Authors

Demelza J Green [(1)] *

Christine J Heales [(2)]

[(1)] University of Exeter, St Luke's Campus, Heavitree Road, Exeter, United Kingdom, EX1 2LU. d.j.green@exeter.ac.uk

[(2)] University of Exeter, St Luke's Campus, Heavitree Road, Exeter, United Kingdom, EX1 2LU. c.j.heales@exeter.ac.uk

Corresponding author: email d.j.green@exeter.ac.uk

Abstract

Introduction

In March 2020, the first diagnostic radiography degree apprenticeship programme in England was launched at the authors' institution. As part of the programme development and design, the programme development team explored and then implemented progress testing into a strand of the programme.

The objective of this educational perspective is to scrutinise the literature around the use of progress testing in higher education programmes, namely medicine, to explain how and why this decision was reached.

Methods

The initial search strategy was developed using the electronic databases CINHAL Complete and SCOPUS. Key words included 'progress test' and 'medicine' or 'health' or 'education' or 'higher education'. Eliminating articles that were not relevant, and also identifying and adding additional articles by key authors and experts resulted in thirty-three key articles being considered for review.

Results

The thirty-three articles were a mixture of review articles, empirical research, case studies and conference presentations. Five key themes were identified which are discussed in this article; the evolution of progress testing; advantages of progress testing, disadvantages of progress testing, developing a test framework and academic progression and student feedback.

Discussion

Progress testing is now well established in pre-registration medical programmes globally. The advantages of progress testing and the use of frequent look rapid remediation appear to be undisputed. Key disadvantages with progress testing were identified as it being an administrative heavy assessment process as well as a perceived bias towards male students undertaking this type of assessment.

Conclusion

Now this assessment practice is established within medicine, it seems reasonable to explore its use in other areas of healthcare, such as radiography.

Title: Progress Testing: A narrative review exploring the rationale for progress testing and its introduction into a Diagnostic Radiography curriculum.

Introduction.

In March 2020, the first diagnostic radiography degree apprenticeship programme in England was launched at the authors' institution. Any such apprenticeship programme must enable registered apprentices to both demonstrate and meet the occupational duties and knowledge, skills and behaviours (KSBs) laid out in the relevant profession or occupation's published apprenticeship standard [(1)]. With any professional healthcare degree apprenticeship the programme needs to additionally reflect the professional standards of the registration body, in the case of diagnostic radiography this being the Health and Care Professions Council (HCPC) [(2)].

Writing a new programme allowed the programme development team to consider, in depth, the curriculum design, delivery and assessment plan of the apprenticeship. With the development team having some knowledge of the use of progress testing already used within the Medical Degree programme within the authors' institution the development team explored this assessment option and decided to introduce this unique form of assessment into one of the academic strands of the diagnostic radiography degree apprenticeship programme.

The aim of this educational perspective is to scrutinise the literature around the use of progress testing in higher education programmes to explain how and why this decision was reached.

Overview of assessment within Higher Education (HE)

Any course within higher education includes assessments of some format and it is widely recognised that educational assessments in higher education have significant impact on the teaching and learning that occurs within institutions and equally within individual programmes[3].

Up until the early 2010s assessment was often viewed as a way in which academics and higher education institutes could measure outcomes of learning; however, at this point in time a wider body of opinion started to form within academic circles which classified assessment as a fundamental part of the learning process [4]. Furthermore, many authors now argue that assessment should be used as a tool to foster and promote effective learning and engagement thus shifting the focus to assessment for learning as opposed to the historical assessment *of* learning (5, 6). Assessment could, therefore, be seen as a way in which academic skill development can be encouraged as well as being directly related to learning outcomes [7] and the thinking around assessment has altered considerably over the last few decades with much more of a focus on assessment for learning and the promotion of deep learning [8]. However higher education institutes across the globe still use assessment results as a final judgement on whether a student has been successful on their chosen programme of study [9]. Therefore, designing and delivering an assessment strategy which promotes engagement with course material, develops academic skills and promotes deep learning is an on-going area of debate for academics.

It is also agreed that assessment should be used to help both students and academic lecturers identify areas of deficiency and weakness, at both cohort and individual level, in order for consequent delivery of learning materials to be tailored to meet the needs and outcomes of the programme [9].

Opinions regarding assessment are now shifting again with many academics agreeing that a truly authentic and reliable assessment in the 21st century is one that is able to demonstrate and develop high level graduate skills [9]. With so much knowledge and opinion readily available at the 'touch of a button' then the question is whether assessments should now be challenging a graduate's critical analysis and evaluation skills opposed to just recalling facts, thus allowing graduates to demonstrate they have acquired the skills they require for life beyond that of the university [10]. However, in many programmes, especially those with professional body registration such as healthcare programmes including radiography, there is still a core element of knowledge which is required for the graduate to be able to register to practice and to enable the graduate to practice safely and effectively in their chosen profession.

Background to progress testing

Even though first developed in the 1970s progress testing was not widely established until the 1990s within Medical Education [10]. Its introduction was seen as a means of promoting deep learning rather than rote memorisation on problem-based learning programmes [5, 11]. Progress testing is a longitudinal assessment method used to evaluate how a student's knowledge develops over the course of their programme [11, 12]. Since the first introduction of progress testing it is also being utilised in other areas of medical and healthcare education including dental education [13]as well as post-qualification programmes such as Radiology, albeit not in the UK [14]. Progress testing can be defined as 'a test of the complete domain of knowledge considered a requirement for a medical student on completion of his or her undergraduate course' [15]. Therefore, the principle of progress testing is that

each test is benchmarked against the competent qualified practitioner's knowledge base [16] ; this means that in terms of undergraduate education a first year student would not be expected to gain a passing mark but would instead improve (progress) their marks throughout the duration of the programme. It is believed that it is this aspect of this unique assessment technique that prevents the culture of learning only for the test and promotes engagement with all content and therefore deep learning.

Therefore, the concept of repeated assessment may have benefits in other healthcare training programmes such as nursing and allied health professions programmes including radiography or any other educational programme where there are measurable knowledge outcomes required upon graduation [1, 2]. Used in this way, and with appropriate support for analysing the test data, progress testing has the potential to enable individuals to identify specific areas where they may need to improve or where they may need increase their knowledge base in relation to their chosen profession. It should also enable an academic team to identify common themes to better target specific areas of knowledge deficit, for both cohorts and individuals, through the programme of study [12, 17].

As progress testing has recently been introduced into the degree apprenticeship diagnostic radiography programme this review was undertaken as it is beneficial to have a thorough understanding of progress testing including the potential pit falls as well as the advantages of this unique assessment process.

Method

Search strategy

The initial search strategy was developed using the electronic databases CINAHL Complete (Appendix Table 1) and SCOPUS (Appendix Table 2).

Article title, abstract and keywords were searched with no limits placed on the searches due to the relatively small numbers of articles returned. Articles not related to higher education were discarded from the search to ensure it remained relevant.

The second database returned the same articles as the first therefore no further databases were explored.

A second search was used looking specifically at the use of true-false-abstain (TFA) examinations (Appendix Tables 3 & 4) as these are the exact same assessment method as those utilised in medical education progress testing. The same databases were used and article title, abstract and keywords searched with no limits placed on the searches as before.

In addition to the above searches citation chaining was utilised to obtain other relevant literature. Additionally there are a number of key authors and experts who have paved the way for progress testing development and continue to influence its format and delivery so work by these individuals and their collaborators were sought.

Results

This resulted in 33 articles being considered for this educational perspective review. These articles were a mixture of review articles, empirical research, case studies and conference presentations.

From the literature a number of key themes were identified:

1. The evolution of progress testing
2. Advantages of progress testing
3. Disadvantages of progress testing including gender bias in progress testing.

4. Developing a progress testing assessment programme, including the framework, test construction including the use of blueprints, question bank writing including collaborations or consortiums, scoring methods and analysing of results.
5. Academic progression and student feedback.

These themes are discussed in turn below.

Discussion

1. <u>The evolution of progress testing.</u>

Although progress testing was initially developed by two separate institutions in the 1970s it took many decades for this assessment method be adopted more widely, with progress testing not really being used extensively until the 1990s [10].This has been attributed to two key reasons: Firstly, it is agreed that the testing procedure utilised by this assessment method is not easily understood and secondly the academic and administrative burden of progress testing is heavy [10].

A key driver in the implementation of progress testing in medicine was the increased use of problem based learning (PBL) within medical curriculums [10]. With a PBL curriculum the learning is primarily directed by the students and therefore not content based as it is with traditional teaching methods often utilised in other course deliveries. Therefore, each PBL group may end up taking their learning in very different directions making the setting of more traditional closed book end of module or end of year examinations difficult. Many studies demonstrate that the use of progress testing promotes and fosters deep learning strategies [6, 15, 18, 19] however there is also an argument that the fostering of these academic deep learning skills comes from the problem based nature of the curriculum delivery as opposed to the progress testing

alone [20]. So, the progress test could be seen instead as a tool to reinforce and consolidate the knowledge gained through problem based learning, show growth of learning and identify any knowledge 'gaps' [6, 11, 21].

2. Advantages of progress testing

It is widely agreed that the biggest advantage of using progress testing as an assessment method is the deep learning that is instilled into the students registered on programmes utilising it. Many authors agree that progress testing enables students to not just 'learn for the test' and instead the learning is a process of continuous knowledge acquisition which can be used in their studies and future careers when the situation requires such knowledge [6, 15, 18, 19].

It is also believed that progress testing promotes increased knowledge retention as the same knowledge is being examined repeatedly. Further studies have consequently shown that repeated testing leads to increased retention and transfer of knowledge [22, 23]. Due to this 'long-term knowledge and knowledge retention is fostered because question content remains relevant long after the knowledge has been learned' [10].

Progress testing also utilises the 'frequent look rapid remediation' philosophy [24] which allows staff and students to receive regular feedback and measures of their teaching and learning strategies and if required amend the delivery methods and content to suit the needs of both individuals and whole cohorts. It also allows for individual students to be given the appropriate support with their studies as required thus enhancing learning [25, 26].

However, the prolonged use of progress testing internationally has seen other advantages surface. As it is felt that progress tests can provide a rich source of

information for both individual students and cohorts as well as for curriculum evaluators and designers [6] this can lead to the following benefits.

If programmes use a longitudinal method of analysis to make academic judgements on progression then the need for unsuccessful candidates to re-sit a test on every test occasion is removed, progress testing equally allows programme teams to identify both learners who are struggling but also high achievers early in the process and thus allows for appropriate tailoring of each groups' curriculum needs. Finally it is considered that the use of progress testing allows for opportunities to benchmark between different medical schools and curriculums, this additionally also allows comparison of graduates and different methods of curriculum delivery to be achieved, [10], potentially promoting 'best practice' across the profession. This also serves to add additional quality assurance measures and opportunities for programmes to assess their provision and evaluate the effectiveness of this [6, 12].

Another advantage of progress testing comes from the in depth psychometric analysis of the data following each test. Analysis of each individual question can be achieved but also analysis of various demographics of the cohorts which aids in ensuring that no individual group is disadvantaged and therefore showing that the assessment method is effective [4] and is valid, reliable, transparent, inclusive, authentic and fair.

These advantages are not disputed in any of the literature however with the use of progress testing there does come some disadvantages. These disadvantages are mainly associated with the intensity and demands of the test administration and processing of marks.

3. Disadvantages of progress testing

Progress testing is an administration heavy assessment process with quality assurance checks required at various stages plus commitment from all stakeholders as well as significant funding investment in all aspects of the design and application of progress testing being key to its success [12].

One of the main problems associated with progress testing is that in order for students to progress though their course the test marks need to be converted to grades, or deemed to have met a threshold satisfactory level which demonstrates their ability when criterion referencing [15]. If norm referencing is utilised there is no clear pass mark and decisions need to be made on what constitutes a satisfactory mark which allows the student to progress onto the next stage. Therefore, the administrative burden on the psychometric team analysing the test is immense [11]. In addition to the norm referencing methods employed each question is analysed with unreliable questions potentially discarded from final scores. Tests are also analysed to ensure no single demographic of student is unknowingly biased or benefitted. If norm referencing, opposed to criterion referencing, standard setting is chosen to be employed as the means for analysis and grading of the results the programme teams will also need to accept that there will always be a percentage of students who will not meet the satisfactory judgement of an individual assessment. However by employing a longitudinal method then the decisions made should be fair and consistent.

Another disadvantage is the perceived notion that the multiple choice questions, with a ‘don’t know’ option, as utilised in progress testing favour male students.

Gender bias with Multiple Choice Questions

Ongoing debate around gender bias towards certain assessments types continues although most studies are focused at primary and secondary school aged children and therefore it is not fully understood if the same is true in higher education [25]. One such debate relates to whether the multiple choice questions (MCQs), or True False Abstain (TFA) questioning, utilised in progress testing allow male candidates to be advantaged over their female peers. It is well documented that females generally are less likely to take risks compared to their male counterparts who tend to be less risk adverse [25]. It can thus be argued that when undertaking TFA questions associated with progress testing that female candidates are less likely to attempt an answer and instead opt for the don't know options whereas males may take that risk.

Two key papers have explored gender bias with progress testing and the use of TFA questions. A 2009 study by Kelly & Dennick took 16 years' worth of year one and year 2 assessment results and used ANOVA (analysis of variance) to identify if there were any significant differences between the scores for each gender. The results showed that even though females performed better across more assessments, when a range of assessment methods were utilised, the most notable differences between the genders were observed in the TFA progress tests with male students significantly outperforming females [25].

Kacprzyk et al. also explored the effect gender has on different assessment methods used in undergraduate science curriculums concluding 'that the only assessment component showing statistically significant bias against female students was the MCQ exam with negative marking' [28].

Even though the use of negative marking and the 'don't know' option are utilised in order to minimise any advantages between the sexes, both of the studies imply that male students are advantaged in comparison to their female peers.

4. The progress test framework.

It has been recognised that as the use of progress testing grew internationally there would be a need for a systematic framework for higher educational institutions delivering medical education to work within [12]. The authors gathered evidence from the published literature supplemented by their own vast experiences of progress testing to produce this framework [12] . As this framework is written by four, internationally recognised, leading academics in medical education and assessment and published by the Association for Medical Education in Europe (AMEE) it is now seen as the primary framework for progress testing, outlining four key areas which are discussed briefly below:

- Test construction,
- Test administration,
- Results and analysis,
- Feedback to stakeholders.

Test construction.

The recommendation is that each test is built upon a blueprint which defines the knowledge required within the subject area. This blueprint is a basic and fundamental requirement on which the progress test relies to ensure valid and reliable test construction [12] as well as consistency between tests. It also enables test questions to be mapped against the curriculum thereby ensuring the test content aligns with the required learning outcomes. Such a blueprint would typically specify how many questions per topic area are to be contained within the test, as well as the level of

difficulty [12] with questions being classified accordingly [12, 17, 29, 30]. It is recognised that writing appropriate, well-constructed multiple choice questions is both challenging and time consuming, and initial and update training is recommended [6, 12] and this needs to be factored in, if deciding to move to this assessment method.

Finally the programme delivery team need to decide upon the length of the progress test and number of questions included within each test. In medical education the number of questions varies considerably internationally ranging from 125 – 200 questions [12]. Research has shown that shorter tests potentially have less validity than those with an increased number of questions [11, 12, 31]; the number of questions per test also needs to be considered alongside test frequency, and the associated workload.

Test administration.

It is recognised that how the use of progress testing, in medical education, is applied to each course varies considerably across institutions. Differences arise from whether the tests are used in either a formative or summative manner and also whether the institutions decide to use accumulative/aggregated results or single test results when making academic progression judgements. Programme teams also need to decide on whether synchronised testing is employed. Synchronised testing is all cohorts of student take the same test at the same time on the same [12]. The advantages of synchronised testing is that this allows for detailed bench marking but can be difficult to organise [6, 12]. Administering different test papers to different cohorts is simpler logistically but means that benchmarking is not possible so other approaches for ensuring test validity are likely to be needed [11, 17].

A further consideration of test administration is related to test frequency and year level. It is reported internationally that most medical schools appear to include all years of their respective programmes in their progress test assessment regime albeit with marked differences in the frequency of testing and the number of questions within each test [12]. Regardless of which method of administration is employed the key to progress testing's success as both an assessment tool and a teaching and learning tool is the analysis of the results and consequent feedback to the students undertaking the tests.

Scoring of progress tests

Different approaches to scoring/marking progress tests are in use. The first method, known as 'number right' scoring, simply gives a raw mark; i.e. the total number of correct answers given. The second method is known as either correction or formula scoring; in this method there is a penalty, in the form of a negative mark, applied for any incorrect answers selected. There is much discussion around the use of number right marking versus correction scoring when calculating the raw score of each individual undertaking the progress test.

As progress tests utilise multiple choice questions one of the major downfalls of using number right scoring is that there is always a possibility that the right answer is achieved by guess work [15, 31] although it is also argued that there is less bias observed in number right scoring opposed to the commonly employed correction scoring methods used in medical education progress testing [21]. With correction scoring students are penalised with a negative score if they select the incorrect answer for a particular question; the theory behind this is that candidates undertaking the test will be discouraged from guessing the answer. In addition to the use of negative marking

in order to discourage guessing, it is common practice for progress tests employing correction formula scoring methods to utilise a ‘don’t know’ option. Correction scoring has, and continues to, cause much debate amongst academics; where there is agreement that the use of correction scoring; in the form of a negative marking set up; reduces the influence of guessing on the overall test mark it means the candidate sitting the exam is not fully rewarded for what they do have knowledge of [21]. Conversely, the number right scoring methods may aide those with limited or partial knowledge as the candidates sitting the assessment have to make a decision on the question if a ‘don’t know’ option is not available [12, 32, 33].

Overall though; it is widely agreed within medical education that adding a ‘don’t know’ option mimics a ‘real world’ environment due to the fact that if an individual does not know, they can, and should, seek advice from a colleague opposed to ‘guessing’ an answer, embedding this into the test structure has a clear educational benefit [33] .

Additionally it is suggested that the don’t know option is only fair in progress testing as it does not penalise students who may not have covered the curriculum content addressed in the question at the time they sit the progress test; this is especially true in the early stages of a programme (15). Those tests which include the ‘don’t know’ option award a score of ‘0’ marks for this selection and this is now regarded as best practice in the setting and administration of medical progress tests [12, 15, 33].

5. <u>Academic progression and student feedback.</u>

As previously discussed for Medical Education, significant effort is put into ensuring the appropriateness of the question test bank [34] however subsequent analysis of results is equally as crucial in providing an effective assessment method. Academic programme teams have the choice of using either norm referencing or criterion

referencing in the analysis of the tests results with each method having its own merits and perceived faults.

With norm referencing the difficulty of the paper can be accounted for as this method of analysis compares the performance of an individual with similar performances of the cohort taking the same paper; the actual scores in effect become irrelevant as analysis is made of how individuals perform in relation to others sitting the same examination. One of the major issues, however, with using norm referencing is that the programme team need to accept that there will always be a percentage of the cohort who will be deemed to have performed poorly in comparison to their peers and may be seen to have 'failed' the test. In medical education these individuals are not awarded a fail but instead are given an unsatisfactory judgement and then offered remediation [15, 24]. For this reason final academic judgements and decisions of progression tend to be made on an aggregate of grades over a number of tests as opposed to a single final end of programme test.

With criterion referencing the examination is marked against pre-determined, fixed criteria and each individual sitting the examination is assessed independently of each other. With this type of analysis no account is made in regards to the difficulty of the examination and generally, at UK undergraduate level, a pass mark/level is set at 40%. This means that from test to test there can be a varied amount of individuals who are successful in the test and those who fail to meet the 40% pass mark.

Consideration also needs to be made on whether academic progression decisions are made related to one single test alone or the accumulation of previously sat tests over that particular stage of the course. Single point benchmarking needs a cautious approach in the interpretation of results and longitudinal benchmarking is arguably a

more accurate way to interpret accurate student achievement and consequently make an academic judgement on an individual student's performance and progress [21].

Using absolute standards in progress test is challenging [33] until there is a bank of questions with a known difficulty scale from which the paper is constructed. Nevertheless, the use of norm referencing in progress test analysis and decisions over academic progression has become cemented within medical education. Other statistical methods have been explored such as 'equating' which aims to control for differences in difficulty [11], the use of prediction based on prior performance and

linear regression analysis [15] but none of these are in routine use.

Student Feedback

Detailed feedback specifying knowledge areas, skills development, comparison with peers and also with individual previous performances allow the student to direct their own learning as they move through the programme. This information can be useful also to staff delivering remedial interventions on an individual basis and thus allowing more personalised support. It also allows academic teams to evaluate any universal areas for development and modify curriculum delivery if and as necessary [12].

Understanding what support students may require.

Very little of the current literature surrounding progress testing specifically looks at student support. It is acknowledged that some of the authors in the literature reviewed consider the frequent look rapid remediation nature of progress testing as a support mechanism within itself [24]. Furthermore, as discussed previously, there are the claims that progress testing reduces stress and anxiety and therefore it could be argued that due to these phenomenon associated with progress testing use then students do not

need any additional support other than the general support given to all students studying on a course.

Programmes which use Progress Testing

It is increasingly being recognised that the progress test approach may have a broader application than just those educational programmes which utilise problem-based learning [34]. However, it is recognised that progress testing is not a suitable assessment method for many programmes of a heterogeneous nature with early specialism [10].

Bennett et al. discuss the applications of progress testing in dentistry which unlike medical education has less of a focus on gaining a wide range of diagnostic skills and competencies but instead has a more specific focus on gaining skills to deliver suitable and often more complex, irreversible treatments [35]. In this paper the fundamental differences in the knowledge and skills acquired by the different programmes (medicine compared with dentistry) are acknowledged and consequently the need for a slightly different approach to progress testing is required. It was recognised that the style of question utilised in dentistry needs to be different and that made question authoring difficult [35]. However fundamentally the assessment ethos of 'frequent look rapid remediation' still remains key regardless of the discipline it is applied to.

Despite this paper it appears that no other higher education programme or discipline either within or outside of healthcare has adopted this unique assessment method; the reasons for this are not clear considering all the advantages detailed in the literature around its use in medical education.

Conclusion

Assessment still remains integral to higher education programmes although the way in which assessment is considered and implemented is changing rapidly with more academics seeking assessment strategies which promote engagement and deep learning as well as the development of high level graduate skills [5-7].

Progress testing is well established in pre-registration medical programmes throughout the world [12]. It is also utilised in post graduate medicine and dental programmes, however there is no evidence of any other discipline either within or outside of healthcare, including radiography, employing this assessment method.

The benefits of progress testing within the field of medicine are well documented with the consensus of academic opinion being that the use of progress testing develops deep learning of the content material of the programme as well as developing high level academic skills [6, 11, 21]. Bennett et al. even believe that the use of this assessment method additionally aids “applications in the early years of professional practice following graduation” [35] and it could be seen that this would also be of benefit within the field of radiography

Whereas the advantages of progress testing appear to be undisputed one of the key disadvantages is the perceived bias towards male students undertaking this type of assessment regime. The challenges of establishing an appropriate test bank, in terms of time and academic staff development should not be underestimated. Likewise key decisions need to be made in terms of test construction and approaches to marking which require an understanding of this test format. Nevertheless, the use of regular progress testing provides opportunities for frequent feedback for students and academic teams alike.

Progress testing is therefore a unique assessment method which is proven in medical education to promote and develop deep learning and high level academic skills and could have wider application including within radiography education. It comes with some challenges, but the ethos of frequent look rapid remediation is key to enabling students to successfully progress into their chosen careers.

This article has sought to highlight and appraise the evidence base around progress testing, thereby explaining why it is has been selected as an assessment tool within the knowledge based component of the diagnostic radiography degree apprenticeship programme at the authors' institution.

REFERENCES

1. Institute for Apprenticeships & Technical Education 2021 [Available from: https://www.instituteforapprenticeships.org 2021.

2. Council HaCP. Standards of Proficiency - Radiography [23/09/2022]. Available from: https://www.hcpc-uk.org/standards/standards-of-proficiency/radiographers/.

3. Baird J, David Andrich D, Hopfenbeck TN, Stobart G. Assessment and learning: fields apart? . Assessment in Education: Principles, Policy & Practice. 2017;24(3):317-50.

4. Race P. The lecturer's toolkit: A practical guide to assessment, learning and teaching 4th ed: Routledge; 2013.

5. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: From assessment of learning to assessment for learning. Med Teach. 2011;33(6):478-85.

6. Van Der Vleuten CPM, Verwijne GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. . Medical Teacher. 1996;18(2):103 - 9.

7. Croy S. Development of a group work assessment pedagogy using constructive alignment theory. Nurse Education Today. 2018;61:49 - 53.

8. Sambell K, Brown S, Race P. Assessment as a locus for engagement: priorities and practicalities. Italian Journal of Educational Research. 2019:45 - 62.

9. Senel S, Senel H. Remote Assessment in Higher Education during COVID-19 Pandemic. International Journal of Assessment Tools in Education. 2021;8(2):181 - 99.

10. Kolhatkar I. The Authenticity of Digital assessment. LTHEchat [Internet]. 2021. Available from: https://lthechat.com/2021/03/10/lthechat-199-the-authenticity-of-digital-assessment/.

11. Langer MM, Swanson DB. Practical considerations in equating progress tests. Med Teach. 2010;32(6):509-12.

12. 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.

13. Ali K, Coombes L, Kay E, Tredwin C, Jones G, Ricketts C, et al. Progress testing in undergraduate dental education: the Peninsula experience and future opportunities. Eur J Dent Educ. 2016;20(3):129-34.

14. Rutgers DR, van Raamt F, van Lankeren W, Ravesloot CJ, van der Gijp A, Ten Cate TJ, et al. Fourteen years of progress testing in radiology residency training: experiences from The Netherlands. Eur Radiol. 2018;28(5):2208-15.

15. McHarg J, Bradley P, Chamberlain S, Ricketts C, Searle J, McLachlan JC. Assessment of progress tests. Med Educ. 2005;39(2):221-7.

16. Albanese M, Case SM. Progress testing: critical analysis and suggested practices. Adv Health Sci Educ Theory Pract. 2016;21(1):221-34.

17. Swanson DB, Holtzman KZ, Butler A, Case Western Reserve University School Of Medicine Cumulative Achievement Testing Study G. Cumulative achievement testing: progress testing in reverse. Med Teach. 2010;32(6):516-20.

18. Blake JM, Norman GR, Keane DR, Mueller CB, 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(9):1002-7.

19. van Berkel H, Nuy H, geerligs T. The influence of progress tests and block tests on study behaviour. Instr Sci 1994;22:317 - 33.

20. Kilroy DA. Problem based learning. Emerg Med J. 2004;21(4):411-3.

21. Muijtjens AM, Schuwirth LW, Cohen-Schotanus J, Thoben AJ, van der Vleuten CP. Benchmarking by cross-institutional comparison of student achievement in a progress test. Med Educ. 2008;42(1):82-8.

22. Butler AC. Repeated testing produces superior transfer of learning relative to repeated studying. J Exp Psychol Learn Mem Cogn. 2010;36(5):1118-33.

23. Roediger HL, 3rd, Butler AC. The critical role of retrieval practice in long-term retention. Trends Cogn Sci. 2011;15(1):20-7.

24. Freeman AC, Ricketts C. Choosing and designing knowledge assessments: experience at a new medical school. Med Teach. 2010;32(7):578-81.

25. Kelly S, Dennick R. Evidence of gender bias in True-False-Abstain medical examinations. BMC Med Educ. 2009;9:32.

26. Kaufman A. Implementing problem-based medical education: Lessons from successful innovations1985.

27. Smith F. "It's not all about grades": accounting for gendered degree results in geography at Brunel University. Journal of Geography in Higher Education. 2004;28(2):167 - 78.

28. Kacprzyk J, Parsons M, Maguire P, Steward G. Examining gender effects in different types of undergraduate science assessment. Irish Educational Studies. 2019.

29. 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.

30. Nouns ZM, Georg W. Progress testing in German speaking countries. Med Teach. 2010;32(6):467-70.

31. Downing SM. Threats to the validity of locally developed multiple-choice tests in medical education: construct-irrelevant variance and construct underrepresentation. Adv Health Sci Educ Theory Pract. 2002;7(3):235-41.

32. Cecilio-Fernandes D, Medema H, Collares CF, Schuwirth L, Cohen-Schotanus J, Tio RA. Comparison of formula and number-right scoring in undergraduate medical training: a Rasch model analysis. BMC Med Educ. 2017;17(1):192.

33. Muijtjens AM, Hoogenboom RJ, Verwijnen GM, Van Der Vleuten CP. Relative or Absolute Standards in Assessing Medical Knowledge Using Progress Tests. Adv Health Sci Educ Theory Pract. 1998;3(2):81-7.

34. Albanese M, Case SM. Progress testing: critical analysis and suggested practices. Adv Health Sci Educ. 2016;21(1):221-34.

35. 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.

Appendix

#	**Search terms**	**Results**
1	"Progress test*"	110
2	"Progress test*" and medicine	23
3	Progress test*not medicine	87

4	Progress test* and health*	17
5	Progress test* and education	82
6	Progress test* and higher education	0

Table 1 – data base search results – CINHAL Complete

#	**Search terms**	**Results**
1	Progress test*	318
2	Progress test* and medicine	48
3	Progress test* and higher education	81

Table 2 – data base search results – SCOPUS

#	**Search terms**	**Results**
1	true-false-abstain	8
2	True and false and abstain*	0
3	True and false and abstain and medicine	0
4	True and false and abstain and assessment	0

Table 3 – data base search results – CINHAL Complete

#	**Search terms**	**Results**
1	true-false-abstain	1
2	True and false and abstain*	1

3	True and false and abstain and medicine	1
4	True and false and abstain and assessment	1

Table 4 – data base search results – SCOPUS