Making Progress: Introducing progress testing approaches to a single semester paramedic subject

James Thompson (✉ james.thompson@flinders.edu.au)

Flinders University <https://orcid.org/0000-0002-0473-3266>

Don Houston

Flinders University

Research article

Keywords: Progress Test, Programmatic Assessment, Paramedic Education, Capstone

Posted Date: July 11th, 2019

DOI: <https://doi.org/10.21203/rs.2.11233/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Abstract Background. Paramedicine is a rapidly evolving profession with continually increasing demands placed upon educating its future workforce. Ensuring graduates are adequately prepared places many expectations on the suitability and effectiveness of university assessment practices, in a discipline where summative credentialing has established traditions of use. Progress tests and programmatic assessment have growingly become common fixtures of medical education, offering longitudinal information about student knowledge, ability and progress, usually across an entire program of study. **Methods.** Our project explored the development, implementation and evaluation of progress testing in a single semester capstone undergraduate paramedic topic. We examined the changes in student performance between two MCQ tests spaced ten weeks apart, and performance in a final oral assessment based on the same test content. Student perceptions and experiences of these events were also evaluated. **Results.** 55% of students indicated it was common practice to guess answers in exams. After introducing of negative marking students achieved 40% mean correct answers on previously satisfied curriculum content in our test 1. Scores increased by 65% by test 2, with substantial declines in numbers of incorrect and don't know responses. **Conclusion.** Our results demonstrate a substantial increase in correct responses between the two tests, a high mean score in the viva, and broad agreement about the significant impact the approaches have had on learning growth.

Background

Since the start of university based paramedic education in Australia two decades ago, educators have faced challenges in preparing graduates for the highly specified paramedic role when using traditional teaching approaches (1). Transitioning from the former vocational, largely 'hands on' model of training, to more theoretical classroom-based approaches has not been without friction. Anecdotally, industry frequently value a graduate's traits of 'road readiness' ahead of academic achievement or GPA. Road readiness as a yardstick which is both difficult to quantify and subject to differing interpretations (2) presents an ongoing challenge to academics striving to meet stakeholder demands in addition to ensuring students achieve key learning milestones prior to graduation. As their undergraduate studies draw towards completion, there are expectations that students have obtained and retained knowledge from all prior curriculum events and are capable of applying it at will. Student assessments during these final periods of study routinely reflect student credentialing, as subject specific increments of learning continue to be satisfied. Seldom are students assessed on the comprehensive knowledge required of the discipline, the breadth of an entire curriculum, or through assessments designed in context of the discipline (2, 3).

To enhance graduate standards, our capstone program was first developed a decade ago, and continues to be refined and responsive to changing student, academic and discipline needs. Central to all iterations of our capstone subject has been the coordinated use of assessment for learning. Our subject has steadily evolved into a model of programmatic assessment for learning (PAL), a design commonplace in medical education which features assessment of student knowledge across an entire broad body of

curriculum representative of expectations of the field of study (4). One of the main tools used to achieve this is the progress test; a comprehensive exam designed to evaluate student mastery of knowledge, administered at regular intervals across all the years of their study (5). We sought to explore whether progress testing could be effectively introduced to paramedicine, and if an approach which is typified by repeated testing over a whole course, could be effectively applied within a single semester subject.

This paper describes the context for our innovation and the collaborative process we used to develop the test instrument. We explain our integration of the progress test into the student learning experience, our various approaches to evaluation and an analysis of our findings and implications.

Capstone paramedic developments

For a decade, we have been developing and evaluating teaching, learning and assessment innovations in a final year, single semester capstone topic of a Paramedic degree in an Australian university. This subject focuses on the goal of 'bringing it all together' for student learning and making sense of all the material previously covered in the degree in preparation for the transition to paramedic practice.

Previous cycles of action research have resulted in modifications to the subject's pedagogy, principally responding to issues relating to the student relationships with assessment and its impacts on their learning. Examination related stress, grade seeking behaviour, students' reluctance to accept critical feedback and poor engagement with learning had routinely challenged teaching staff. Incremental subject changes since the initial redesign place far greater emphasis on formative learning, feedback to students and far deeper levels of student understanding. While the new subject assessments were clear improvements on previous practice, opportunity existed for further improvement. Two unresolved issues remained: validating the expected knowledge for practice in the discipline, and addressing assurances that grades accurately reflected overall student capability and knowledge.

Progress testing was introduced and evaluated as an additional design element in the second semester of 2018 academic year.

Progress Testing (PT)

An established feature of medical degrees in the Netherlands for several decades (6), the test enhanced learning approach is now a global phenomenon (7). Initially introduced in response to the effects that examinations were having on driving rote learning among students, the progress test (PT) seeks to develop deeper student understanding (8). It also answered a call for a suitable assessment strategy to respond to self-directed PBL based curriculum models (9) (10). With PTs, conventional single summative tests are replaced by a series of similar repeated tests dispersed across an entire program of study with every enrolled student across all year levels sitting the same test simultaneously (11). Students' broad understanding is tested and retested (12). This acknowledges that the outcomes of a single test are likely to be a less reliable indicator of student ability than multiple samples of testing dispersed over a period of time (13). A carefully choreographed suite of low stakes assessments providing maximal feedback, enable the student to be self-aware of their ability levels and development (12, 14). Where

traditional teaching programs offer insight to student incremental learning steps, they are unable to validate student mastery of the full, inter-related curriculum (15). Progress tests are designed to represent the full breadth of functional knowledge required for the discipline, are not aligned to any one subject or student year level, and commonly involve large samples of questions drawn from large question pools (4) (16). Sampling such a holistic breadth of curriculum, it is considered near impossible for a student to cram or binge learn as preparation for the test, instead learning is more evenly regulated across a full program(17) (8). This type of test directed learning is linked to both learning enhancement as well as offering educators more reliable indicators of student knowledge retention (13). Replacing simple passive measurement of student knowledge with a tool which is an active driver for learning has led many to re-think how they regard assessment (18). Commonly multiple-choice exam formats are used for testing (19), however, examples of alternative assessment styles such as OSCEs in medicine have been recently emerging (20). The merits of PTs have also seen their broader application to disciplines beyond medicine such as dentistry (21), although we found no reported instances of PT use within paramedic education. Despite our subject comprising only a single semester of a 3 year teaching program, we felt that a comprehensive test series which could provide students with rounds of feedback set against the discipline's knowledge requirements, matched the ethos of our capstone approach. Our initial steps were to establish and validate the knowledge expectations of the discipline.

Methods

Determining Paramedic Knowledge: Paramedic learning list

Australian universities offering paramedic education are guided by the 'Paramedic Professional Competency Standards' produced by the Council of Ambulance Authorities (CAA) (22). This document broadly specifies the expectations for paramedic practice within industry, which by inference determines the goals of any underpinning education and training (1). Broad statements are presented under the headings of 'Professional Expectations of a Paramedic' and 'Knowledge, understanding and skills required for Practice'. These are neither an exhaustive list of knowledge or skill components, nor specific instructions, but represent an equivocal set of points which are able to be translated for the vastly differing dialects of Australia's ambulance services and education providers. In the absence of definitive detail to inform the specific graduate knowledge, we set about establishing and validating these requirements. Using an existing undergraduate curriculum that had been recently endorsed by an external national accreditation board, each learning outcome, all teaching and assessment artefacts were reviewed and itemised, simultaneously alongside all clinical practice guidelines (CPGs) which represent instructions for practicing local paramedics. Clinical skills and knowledge required to enact these instructions were incorporated into extensive item lists that were then subject to processes of validation. Academic staff responsible for teaching design and delivery reviewed the lists in relation to their own teaching and curriculum priorities. In addition, several senior paramedic clinicians from the local industry were invited to review items, offering opinion regarding the significance of items to the practice of paramedics. A third phase of development involved mapping and linking each of these items, a process which united concepts normally the domain of one topic with those from others and with features of

practice requirements drawn from the CPGs. This scaffolding process integrated concept themes such as anatomy, pathophysiology, pharmacology, clinical skills and field instructions. Despite these broad subject areas usually representing pre-requisite requirements, student knowledge usually had been evaluated in isolation. An exhaustive process of itemising, accounting and organising key items of learning and paramedic practice enabled us to produce a template comprising of primary items and four connected subsidiary items. As an example, a primary theme of 'myocardial infarction' was linked with subordinate themes of referred pain pathways, interpretation of ischaemic 12 lead ECG changes, the contraindications for GTN use, and the anatomical location and features of coronary arteries. In this case a primary concept with a focus on pathophysiology, draws together physiology, diagnostic skills, pharmacology and anatomy knowledge. This exercise in mapping, distilling and aligning concepts generated a framework which underpinned our progress test. This was again considered by our review team for legitimacy and perceived relevance to both university curriculum and paramedic practice. The product which resulted was a prioritised and validated list of 100 primary concepts, each aligned to 4 sub-concepts (400 in total).

The test design: TEST QUESTIONS

Once the list had been established and validated, constructing test questions commenced, with the list providing the framework for both question and four potential responses (one correct and three distractors). Throughout the design process a goal was to ensure that assessments represented a faithful measurement of student knowledge. We deliberately sought to reduce student results which were obtained through chance or through only partial topic knowledge used to eliminate obvious incorrect distractors. We aimed to create items which a student 'who knows' would be able to get correct but a student 'who doesn't know' would be unlikely to get correct. Consequently, test outcomes would be less likely to reflect false positive or false negative performances (23). Literature and resources on optimal assessment design were consulted, and the revised taxonomy of multiple-choice item writing guidelines tool (24) was applied as a filter during question composition and editing phases. Consistency in response item length and opening wording were carefully considered to ensure that item structure was unlikely to be a factor influencing the student response decision. A small working group of academic staff, recent graduates and senior paramedics then participated in a series of question review sessions to ensure relevance, non-ambiguity, fairness and balance, as a final validation of the question set with particular attention to content, format, style and writing.

Marking & Grading Decisions

Considerable disagreement exists on optimal test marking approaches. Central to the debate is the capacity of differing approaches to provide a true account of student knowledge(25) (26). In the case of simple 'marks for correct answer' approach, criticism relates to assessors being unaware of the extent a final score is achieved from chance (27). Alternatively, negative marking approaches which seek to discourage students guessing through penalising incorrect answers attract criticism for the additional test related anxiety these create for some students, while others suggest that it is infrequent that students

entirely guess a response, but instead use deduction informed by some knowledge (25) . One point of consensus is that there is no one optimal measure, but instead a need for assessment design to consider local need and context (28). It was the specific context of our discipline which ultimately informed our grading decisions. Paramount to the practice of paramedics is the requirement that all clinical decisions are founded on effective knowledge for practice, with a high degree of risk aversion and clinician recognition of their own limitations (22). It was our desire for test practices to echo this philosophy.

Conscious of the critical negative marking rhetoric it was still felt that reducing chance results and encouraging students to self-identify material they had not yet mastered was consistent with our wider topic intentions. The construct of our test distractors involving differing domain knowledge was intended to counter deductive elimination based on partial knowledge guesses. In the case of a student who was unsure of the correct answer, our preference was that they choose the 'don't know' option and received the structured learning support featured within our subject pedagogy. Our final student test scores were designed to reflect a summary of correct minus incorrect responses.

Test Implementation

Progress Test 1

Progress test 1 was administered on the first day of the semester. Typically progress testing is introduced with no prior exposure to material being examined. By contrast our students had previously covered nearly all content across two and a half years of the teaching program. While they had previously satisfied the assessment milestones of pre-requisite subjects, their knowledge had been examined solely within the boundaries of individual subjects and not the broader context of the pre-hospital setting requirements.

Students were required to select a single correct answer from four possible options or 'don't know', with three options being distractors. The first test event was entirely formative introducing students to the PT experience and offering early performance benchmarking and self-reflection opportunities. Negative marking applied to incorrect answers and students received a zero mark for each unanswered or declared unknown answer. As each question shared a direct relationship to knowledge expectations for practice, we wanted incorrect answers to show that there would be foreseeable consequences associated with judgement or practice errors, while also considering areas of strengths and weaknesses in their understanding of the curriculum.

A common practice with PTs is to provide students with a copy of their exam questions, to encourage students to continue to reflect beyond the test, noting problems encountered when tests feedback is withheld (29). We made a decision to deviate from this and provided students access to their results and a copy of the learning list which corresponded directly with each individual question. This directed student learning towards identified knowledge gaps (incorrect questions) with a corresponding learning list, supporting learning while also preserving the question set for subsequent test use and enabling us to make direct comparisons on the two tests. Replacing the exam questions with the learning list was intended to encourage the development of broader student understanding and discourage students from

being distracted by debating question semantics rather than investing effort in learning. We offered students the opportunity to seek additional clarification in a face to face meeting with staff, where additional feedback or concerns could be explored.

Learning List in Teaching

In addition to providing the framework for exam questions the learning list permeated all other areas of teaching. Classroom problem-based learning (PBL) sessions constructed around authentic cases steered students through selections of items from the learning list presented in context of actual patient cases and reasoning challenges. Students were encouraged to recognise the context around material that had been tested and to challenge their understanding through collaborative PBL activities. At the close of each class, each student was required to self-nominate a list item to research further before reporting back to a group shared wiki platform. Students co-constructed a collective body of information and sourced links supporting the learning of the group. Over the semester each group compiled entries for the all items on the learning list producing a comprehensive database of shared study resources which corresponded to the PT content.

Practical classes were also mapped to the learning list to encourage a hands-on application of required knowledge. Simulated scenarios mimicking 'on-road' events, required students to work through a defined discipline specific paramedic process of care (30). Student responses, performance and judgement formed the basis of these events, which were calibrated against the guidance of their paramedic tutors in a consensus-based assessment approach, a largely self-regulated approach to learning requiring student self-critiques of their efforts (31). This format of alternating PBL, practical classes and online wikis connected through the learning list was used for a ten-week cycle prior to students repeating the identical progress test for the second time.

Progress Test 2

The identical test was re-administered at week 11 with student marks this time contributing to their final grade for the subject. Questions were again retained by staff at the close of the exam and similarly feedback on performance was channelled through the learning list. This time the test was also used as a diagnostic tool with results informing a personalised oral exam, unique to the gaps identified for each student.

Viva/ Oral Exam (Test 3)

Ambulance industries routinely use a viva style approach as a means to determine knowledge or competence particularly during recruitment. Despite the importance placed on a graduate's ability to respond well, students have little exposure or preparation, which influenced our inclusion of vivas within our assessment strategies. Vivas are noted for enabling face to face judgements of student competence beyond what is achievable within a written exam (32). For many of our graduates these also represents one of the next major hurdles they will encounter potentially with high stakes attached to their

performance (2). Industry vivas often require the participant to respond to broad, open-ended questions before a small panel, who evaluate the quality of the candidate's response. Unlike the unknown element of the questions they may encounter within a job interview we were transparent about the possible questions they would face. Students were informed they were to be questioned on items they had been unable to answer correctly in PT 2 and had approximately four weeks following the PT2 to target remaining gaps in their knowledge which had been identified for a second time. The strategy directing maximal learning efforts towards students' weakest areas of understanding. Each viva was assessed by two practicing paramedics who were provided with copy of each student's PT2 results. Once again using the learning list mapped to student PT results assessors selected 3 items from the student's unique results profile. During a 15-minute interview the student shared their understanding of these 3 items, paramedics awarding summative scores for accuracy, depth and breadth of information provided. The viva marking the final step of the interrelated test-driven learning experience illustrated in figure 1.

Results

Evaluation of the innovation

We analysed student performance in the identical tests administered 10 weeks apart as well as performance in the final viva. Additionally, a questionnaire was administered to participants recruited from the student cohort. (Ethics Approval was obtained through the Social and Behavioural Research Ethics Committee.) Students were notified of the study by email prior to commencing the subject and advised their participation was entirely voluntary and assured their responses would not be identifiable. Participant responses were obtained via a paper questionnaire administered to students during the last contact day of the subject with students required to rate their level of agreement with statements as well as being provided an option for free text to provide additional comment. All responses were de-identified and responses entered to a spreadsheet for analysis.

All 103 students (101 internal and 2 distance education) attempted both progress tests. Initial measures considered the differences in student responses over the 2 tests (correct, incorrect and don't know responses). Key characteristics of the results are presented in table 1 and figure 2.

Our data shows a noticeable positive shift in all responses. Analysis showed a 64% increase in the total number of correct responses between the tests, conversely there was narrowing of the response ranges and mean number of responses in both incorrect and 'don't know' categories. Together these data suggest that substantial learning occurred between the two tests.

Final Viva Results:

103 students participated in oral viva assessments in week 15 of the semester. A class mean correct answer rate in PT2 resulted in the average number of 35 different potential viva topics across students, each with 4 related themes which could be explored (140 total potential discussion items). The mean

student score in this assessment was 71% with the range: 23-100%. More than 10% of the class achieved 100%.

Student Perceptions:

88 students (91%) voluntarily completed the survey directly following the final viva. The table below illustrates the level of broad agreement obtained from the survey. Survey questions were designed to capture student perceptions relating to their experiences with the test, its effects upon their learning, and value of the approaches.

Free text responses proved additionally informative. Thematic analysis of this data showed that student feedback could mostly be organised within a small number of different classifications. Themes were most commonly reflected; *1. Challenging Experience Good, 2. Challenging Experience Bad, or 3. Personal Development/ Achievement.* When these were considered in parallel with the quantitative responses it appeared consistent that students found the test to reflect the breadth of curriculum and effective at identifying knowledge gaps and challenging them to learn, however students appeared divided over how well they received and responded to this.

Comments included;

“forced me out of my comfort zone”, “it was terrifying but very helpful in the end”, “Stressful but effective”, “challenging....definitely learned a lot”“more confident”

Similarly critical reviews *“felt discouraged from choosing (when didn’t know answers)”, “A lot of content to cover in a short time which made me feel pressured & stressed”, “difficult if you are not comfortable being scrutinised”, “stressful to get my abilities to expected standards”*

Student comments also reflected on the learning process.

“Made to learn in a comprehensive manner”, “good preparation for the future” “very useful -Broad study was required ...exactly what we need...” “learning to self-learn is more valuable than being spoon-fed information”

The 2 test scores, quantitative ratings and qualitative reports appear to support a similar conclusion: the subject was challenging but highly effective at generating learning and engaging students.

Discussion

It could be argued that a rigorous national industry course accreditation process attends to the task of defining what a paramedic needs to know, as well as determining the effectiveness of universities to deliver on this. When it comes to curriculum and teaching detail however, reviews consider a mostly macro level of education. Development of our learning list in collaboration with members from industry was critical to the identification, interpretation and validation of specific content detail. Previously

university curriculum and industry-based practice guidelines had been considered and developed by both groups in isolation, or with ad hoc opinions sought. Our collaborative test building approach advanced a mutual appreciation and addressed assumptions from each camp and our decision to include several recent graduates to the review committee provided invaluable insight to student reactions and test strategies during design.

Capstone subjects and progress tests may appear incompatible at first glance. PTs offer longitudinal student performance data, encouraging paced learning across a whole program and discouraging intensive bursts of isolated study for tests, where capstones represent a final learning push. PT avoidance of cramming and binge learning (13) is challenged in intensive single semester delivery. They do however share some important common ground. Both aim to facilitate learning through immersing students into a full experience of the discipline, its practices, knowledge and expectations. We accept that the confines of a single semester period mean we forgo the beneficial longitudinal performance data. However, data from 3 tests (2 MCQ and 1 viva) is a marked improvement on student data achieved from the former single summative test. A conventional PT philosophy discourages student focus on test preparation as a strategy to avoid superficial and less sustainable rote learning. In contrast we repeatedly promoted our learning list, openly advertising the 400 items (relating to 100 questions). Essentially these represented a comprehensive set of mini learning outcomes which students were to be measured against on 3 occasions during the single semester. Where PTs direct student focus to the wider curriculum instead of a test, we potentially met this ideal part way with our design.

Comparing PT1 and PT2 results the 64% increase in total correct student responses and reductions in incorrect (33.5%) and don't know responses (47%) in addition to the student reported experiences, and the observed paramedic assessor feedback suggest considerable learning growth. Improvements to student test scores in an examination they had previously attempted, following 10 weeks of focussed teaching and learning design may seem unremarkable and likely a predictable result, but this does not represent the complete picture. This was by far the most comprehensive test the students had encountered in the history of our degree delivery and represented knowledge critical to their future work as paramedics. Mastery of 400 learning items deemed essential to on-road practice places greater stakes beyond a simple test score result, with foreseeable consequences linked to knowledge gaps or poor decisions. Until now students had not been measured on their 'whole knowledge' and the broader expectations of the paramedic role. Nor had they been previously exposed to a correct minus incorrect scoring approach. To a cohort of previously high achieving students embarking on their final academic phase, many already with conditional offers of employment, an adjusted class mean score of 14% on PT 1 close to the end was extremely confronting. We were very interested to explore the effect our first use of negative marking had upon student test behaviour and posed a survey question about what amount of negative weighting it would take to deter students from guessing an answer in a test. While the responses varied, -1 was the most common response with 35% supporting this. Remarkably 8.3% indicated that no weighting amount would stop them from guessing to potentially optimise their scores. More than half of the respondents (55%) indicated it was normal for them to guess answers in exams. With the PT almost entirely reflecting curriculum students had previously satisfied, these responses coupled with a PT1 correct score of only

40% led us to contemplate the extent to which earlier pre-requisite degree milestones had been a product of chance.

While the numbers of correct, incorrect and don't know responses all showed pleasing shifts between the tests, student attitude towards the PT1 result proved pivotal to their success. Students more willing to accept the critical PT1 results proved far quicker to engage with the learning structure of the subject and respond to knowledge gaps. The free text feedback results echoed this with the identified 'challenging experience good' and 'challenging experience bad' themes, suggesting that most students were challenged by the testing process, however personal 'like' or 'dislike' for the process of being challenged featured prominently. Many embraced the testing format and opportunities to target knowledge gaps, while others struggled with receiving such extensive critical feedback and vehemently defending a right to chance test results. The perceived impacts on GPA close to their course completion outweighed any learning benefit of the innovation for these few. PT claims to being linked to a reduction in test related anxiety (4) was certainly different to our own experience when applied to a single semester topic and for the first time for our students. Regardless of the purely formative nature of PT1 the results were clearly inconsistent with many student's performance expectations. Student awareness that the next time they would face a summative test on the same instrument which had left much knowledge very exposed proved a source of some nervousness for much of the semester.

By retaining and re-using the same test questions for PT 1 and PT 2 we attempted to address concerns of students memorizing questions ahead of prioritising substantive learning through frequent requirements of students to demonstrate their knowledge in PBL and practical exercises. The inclusion of an oral viva further encouraged deeper student understanding. We have no way of establishing if students did or did not memorise any of the questions, however during a subject exit interview, students shared how unfamiliar they felt with the specific questions wording after having been so focussed on the learning list, with several students conveying their genuine surprise that the 2 tests were identical.

Our decision to include a viva to the PT offered a different twist on the versatility of PTs. Although there are examples in the literature of alternatives to MCQ PTs, such as OSCEs (20), we were unable to find reports of the use of PT content across several linked assessment formats. We had introduced the viva assessment in an earlier iteration of the subject and have found it continues to be well received by students.

Regardless of whether they liked or disliked the test driven design, there was clear consensus the method had been highly effective at contributing to relatively rapid learning growth.

Conclusions

100 core themes and 400 related concepts required for paramedic knowledge were defined and validated before we set about implementing teaching and learning strategies to achieve student mastery of this these. Our context was a single semester capstone undergraduate subject tasked with addressing graduate preparation for entering the profession of paramedicine. We introduced principles of the

progress testing approaches which has had extensive success in medical education, adapted and applied them to a single semester paramedic subject. We are confident that our methodology was successful in testing students' deeper understanding and made a significant contribution to their learning at this critical phase as beginning clinicians. The capstone subject seeks to 'bring it all together' for students at the closing stages of their studies and the PT use was able to clearly share knowledge expectations and transparently measure students against these, making it a valued inclusion to the teaching design. While many students found the experience challenging, our assessments and expectations remained transparent from the outset. Our design offers an example for other single subjects and the learning list approach suggests alternative approaches to releasing exams to students.

Abbreviations

CAA Council of Ambulance Authorities

CPG Clinical Practice Guidelines

ECG Electrocardiogram

GPA Grade Point Average

GTN Glyceryl trinitrate

MCQ Multiple Choice Question

OSCE Observed Structured Clinical Examination

PAL Programmatic Assessment for Learning

PBL Problem Based Learning

PT Progress Test

References

1. O'Brien K, Moore A, Dawson D, Hartley P. An Australian story: paramedic education and practice in transition. *Australasian Journal of Paramedicine*. 2014;11(3).
2. Thompson J, Grantham H, Houston D. Paramedic capstone education model: Building work ready graduates. *Australasian Journal of Paramedicine*. 2015;12(3).
3. Houston D, Thompson JN. Blending Formative and Summative Assessment in a Capstone Subject: 'It's not your tools, it's how you use them'. *Journal of University Teaching & Learning Practice*. 2017;14(3):2.
4. Heeneman S, Schut S, Donkers J, van der Vleuten C, Muijtjens A. Embedding of the progress test in an assessment program designed according to the principles of programmatic assessment. *Medical*

- teacher. 2017;39(1):44-52.
5. Wrigley W, Van Der Vleuten CPM, Freeman A, Muijtjens A. A systemic framework for the progress test: Strengths, constraints and issues: AMEE Guide No. 71. Medical Teacher. 2012;34(9):683-97.
 6. Vleuten CVD, Verwijnen G, Wijnen W. Fifteen years of experience with progress testing in a problem-based learning curriculum. Medical Teacher. 1996;18(2):103-9.
 7. Howe A, Campion P, Searle J, Smith H. New perspectives—approaches to medical education at four new UK medical schools. Bmj. 2004;329(7461):327-31.
 8. Van Berkel HJ, Nuy HJ, Geerligs T. The influence of progress tests and block tests on study behaviour. Instructional Science. 1994;22(4):317-33.
 9. Vleuten CPMVD, Verwijnen GM, Wijnen WHFW. Fifteen years of experience with progress testing in a problem-based learning curriculum. Medical Teacher. 1996;18(2):103-9.
 10. Tio RA, Schutte B, Meiboom AA, Greidanus J, Dubois EA, Bremers AJ. The progress test of medicine: the Dutch experience. Perspectives on medical education. 2016;5(1):51-5.
 11. Coombes L, Ricketts C, Freeman A, Stratford J. Beyond assessment: feedback for individuals and institutions based on the progress test. Medical teacher. 2010;32(6):486-90.
 12. Muijtjens AM, Schuwirth LW, Cohen-Schotanus J, van der Vleuten CP. Differences in knowledge development exposed by multi-curricular progress test data. Advances in Health Sciences Education. 2008;13(5):593-605.
 13. Schuwirth LW, van der Vleuten CP. The use of progress testing. Perspectives on medical education. 2012;1(1):24-30.
 14. Hauff SR, Hopson LR, Losman E, Perry MA, Lypson ML, Fischer J, et al. Programmatic assessment of level 1 milestones in incoming interns. Academic Emergency Medicine. 2014;21(6):694-8.
 15. Verhoeven B, Verwijnen G, Scherpbier A, Van Der Vleuten C. Growth of medical knowledge. Medical education. 2002;36(8):711-7.
 16. McHarg J, Bradley P, Chamberlain S, Ricketts C, Searle J, McLachlan JC. Assessment of progress tests. Medical education. 2005;39(2):221-7.
 17. Muijtjens A, Hoogenboom R, Verwijnen G, Van der Vleuten C. Relative or absolute standards in assessing medical knowledge using progress tests. Advances in Health Sciences Education. 1998;3(2):81-7.
 18. Freeman A, Van Der Vleuten C, Nouns Z, Ricketts C. Progress testing internationally. Medical teacher. 2010;32(6):451-5.
 19. Ricketts C, Freeman AC, Coombes LR. Standard setting for progress tests: combining external and internal standards. Medical education. 2009;43(6):589-93.
 20. Pugh D, Touchie C, Wood TJ, Humphrey-Murto S. Progress testing: is there a role for the OSCE? Medical education. 2014;48(6):623-31.
 21. Bennett J, Freeman A, Coombes L, Kay L, Ricketts C. Adaptation of medical progress testing to a dental setting. Medical teacher. 2010;32(6):500-2.

22. CAA. Council of Ambulance Authorities : Paramedic Professional Competency Standards Version 2.2. 2013(June 12, 2019).
23. Schuwirth LW. Assessing medical competence: finding the right answers. The clinical teacher. 2004;1(1):14-8.
24. Haladyna TM, Downing SM, Rodriguez MC. A review of multiple-choice item-writing guidelines for classroom assessment. Applied measurement in education. 2002;15(3):309-33.
25. Lesage E, Valcke M, Sabbe E. Scoring methods for multiple choice assessment in higher education– Is it still a matter of number right scoring or negative marking? Studies in Educational Evaluation. 2013;39(3):188-93.
26. Burton* RF. Multiple-choice and true/false tests: myths and misapprehensions. Assessment & Evaluation in Higher Education. 2005;30(1):65-72.
27. Burton RF. Quantifying the effects of chance in multiple choice and true/false tests: question selection and guessing of answers. Assessment & Evaluation in Higher Education. 2001;26(1):41-50.
28. Burton* RF. Multiple choice and true/false tests: reliability measures and some implications of negative marking. Assessment & Evaluation in Higher Education. 2004;29(5):585-95.
29. Wade L, Harrison C, Hollands J, Mattick K, Ricketts C, Wass V. Student perceptions of the progress test in two settings and the implications for test deployment. Advances in health sciences education. 2012;17(4):573-83.
30. Carter H, Thompson J. Defining the paramedic process. Australian journal of primary health. 2013.
31. Thompson J, Houston D, Dansie K, Rayner T, Pointon T, Pope S, et al. Student & tutor consensus: a partnership in assessment for learning. Assessment & Evaluation in Higher Education. 2016:1-11.
32. Torke S, Abraham RR, Ramnarayan K, Asha K. The impact of viva-voce examination on students performance in theory component of the final summative examination in physiology. Journal of Physiology and Pathophysiology. 2010;1(1):10-2.

Tables

Table 1. Student comparative results Progress Test 1 & Progress Test 2

	Range correct	Mean correct	Range incorrect	Mean incorrect	Range 'don't know'	Mean 'don't know'
PT1	1-68 (67)	40	1-68 (67)	28	1-98 (97)	31
PT2	23-97 (74)	65	3-50 (47)	19	0-58 (58)	17

Table 2. Survey Question Response Ratings

Questions	% Broad Agreement
Test content effectively reflect the breadth of the undergraduate curriculum	89.7%
Questions challenged my understanding	95.5%
Test 1 was effective identifying gaps in my knowledge & understanding	95%
Re-sitting the identical test was an effective way to measure personal development	86.4%
I was satisfied with the amount I learned between the 2 tests	76.1%
Negative marking discouraged me from guessing answers	85.2%
I normally guess answers in exams	55.7%
The viva encouraged me to effectively target personal knowledge development	93.2%
Explaining my answers verbally enabled me to demonstrate my understanding	83%
It was beneficial to include this type of industry assessment approach to university teaching	89.7%

Declarations

Declarations

- Ethics approval and consent to participate

Flinders University Social and Behavioural Research Ethics Committee Approval obtained. Project Number 8034. Participant consent gained prior to study. Participation was entirely voluntary and students provided verbal consent to be involved in the study. Participants electing to not be involved simply did not complete the paper based questionnaire or returned the unanswered document. The Social and Behavioural Research Ethics Committee considered this study met the “Low Risk” classification and were satisfied with voluntary consent to participate.

- Consent for publication

No identifiable individual participant information was presented. No images or recordings included. Participants consented for study findings to be included for publication. Ethics approval acknowledges the study results may be published.

- Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

- Competing interests

The authors declare that they have no competing interests

- Funding

There was no funding provided to this project.

- Authors' contributions

James Thompson was the lead author. He was responsible for study design, evaluation and production of the manuscript

Don Houston was responsible for guidance, advice, editing and broad contributions throughout the paper

- Acknowledgements

Nil

- Authors' information (optional)

Figures

Figure 1

Summary of the assessment design

Figure 2

Distribution Curves Student Responses for Progress Tests 1 & 2