). Setting the standard of questions at the level of a new graduate, progress testing facilitates assessment of functional knowledge and breaks the relationship between the taught programme and assessment (1, 2, 3). In addition, progress testing provides huge opportunities for feedback for internal and external evaluations. It allows students and their academic supervisors to identify areas of weakness to improve performance in successive years (3, 4). Nevertheless, progress testing also presents several challenges and potential disadvantages: Development

and maintenance of a question bank, administration of tests, psychometric analyses of data entail intense involvement of academic and administrative staff which is costly (1). From a students' perspective, results of progress testing may be demoralising for new entrants to the course. However, this issue can be addressed through appropriate guidance and support by the academic staff to manage students' expectations in the first year. In addition, due to limited knowledge in the first year, the scores may have a low predictive validity.

Progress testing has become increasingly popular over the years and is now firmly established in undergraduate medical curricula internationally (6-11). In addition, there is a growing trend of using progress testing in postgraduate medical education (12). However, it has not been used in dentistry, despite similarities in curricula, educational approaches and desired educational outcomes.

In this paper we report our experience of developing progress testing for use as a major summative assessment tool in an undergraduate dental programme over a seven-year period from 2007-2014.

Peninsula Dental School

Established in 2007, Peninsula is one of the first dental schools in UK to be founded in a primary care setting with a highly innovative, problem-based, patient-centred curriculum (13, 14). With an annual intake of 64 students, the school follows graduate-entry criteria to its four-year (five-year equivalent) Bachelor of Dental Surgery (BDS) programme. Teaching and learning is delivered in the context of the patient scenarios. Students are signposted to essential topics defined in the learning objectives for each scenario with the aid of plenary lectures and small-group, interactive sessions. These sessions provide the students with

opportunities to interact with subject specialists. Students learn their clinical skills in the simulated dental learning environment (SDLE), and the learning outcomes for these sessions are similarly linked to the cases. Students gain early clinical exposure and start seeing patients midway through Year 1. Time spent on clinic increases progressively from 1 day per week in Year 1 to 4 days in Year 4. The subjects introduced to Year 1 are revisited and developed sequentially in subsequent years. The breadth and depth of student learning at each 'visit' to a topic increases and helps students to build layers of knowledge, skills and attitudes in a spiralling manner.

Progress testing at Peninsula

Peninsula Dental School was the first to use progress testing in undergraduate dental education¹⁵. The dental progress tests at Peninsula require students from all 4 years to sit the same test simultaneously. Each progress test is benchmarked to the level expected of a newly qualified dentist. Tests are repeated twice a year for the duration of the course, and progress is indexed by a steady increase in scores achieved. The progress tests are formative in years 1 and 2 and summative in subsequent years. Each test is based on 100 single best answer multiple choice items. Where possible, questions are written around an appropriate dental vignette setting the test item within a particular clinical context with the objective of testing the analysis, synthesis and application of knowledge as distinct from simple factual recall. The students choose their answer from 5 options or can choose a 'Don't Know' (DK) option. A score 1 mark is awarded for each correct answer; minus 0.25 for an incorrect answer and 0 for 'Don't Know'.

Test Production

The entire question bank for progress testing has been developed in-house and involves a two-stage quality assurance process before final inclusion in the bank. All new questions are submitted to subject specialist panels headed by senior academics. Questions deemed to be appropriate are then submitted to a Dental Question Review Group (DQRG) which considers all details of each question before acceptance; questions remain subject to further scrutiny in the pre-test and post-test meetings. The entire question bank and individual progress tests are blueprinted against the learning outcomes of the General Dental Council (GDC), UK¹⁶. The learning outcomes for dental graduates are outlined by the GDC under four key domains: (1) Clinical; (2) Communication Skills; (3), Professionalism and (4) Management and Leadership skills. Although progress testing is often considered to involve administering the same single test repetitively, it does not always imply repeating the same questions too. In the context of progress testing at Peninsula, the format and standard of the test are aimed to remain constant i.e., 100 MCQs (single best type) set at the level of knowledge expected from a new dental graduate. While the question bank is blueprinted against the GDC learning outcomes, it is not possible to assess all the learning outcomes with 100 questions. The aim is to include questions mapping to the four major domains of the learning outcomes in each sitting. Moreover, we also aim to ensure that questions related to all learning outcomes which are amenable to assessment in a multiple-choice exam format, are included during eight sittings which each cohort sits for the duration of the BDS course.

Standard Setting

The standard of the questions is benchmarked against the knowledge expected from a newly qualified dentist. The standard expected of a new graduate is set by criterion-referencing

using a combination of Angoff and Hofstee methods (10, 17). This involves a panel of experienced faculty members including clinicians, subject specialists and other academic staff. Members establish difficulty level of each question in a progress test by indicating if a borderline (minimally acceptable) student is expected to answer the question correctly using a “yes” or “no” response (Angoff method). Also the members indicate the minimum and maximum pass marks along with minimum and maximum failure acceptable for a given test (Hofstee method). The ratings are averaged across panel members for each item and then summed to obtain a panel-recommended raw cut-off score which represents the score expected from a minimally competent candidate (newly qualified dental graduate). The standard setting is then deliberated further in a group setting during a “Moderation” meeting to finalise the cut-off score for Year 4. Progress tests are also summative for Year 3 but the cut-off score is lower than Year 4 to account for differences in difficulty and variance between the two cohorts. The Year 4 cut-off score can be converted to a z-score which expresses the number of standard deviations the pass mark is away from the Year 4 mean. The Year 3 standard is placed relative to the Year3 mean at the same number of standard deviations,

Feedback

An essential element of progress testing is to provide immediate and comprehensive feedback on performance to the students. In addition to the test scores each student receives their ranking within the cohort, progress in relation to previous sittings and details of correct, incorrect and “don’t know” responses. Moreover, with receipt of their scores, students receive a short statement outlining the main learning outcome being addressed by each test item.

Analysis of Progress Test Data

Analyses of data are undertaken using the R statistical language and environment¹⁸ (R core Team 2013). Analyses include simple descriptive statistics and distribution plots, reliability measures, demographic analysis to ensure that the test does not discriminate against particular groups and classical test theory based item analyses¹⁹. This information is used post-test to review the standards of the assessment and locate items for review and possible exclusion from the final test results. While statistical analyses guides this process, the decision to remove any item always rests with the content experts at the post-test committee.

Results

Up to the end of the academic year in 2013-14, there were two progress test sittings annually leading to a total 14 test occasions. However, a given cohort undertakes only 8 tests in total throughout their four years of study. Although each test is taken by all students, analysis of the performance for each cohort is carried out separately, so there have been a total of 42 test/cohort combinations over the last 7 years.

Progress of students across the 8 tests students in each cohort taken during the BDS programme is depicted in Figure 1. The largest increase in knowledge and reduction of “don’t know” responses was observed between tests 1 and 5. This indicates students acquired most of their knowledge prior to their final year of study, and highlights just how much knowledge students acquire each academic year. The difference between the Year 2 and Year 3 (tests 4 and 5) being particularly noticeable which coincides with the test becoming summative in Year 3 (Table 1). The progress of individual cohorts in successive progress tests taken during the course of the BDS programme is depicted in Figure 2.

To assess the reliability of each test we use Cronbach’s Alpha. The average reliability of the test was 0.753 (\pm SD0.08) with the values for Years 1 to 4 being 0.82 (range 0.74-0.90), 0.78 (range 0.71-0.90), 0.71 (range 0.52-0.76) and 0.66 (0.56-0.74) respectively. The higher reliability observed in Year 1 may indicate consistency in gaps in each student’s knowledge which serve to increase the internal consistency of the test.

| Test | Score | Std. Dev | Correct | Incorrect | Don't Know |
|------|-------|----------|---------|-----------|------------|
| 1 | 19.32 | 8.55 | 25.36 | 24.33 | 50.31 |
| 2 | 24.30 | 9.51 | 31.59 | 29.64 | 38.77 |
| 3 | 37.01 | 10.52 | 42.93 | 29.47 | 27.59 |
| 4 | 42.21 | 11.65 | 47.36 | 32.82 | 19.82 |
| 5 | 58.47 | 12.44 | 62.62 | 27.19 | 10.18 |
| 6 | 58.29 | 10.88 | 61.93 | 28.30 | 9.76 |
| 7 | 66.70 | 8.06 | 69.08 | 24.57 | 6.35 |
| 8 | 66.57 | 8.36 | 68.74 | 24.94 | 6.35 |

Table 1: Percentage of Correct, Incorrect and Don't Know Responses.

(Test is the test number in each student's sequence of 8; N is the number of cohorts used to calculate the Mean Test Score and subsequent columns.)

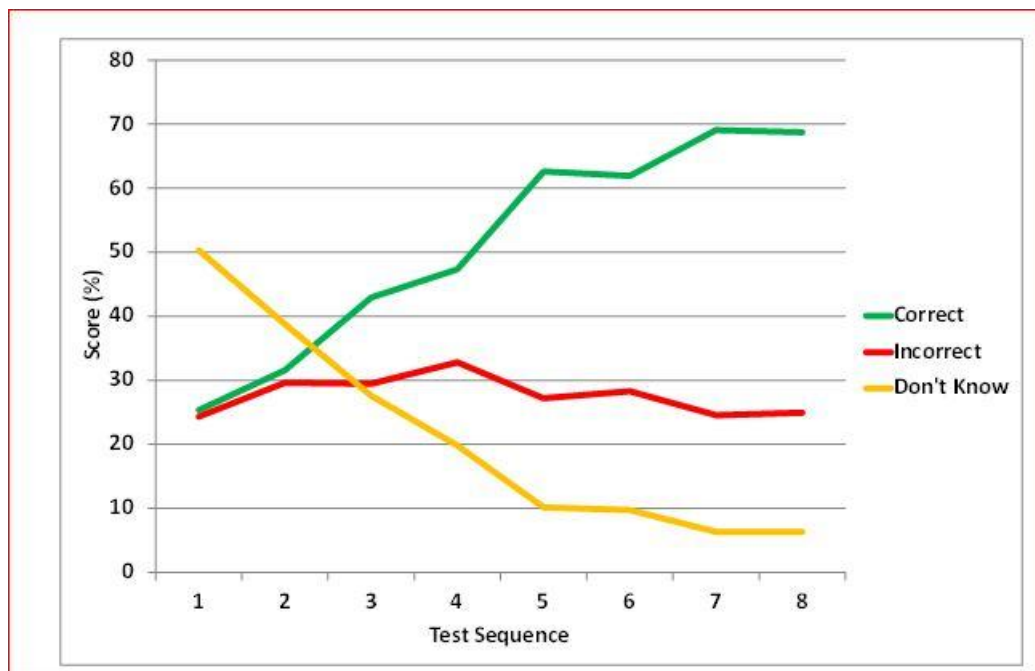


Figure 1 Graphical Representation of Growth in Applied Dental Knowledge

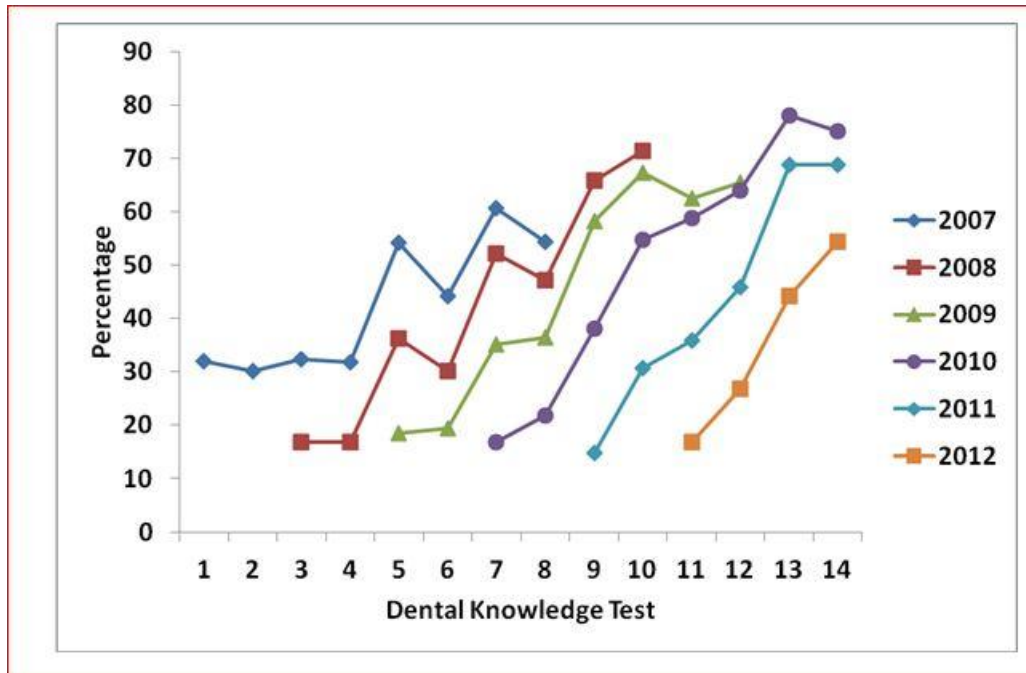


Figure 2 Graph shows progress of individual cohorts in successive progress tests taken during the course of the BDS programme.

Discussion

This is the first study to report the establishment and use of progress testing as the principal form of written summative testing in an undergraduate dental curriculum. Progress testing has been used over seven years, sufficient time for the test to be established and cohorts of students to process through from matriculation to graduation. Evidence from the literature supports progress testing as a valid and reliable form of assessment in medical programmes to measure the growth in knowledge with a positive influence on student learning (1, 2). Our results are broadly consistent with the medical experience. The key objectives of progress testing are to provide repeated and comparable integrated assessments across the range of topic areas covered throughout the duration of some or all of a programme (20). Comparisons of scores from successive tests are used to give an indication of knowledge growth. Frequent testing allows early recognition of learning difficulties and to identify the need for early intervention. In common with its use in our sister medical school (21), the application of progress testing to a dental programme has achieved this objective.

A desirable goal of progress testing is to focus assessment on the application of dental knowledge to clinical situations rather than encouraging rote learning. Achievement of this objective across the board may be restricted by the exclusive use of a single best answer out of five multiple choice format. However, test items based around a carefully constructed clinical vignette are more likely to achieve this than non-contextualised test items and it is the design of individual test items as distinct from the format of the test which determines success of individual questions (22).

In common with other authors (1), preparation of test items in sufficient numbers to populate repeated tests is time consuming and has placed a considerable load on the staff involved; indeed, because of the load placed on staff and resources, progress testing has been discontinued in one of the colleges pioneering this form of assessment (23). Because of the challenge this poses there has been a growing trend of international collaboration amongst medical institutions in the use of progress testing with exchange of questions or shared question banks (24-27). In Dentistry this has been a particular challenge because, unlike Medicine, there are no shared banks of dental questions from which test items can be drawn¹⁵. If progress testing is to become widespread, dental educators will need to accept the challenge of establishing of shared databases. Although progress testing is typically based on multiple choice questions (MCQs) the use of progress testing is not restricted to these or other similar formats requiring cued answers. Other question styles have been tried including an un-cued short answer format (28).

The results demonstrate growth of students' knowledge with progression through the programme and corroborate the growth of medical knowledge amongst medical students following contemporary problem-based curricula (20). Of particular note is the observation that knowledge growth is greatest during the early years of the programme, tailing off in the final year. Given that a spiral design of the curriculum is used, Year 3 may be a year of consolidation which revisits the biological and clinical knowledge of the previous years and this may also account for a marked increase in scores during Year 3. Early exposure to progress testing students to work towards the level of knowledge expected from them upon graduation. Although the tests are formative in the first two years of the programme, our test-data provide evidence for steady growth in knowledge. Additional studies are required

to explore the perceptions of students regarding progress testing in general and more specifically to evaluate the impact of formative and summative progress testing on growth in their knowledge. There are few data available addressing how optimal progress testing should be applied; whether it is of value across the whole programme or, if it is of greatest use in the early years only and warrants additional forms of assessment in the final year. Our results raise the possibility that this might be the case. Investigation of these issues will provide greater insight into the dynamics of student learning throughout the latter stages of the dental programme and contribute to our understanding of how progress testing may be optimally applied.

At Peninsula Dental School, we use negative marking for incorrect responses with a penalty of minus-0.25 mark per incorrect response. This approach is in line with our sister medical school (21). Although negative marking is not used consistently in progress testing, the rationale is to discourage guess work by students which may potentially advantage or disadvantage students and lead to falsification of test score. This may be appropriate and perhaps more reliable for specialised assessments like progress testing while a “number right” (scores based on right answers only) may be preferred for normal achievement tests (29). Students in early years have limited knowledge and may prefer to use the “don’t know option” rather than resort to guess work. Inclusion of a “don’t know” option is also appropriate to reinforce the need to recognise personal limitations in knowledge in clinical practice (29). However, further research is required to evaluate the practice of negative marking in progress testing.

Progress testing provides a huge opportunity for feedback (21, 30). All students at Peninsula receive detailed, longitudinal feedback after each test and are required to discuss their performance with their academic tutors at regular portfolio appraisals (RPAs). The feedback allows identification of gaps in knowledge and remediation for under-performing students. Even in the early years when progress tests are formative and the results do not contribute to decisions on advancement of students in the programme, they are expected to sit the tests. Failure to engage with the progress testing process is regarded as a serious professionalism issue to be considered at their regular portfolio appraisals.

Development and review of progress testing by subject experts and mapping of questions to GDC learning outcomes provide face and content validity. Moreover, review by external examiners contributes to external validity of the tests. Although other studies support the predictive validity of progress testing (5), we did not perform separate analysis to measure predictive validity of progress test data and perhaps this could be addressed in a separate study. At 64 students per year, dental student numbers are smaller than comparable medical courses. This may contribute to greater variability in measures of reliability such as the Cronbach's alpha. Coupled to this, explanation of the higher reliability in the early years requires some consideration of the nature of dental knowledge. Students in their early years will be equipped to answer a limited number of straightforward questions and account for a greater consistency amongst students. More complex questions requiring clinical judgement are less likely to be attempted, leading to a greater internal consistency. With progression through the course, students acquire the skill to answer a great many more questions, some of which have no absolutely correct answer, requiring instead a 'best fit' reflecting the most appropriate choice of several provided. More than one response might be possible, and a

student may select an option which is neither incorrect nor the best fit. Answering a greater number of questions, some of which fall into this category could putatively lead to reduced internal consistency. Given that internal consistency is a product of the calculation of alpha, the function is ultimately based on variance. Thus as the variance in the cohort falls so does alpha. This would be consistent with the lower alphas in Year 4.

There is paucity of published literature on the use of progress testing in dental settings and the application of this form of assessment to undergraduate Dentistry and beyond requires further exploration. Dental schools employ a spectrum of pedagogic approaches from traditional, through hybrid approaches to problem-based learning. Yet, the entry point to and exit points from dental programmes is broadly similar. If applied across programmes, progress testing could be used as a longitudinal assessment tool to compare the growth in applied dental knowledge at between schools or different jurisdictions. Furthermore, if used to evaluate new graduates it could be used to explore knowledge change when crossing the divide between the senior student and the new practitioner.

Conclusion

Although progress testing is a resource and labour-intensive form of assessment, it offers a valid and reliable tool to measure growth in knowledge during a dental educational programme. It is being increasingly used in medical education and there is merit in considering its more widespread use for the assessment of dental students.

Acknowledgements

The authors would like to thank all members of the academic staff who have contributed to Dental Progress Testing. We would also like to thank Dr. Tom Gale (Director of Assessment), Dr. Stephen Bur (Deputy Director of Assessment, Medicine), Dr. Vehid Saleh (Deputy Director of Assessment, Dentistry), Kay Allen (Examinations Officer) and the entire team of Psychometricians who have provided administrative and technical support for the delivery and standard setting of progress tests.

References

1. Van der Vleuten CPM, Verwijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teach*. 1996; 18:103–10.
2. Arnold L, Willoughby TL. The quarterly profile examination. *Acad Med*. 1990;65:515–6.
3. Blake JM, Norman GR, Keane DR, Barber Mueller C, Cunnington J, Didyk N. Introducing progress testing in McMaster University's problem-based medical curriculum: psychometric properties and effect on learning. *Acad Med*. 1996; 71:1002–7.
4. Freeman A, Ricketts C. Choosing and designing knowledge assessments: experience at a new medical school. *Med Teach*. 2010; 32:578–81.
5. Schuwirth L, Bosman G, Henning R, Rinkel R, Wenink A. Collaboration on progress testing in medical schools in the Netherlands. *Med Teach*. 2010; 32:476–9.
6. Nouns Z, Georg W. Progress testing in German speaking countries. *Med Teach*. 2010; 32:467–70.
7. Tomic ER, Martins MA, Lotufo PA, Benseñor IM. Progress testing: evaluation of four years of application in the school of medicine, University of São Paulo. *Clinics (Sao Paulo)*. 2005 Oct; 60(5):389-96.
8. Al Alwan I, Al-Moamary M, Al-Attas N, Al Kushi A, AlBanyan E, Zamakhshary M, Al Kadri HM, Tamim H, Magzoub M, Hajeer A, Schmidt H. The progress test as a diagnostic tool for a new PBL curriculum. *Educ Health (Abingdon)*. 2011 Dec; 24(3):493.
9. Aarts R, Steidell K, Manuel BAF, Driessen EW. Progress testing in resource-poor countries: a case from Mozambique. *Med Teach*. 2010;32:461–3.
10. Verhoeven BH, Van der Steeg AFW, Scherpbier AJJA, Muijtjens AMM, Verwijnen GM, van der Vleuten CPM. Reliability and credibility of an Angoff standard setting procedure in progress testing using recent graduates as judges. *Med Educ* 1999; 33:832–7.
11. Freeman A, Van der Vleuten C, Nouns Z, Ricketts C. Progress testing internationally. *Med Teach*. 2010; 32:451–5.

12. Dijksterhuis MG, Scheele F, Schuwirth LW, Essed GG, Nijhuis JG, Braat DD. Progress testing in postgraduate medical education. *Med Teach*. 2009 Oct; 31(10):e464-8.
13. McHarg J, Kay EJ. Designing a dental curriculum for the twenty first century. *Br Dent J* 2009; 207: 493–497.
14. McHarg J, Kay EJ. The anatomy of a new dental curriculum. *Br Dent J* 2008; 204: 635–638.
15. Bennett J, Freeman A, Coombes L, Kay L, Ricketts C. Adaptation of medical progress testing to a dental setting. *Med Teach*. 2010; 32(6):500-2.
16. General Dental Council. "Preparing for practice-dental team learning outcomes for registration." London: General Dental Council (2011).
17. Cizek, G. J., & Bunch, M. B. (2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. London: Sage
18. R Core Team (2013). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>
19. McAlpine M. (2002) *A Summary of Methods of Item Analysis*. CAA Bluepaper Number 2. CAA Centre, Glasgow. ISBN 1-904020-02-X
20. Schuwirth LW, van der Vleuten CP. The use of progress testing. *Perspect Med Educ*. 2012 Mar;1(1):24-30.
21. McHarg J, Bradley P, Chambelain S, Ricketts C, Searle J, McLachlan J. Assessment of progress tests. *Med Educ*. 2005; 39:221–7.
22. van der Vleuten CP, Schuwirth LW, Muijtjens AM, Thoben AJ, Cohen-Schotanus J, van Boven CP. Cross institutional collaboration in assessment: a case on progress testing. *Med Teach*. 2004 Dec; 26(8):719-25.
23. Drees B.M., Arnold L., Jonas H. (2007) The University of Missouri-Kansas City School of Medicine; Thirty-Five Years of Experience with a Non-traditional Approach to Medical Education. *Academic Medicine*. V 82(4). pp 361-369.
24. De Champlain AF, Cuddy MM, Scoles PV, Brown M, Swanson DB, Holtzman K, Butler A. Progress testing in clinical science education: results of a pilot project between the National Board of Medical Examiners and a US Medical School. *Med Teach*. 2010; 32(6):503-8.
25. Finucane P, Flannery D, Keane D, Norman G. Cross-institutional progress testing: feasibility and value to a new medical school. *Med Educ*. 2010 Feb; 44(2):184-6.

26. Swanson DB, Holtzman KZ, Butler A, Langer MM, Nelson MV, Chow JW, Fuller R, Patterson JA, Boohan M, The Multi-School Progress Testing Committee. Collaboration across the pond: the multi-school progress testing project. *Med Teach.* 2010; 32(6):480-5.
27. Wrigley W, van der Vleuten CP, Freeman A, Muijtjens A. A systemic framework for the progress test: strengths, constraints and issues: AMEE Guide No. 71. *Med Teach.* 2012; 34(9):683-97.
28. Rademakers J, Ten Cate TJ, Bär PR. Progress testing with short answer questions. *Med Teach.* 2005 Nov; 27(7):578-82.
29. Muijtjens AMM, van Mameren H, Hoogenboom RJI, Evers JLH, van der Vleuten CPM. The effect of a don't know' option on test scores: number-right and formula scoring compared. *Med Educ* 1999;33:267–75.
30. Coombes L, Ricketts C, Freeman A, Stratford J. Beyond assessment: feedback for individuals and institutions based on the progress test. *Med Teach.* 2010; 32(6):486-90.