

The impact of a change in assessment weighting on the academic performance of first year medical students

Damian H. Cohall and Desiree H-A. Skeete

Faculty of Medical Sciences, The University of the West Indies, Cave Hill, Barbados

The Phase I component of the Bachelor of Medicine, Bachelor of Surgery (MBBS) programme of the Faculty of Medical Sciences at The University of the West Indies (UWI), Cave Hill, Barbados recently underwent a change in its assessment policy for the preclinical courses. This change was an adjustment of the preclinical courses' assessment weighting to improve learning outcomes in students approaching the clinical training years. The Fundamentals of Disease and Treatment (FDT) course is an important introduction to the integrated approach that is used in the delivery of the system-based courses in the preclinical phase of the medical programme. The change in assessment of this course included a transition from the heavier weight in the coursework component (60%) to the final exam component, formerly 40%, and a stipulation that the students had to pass the final exam in order to successfully complete the course. This study investigated the effect of the modification of the assessment weighting on the academic performance of medical students taking the FDT course at UWI, Cave Hill during the academic year 2012/2013. The study quantitatively assessed the first year students' performances before and after the change in the course assessment during the academic years 2011/2012 and 2012/2013 respectively. In-course and final exam grades were categorised and analysed separately. T-tests were used to determine any significant differences between the means of academic performance between the cohorts over the study period. There were significant improvements in mid-semester tests, final course exams and overall course grades in the subsequent student cohort after the change. The results from the overall data analysis would suggest some relationship between assessment weighting and academic performance and that this relationship may be a result of the students' interpretation and internalisation of the rubrics that guided the assessments

Key words: assessment, weighting, coursework, academic performance

Introduction

The Fundamentals of Disease and Treatment (FDT) course at The University of the West Indies (UWI), Cave Hill is an important introduction to the integrated approach that is used in the delivery of the system-based courses in the preclinical phase of the Bachelor of Medicine, Bachelor of Surgery (MBBS) programme (Cohall, 2012). The FDT course is the largest course, in terms of teaching sessions, of the nine courses taught during the first year of the medical undergraduate curriculum.

It is taught during the first year of the medical undergraduate programme and spans both semesters. The course contributes six of the required 117 credits necessary for the attainment of the MBBS degree; and accounts for six of 29 credits available from first year courses.

The FDT course provides an early introduction to basic disease processes such as infection, inflammation, genetic disorders, tumour pathology and disorders of growth. In addition, it provides an introduction to the chemical structures and families of drugs that are used in the treatment of patients and how they work to modulate disease processes (Cohall, 2012; Cohall & Skeete, 2012; UWI, 2009). The course is delivered in a series of didactic lectures, tutorials, laboratory sessions, using an interdisciplinary approach. The members of the teaching staff are primarily from the disciplines of Pharmacology, Pathology, Physiology and Microbiology. As a result of its content, FDT is frequently considered to be a major obstacle to students during the first year of medical undergraduate training at the Cave Hill campus (Cohall & Skeete, 2012). Students tend to have a poor attitude towards the course (Cohall & Skeete, 2012) and the instructional methods, material overload and assessment procedures have been noted in student-staff liaison meetings and students' evaluations of the course as contributors to the student's poor attitude and their mediocre performances.

In addressing the need to improve student performance in fundamental preclinical courses, the Faculty of Medical Sciences at UWI, Cave Hill, recently changed its assessment weighting in preclinical courses. This change involves a transition from the heavier weight in the coursework component (formerly 60%) to the final exam component (formerly 40%) and also resulted in a final exam with a requirement of "must pass". This change in assessment policy has been well documented and formalised within the medical programme in the academic year 2012/2013. The rationale for the change in assessment policy is that it was expected to lead to an improvement in the academic performance of preclinical students which should provide them with better outcomes of their learning in relation to clinical training.

Assessment

Assessment is defined as an activity through which the evidence of learning is collected in a planned and systematic way and is used to make a judgement about learning (Harlen & Deakin Crick, 2002). Assessment is an instrumental component of learning, especially in higher education (Dunn & Mulvenon, 2009; Harlen & Deakin Crick, 2002; Knight, 2002) and is further categorised as summative and formative. Summative assessment is defined as the collection of data for the purpose of determining academic progress at the end of a specified period and with the purpose of establishing a student's academic standing relative to some established criterion. Formative assessment is the collection of data for providing feedback to all stakeholders in the teaching and learning process (Dunn & Mulvenon, 2009). Assessment is formative when the intent is to assist in decisions about how to

advance learning through the use of the judgements that indicate the next steps in closing the learning gaps. It is summative when the purpose is to summarise the learning that has occurred in order to grade, certify or record progress (Dunn & Mulvenon, 2009; Harlen & Deakin Crick, 2002). While both forms of assessment are important and have their roles in the learning process, the summative approach, despite its 'high stakes' nature, is considered to be critical to the learning process in higher education (Knight, 2002). Dunn and Mulvenon's (2009) definition of formative and summative assessment further support the view that summative assessment and its evaluation are esteemed in the higher education system.

In the 1980s, higher education began the paradigm shift from the traditional testing of knowledge and the instructive view of learning to a more student-centred approach to learning (Howell, 2011). The assessment of the new paradigm was commensurate with an understanding of the different 'styles' of learning and included the use of assessment rubrics to reflect criteria-based learning and assessment (Howell, 2011). These assessment rubrics align with the summative evaluation process and have become mainstream in many higher education programmes with extensive coursework and final examination components. In its most basic form, an assessment rubric is a matrix that shows levels of achievement for a set of criteria or dimensions of quality for a given type of performance (Howell, 2011). They are usually regarded as analytic or holistic. Analytical rubrics assess multiple aspects of performance while holistic rubrics assess the overall quality of the scholastic activity (Howell, 2011).

A rubric has three essential features: evaluation criteria, quality definitions and a scoring strategy (Reddy & Andrade, 2010). The notion of a scoring strategy is straightforward and is long-established. Evaluation criteria are the factors that an assessor considers when determining the quality of a student's work. Quality definitions provide a detailed explanation of the portrayal of the skill being assessed and are the proficiency levels to describe particular levels of achievement (Reddy & Andrade, 2010). Rubrics serve several functions such as defining standards for achievement; providing an objective basis to grade scholastic work; providing students with feedback, and making learning expectations clear to students (Howell, 2011). They support a student-centred approach by allowing students to understand the targets and quality standards of their learning as well as allowing them to inform their revision and improvements (Reddy & Andrade, 2010).

Many current course development strategies include strategies to enhance student learning. Courses are designed with specific learning outcomes that are linked to specific course content and methods of assessment, with the guidance of assessment rubrics for monitoring students' aptitude for the coursework (Howell, 2011). This multifaceted approach to course development is strongly aligned with the constructivist view of the learning process – where course development is something that is specifically designed to positively benefit academic performance (Cohall & Skeete, 2012; Patel, 2003).

A review by Reddy and Andrade (2010) relating to studies of students' responses to rubric use suggests that graduate and undergraduate students regarded

rubrics highly because they clarify the targets for their work; allow them to regulate their progress, and make grades or marks transparent and fair. The same study revealed that providing rubrics when handing out and explaining an assignment gives students relevant assignment prompts. Such assignment briefs, alongside the use of rubrics for grading or marking by the instructor, are associated with better student attitudes about fairness and satisfaction with grading. Reddy and Andrade (2010) also provided evidence of both positive responses and resistance to rubric use by college and university instructors where three studies reported positive instructor perceptions of rubrics as scoring guides and one study indicated that the students and instructors had no significant change in attitude towards the use of rubrics in the grading process.

The review by Reddy and Andrade (2010) also addressed the link between the use of rubrics and academic performance with results generally suggesting higher achievement and deeper learning by students who have rubrics to guide their work (Reddy & Andrade, 2010). There have been few quantitative research articles investigating the association between the use of rubrics and academic performance (Howell, 2011) with a meta-analysis by Jonsson and Svingby (2007) indicating that only 10 of 75 studies on grading rubrics published within the past 40 years investigated the impact of their use on student improvement. The majority of these studies did not evaluate the outcome of academic performance and the results were inconclusive (Howell, 2011).

Howell (2011) was able to demonstrate quantitatively that the use of rubrics improved academic performance and precluded a question asking what element of the use of rubrics positively impacted on student performance. The findings of Howell (2011) and Reddy and Andrade (2010) are examined in this research in relation to the FDT course at UWI, Cave Hill. In changing the assessment weighting, curriculum policymakers expected that the mean academic performance of students on the course would increase. This expected improvement in the students' academic performance emanates from the assumption that a reasonable psychological or motivational influence may be one of the chief mechanisms used by the students to overcome any perceived disadvantage of a greater weighting on the final course exam which has a "must pass" requirement compared to the previous year (Harlen & Deakin Crick, 2002). Therefore, the aim of this research project was to investigate the effect of modifying the assessment weighting on the academic performance of medical students taking the FDT course at UWI, Cave Hill during the academic year of 2012/2013. The possible factor(s) which may contribute to the positive effects of assessment rubrics on academic performance were also investigated.

Research question

To what extent did the change in the assessment weighting from 60% coursework and 40% examination to 40% coursework and 60% examination affect the students' performance in the Fundamentals of Disease and Treatment course?

Methodology

The investigation into the effect of this change follows the core concepts of action research in education which describe the integration of action or the implementing of a plan with research. Through this approach it is then expected that an understanding of the effectiveness of this implementation will be developed (Kaur, 2013).

The academic performance of two cohorts of students was examined. Cohort 1 had a weighting of 60% coursework and 40% examination and cohort 2 had a weighting of 40% coursework and 60% examination. Cohort 2 was also supplied with appropriate assessment rubrics. Throughout the period 2011-2013, grades pre and post the change in assessment weighting for the preclinical course were collected. The grades for the in-course assessments and final course exams were recorded and summarised systemically for the first year medical students during the academic years 2011/2012 and 2012/2013 respectively. The student enrolments for the two academic years are denoted as cohorts 1 and 2 respectively. In course and final exam grades were categorised and analysed separately and also combined for a comparison of the final course grades.

The data were evaluated by comparing the range of performance of the student cohorts in the respective academic years. Descriptive cohort data were assessed and recorded accordingly. The range of highest and lowest scores was used to develop the mean for each assessment component and t-tests were used to determine any significant differences in the means of academic performance between the cohorts over the study period.

Results

Seventy and 91 students were enrolled in the FDT course during the academic years 2011/2012 and 2012/2013 respectively. There was a significant improvement in the mid-semester tests, end of course exams and the overall course grades in the second cohort of students ($p < 0.05$). There was also an improvement in practical assessments but this was not significant ($p > 0.05$). Interestingly, there was a significant decline in the performance in assignments among the students in cohort 2. Figure 1 and Table 1 show the two cohorts' overall mean academic performances and their mean academic performances in the various course assessments.

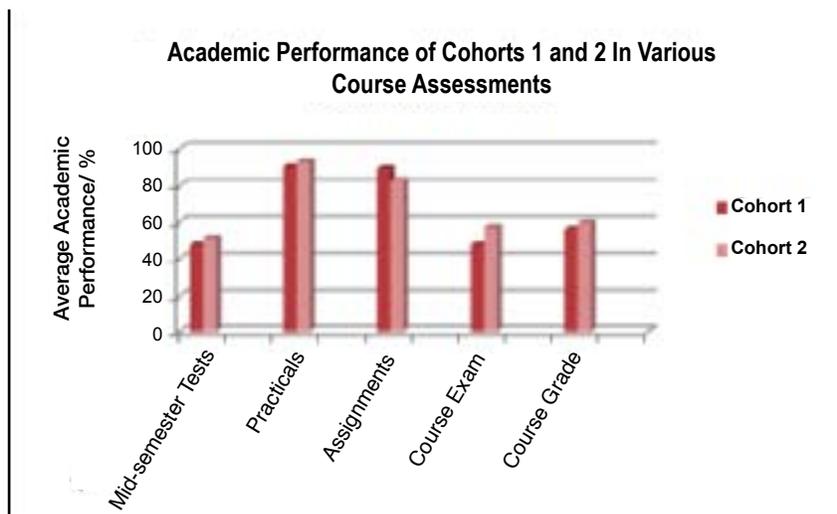


Figure 1. Mean academic performances in the various course assessments and the overall course grade

Table 1. Mean academic performances and the differences between the two cohorts

Assessments	Cohort 1 Mean with CI 95%	Cohort 2 Mean with CI 95%	Statistic	Significance (P value)
Mid-semester Tests	47.40 (44.68 – 50.12)	51.00 (48.88 – 53.12)	-2.106	0.037*
Practical Assessments	90.33 (86.64 – 94.01)	93.03 (91.50 – 94.56)	-1.483	0.140
Assignments	89.32 (84.00 – 94.64)	82.48 (78.55 – 86.41)	2.108	0.037*
Course Exam	47.45 (44.28 – 50.62)	57.04 (54.60 – 59.49)	-4.848	0.000**
Overall Course Grade	55.87 (53.23 – 58.51)	59.52 (57.56 – 61.48)	-2.259	0.025*

* Statistically significant difference in means at the 95 % confidence level.

**Statistically significant difference in means at the 99 % confidence level.

Discussion and conclusion

The change in the assessment weighting (as mentioned in the background of this study) was based on a Faculty assessment strategy. This change rendered the professional programme more appropriate to the needs of the students and was similar to assessment approaches by the other campuses of The University of the West Indies where full MBBS programmes are taught. It was also touted among faculty that this change would improve the students' learning achievement in the clinical training years. Unfortunately, the scope of this study is not able to answer that specific question. However, such a question may be explored by comparing the performance of the two cohorts upon completion of their final MBBS exams. Nevertheless, it is generally accepted that better performances in foundation courses help students as they approach their clinical training. The impact of changing the course assessment weighting resulted in significant improvements in

the mid-semester test, the course exam and the overall course grade. There was also an improvement in the practical assessment, although the t-tests did not show this to be significant (*P value* 0.140). There was a significant drop in one area where the mean performance on assignments fell.

It must be noted that students of the academic year 2011/2012 were also exposed to assessment rubrics so it could be surmised that this particular change could have contributed to the overall improvement in academic performance observed in this study. Reddy and Andrade (2010) and Howell (2007) demonstrated in their review of the literature that the use of rubrics tends to result in higher achievement and deeper learning by students as these rubrics provide students with guidelines and make learning expectations clearer. This engages the students and supports a constructivist approach to the learning and teaching process (Patel, 2003).

While the findings of Reddy and Andrade (2010) and Howell (2007) support the evidence of improvement when students are provided with course rubrics, the number of quantitative research projects investigating the association between the use of rubrics and academic performance has been found to be low in comparison to studies reporting associations of academic performance and other variables and qualitative findings were inconclusive as to the cause and effect relationship (Howell, 2007). Howell's (2007) work was also able to demonstrate quantitatively that the use of rubrics improved academic performance and essentially highlighted the need to further investigate the causal relationship.

In this study, we evaluated the overall findings based on the students' coursework and their final examination performance. The change in the course's assessment was intended to assign a higher weight to the high-stake assessment component. This change realigned the courses' objectives; the required student output, and the assigned marks for the respective components. The students in cohort 2 performed better in their practical assignments and performed significantly better in their mid-term test and final course exam. There was a fall in the mean assignment grade but this fall was not statistically significant. In all, these changes in performance meant that there was a significant increase in the mean overall course grade.

The matriculation criteria for students, teaching modalities, course work and assessments were similar in both cohorts. Assessments were standardised and consistently mapped to the learning objectives of the course in both academic years to reduce observer bias. The impact of practice effects on the study's main outcome measure was minimised by ensuring comparability of the learning experiences in both cohorts of students and ensuring consistency in the cohorts' timetables. It was possible that changing the weight of the coursework and final exam components to the extent reported in this study would negatively impact the students' performance as it deviates from emphasis on the formative to more summative approaches to learning (Udoukpong & Okon, 2012); however, this was not the case. It is therefore reasonable to infer that a psychological or motivational influence may have been one of the chief mechanisms used by the students to nullify the anticipated handicap

of a greater weighting on the final exam (Harlen & Deakin Crick, 2002), which was a “must pass” exam. It is also reasonable to conclude from these formal academic performance parameters that the students were better prepared and hence received better marks, and that this increase in preparedness may have been the result of the introduction of assessment rubrics. While it is suggested that psychosomatics may have contributed to this finding, no attitudinal or psychological scales were used in this study to empirically conclude that a motivational influence was chiefly responsible for the better academic performance of the subsequent first year cohort of students. Therefore a follow-up study with a qualitative component or a quantitative approach with attitudinal scales would be useful to determine if the attitudes changed in the second cohort of students along with their academic performance.

The second cohort of students did better in mid semester exams which might also directly affect the outcome of their performance in the final course exam and overall grade due to the formative feedback approach (Carillo-de-la-Pena et al, 2009; Udoukpong & Okon, 2012). Interestingly, the second cohort’s performance on the assignments was significantly lower than the previous year and it could be surmised that this inconsistency may have been due to smaller weighting on the assignment components and hence the students were not necessarily focused on maintaining the standards set by the cohort in 2011/2012. This finding could also be a contributor to the inconclusive evidence of the casual association of the use of rubrics and students’ academic performance as reported in the meta-analysis of Jonsson and Svingby (2007). In conclusion, the results from the overall data analysis justify the possibility of some causal relationship between assessment weighting and academic performance and that academic performance in specific areas of assessment may be a direct result of the students’ interpretation and internalisation of assessment rubrics.

References

- Carillo-de-la-Pena, M.T., Bailles, E., Caseras, X., Martinez, A., Ortet, G. & Perez, J. (2009). Formative assessment and academic achievement in pre-graduate students of health sciences. *Advances in Health Science Education, 14*, 61–67.
- Cohall, D.H. (2009). Course Outline for Fundamentals and Disease & Treatment. Bridgetown, Barbados: The University of the West Indies.
- Cohall, D.H. & Skeete, D. (2012). The impact of an attendance policy on the academic performance of first year medical students taking the Fundamentals of Disease and Treatment course. *The Caribbean Teaching Scholar, 2*(2), 115–123.
- Dunn, K.E. & Mulvenon, S.W. (2009). A critical review of research on formative assessment: The limited scientific evidence of the impact of formative assessment in education. *Practical Assessment, Research & Evaluation, 14*(7), 1-11.
- Harlen, W. & Deakin Crick, R. (2002). A systematic review of the impact of summative assessment and tests on students' motivation for learning (EPPI-Centre Review, version 1.1). *Research Evidence in Education Library*. Issue 1. London, England: EPPI-Centre, Social Science Research Unit, Institute of Education.
- Howell, R.J. (2011). Exploring the impact of grading rubrics on academic performance: Findings from a quasi-experimental, pre-post evaluation. *Journal on Excellence in College Teaching, 22* (2), 31-49.
- Jonsson, A. & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*, 130–44.
- Kaur, S. (2013). Action research in education. *International Journal of Research in Education Methodology, 3*(2), 255-258.
- Knight, P.T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-286.
- Patel, N.V. (2003). A holistic approach to learning and teaching interaction: Factors in the development of critical learners. *The International Journal of Educational Management, 17*, 272-284.
- Reddy, Y.M. & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435-448.
- The University of the West Indies (UWI) (2009). Medical Students' Handbook. Bridgetown, Barbados: The University of the West Indies.
- Udoukpong, B.E. & Okon, C. P. (2012). Perception of formative evaluation practices and students' academic performance in junior secondary certificate examination in social studies. *International Journal of Business and Social Science, 3*(15), 204-212.