Measuring the quality of the assessment process: dealing with grading inconsistency

Practitioner Research
In Higher Education

Vol 8 (1) pages 32-40

Dawid Wosik
Higher Colleges of Technology, Fujairah Women's College, UAE.
dwosik@hct.ac.ae

Abstract
This paper examines an assessment process within the context of quality, accountability and institutional effectiveness in academia. Considered as one of the most important elements in higher education, the assessment process plays a significant role in contributing to holistic quality as well as the performance of an academic institution. As a result, appropriate requirements need to be established, applied consistently and communicated effectively among different groups of stakeholders. The lack of a systematic approach in this matter as well as the lack of specific performance measures of the assessment process shape a false image of students', and consequently, graduates' academic capacity.

Keywords
Quality; institutional performance; grading consistency; grade inflation; quality measures.

Introduction

Quality in higher education can be defined as a multi-dimensional and multi-layered concept which depends on the requirements set by different stakeholders (Vlãsceanu et al., 2007). Likewise, quality of the assessment process will be described differently depending on different stakeholders' requirements (Bloxham, 2008). Regardless of the perspective taken, the importance of the assessment process and the need for its consistency is undeniable. 'Important decisions are based on information derived from classroom assessments; it is imperative that the information be of high quality: accurate, dependable, meaningful, and appropriate' (Brookhart, 1999). These decisions often affect students' motivation and their academic choices. It is obvious that grades should reward the best students for their hard work and encourage the mediocre and low performing ones to put more effort in their learning (Johnson, 2003). Inconsistency in assessing students distorts the signaling and reward function of the entire assessment process and is misleading. Lack of objective differentiation between students' performance also affects potential employers and graduate schools. It is simply not possible to distinguish between students 'who received good grades because they are bright and worked hard, or because they sought out faculty members who give mainly A's' (Pressman, 2007).

Therefore the assessment process should produce scores which differentiate between higher and lower performing students. Otherwise the grades effectively measure nothing. It needs to be under institutional control to minimise variation of the process and at the same time provide evidence of its reliability as well as uniformity. It is of the highest importance as 'looking at variation helps management to much fully understand the real performance of a business and its processes' (Pande et al., 2000).

This paper outlines a practical approach to measuring the quality of the assessment process in order to support decision-making at the institutional level. It argues for the importance of performance

Citation
Wosik. D. (2014) Measuring the quality of the assessment process: dealing with grading inconsistency, *Practitioner Research in Higher Education Journal*, 8(1), January, 32-40.

indicators that address the issue of grading inconsistency. The article gives an overview of research relating to the subject matter and translates it into the context of the Higher Colleges of Technology (HCT) in the United Arab Emirates. Grade distribution, common examinations, differences between course work and final exam results are presented in the paper to offer practical ways of analysing the quality of grading with special focus on grading consistency. The paper is a result of the author's professional practice and research in the field of quality management in higher education. The article's distinctive contribution relates to conceptualising the approach to measuring the quality of the assessment process in a specific educational setting.

Established in 1988, HCT consists of 17 campuses with an enrollment of approximately 20,000 students – it is the largest federal institution of higher education in the UAE. There are about 2,000 faculty and staff from more than 60 countries working at HCT. All programs are recognised internationally as they are either accredited or validated by various institutions from all over the world[1]. In this paper the author distinguishes between quality assurance (QA) processes developed by HCT as a system and those specific to individual campuses (e.g. Dubai Women's College, Fujairah Men's or Fujairah Women's Colleges). HCT students are taught in English (their second language) according to centrally-developed and common course outlines that apply to all 17 campuses in the HCT system. This creates a particular setting for establishing various quality measures of the assessment process.

Grading consistency as a measure of quality

There are so many definitions of quality that its meaning is often confused. Vroeijenstijn (1991) says: 'it's a waste of time to try to define quality'. It is like love: 'everybody knows and feels when there is love. Everybody recognises it. But when we try to give a definition of it, we are left standing with empty hands' (Vroeijenstijn, 1995). Quality is an elusive concept and we all think we know what it means, however it is very difficult to define (Burrows et al., 1992).

The point is that quality in general doesn't mean anything – it is just a measure. It means something only if it is put into a context – a particular setting where it can be described by various attributes. Quality is a 'degree to which a set of inherent characteristics fulfills requirements' – the degree to which a product or service is fit for purpose (ISO 9000:2005, Burrows et al., 1992). In order to constitute quality, it is important to determine what those characteristics and requirements are. These will always vary depending on the perspective of the one describing it. In terms of quality in higher education as well as the assessment process, there are various stakeholders who see quality differently, e.g. faculty, students, employers (Bloxham, 2008; Burrows et al., 1992).

Quality is a result – the result of being aware (Wosik, 2009). The complexity of its description will depend on the state of awareness of the requirements and expectations shaped by various interested parties in a particular context. Once we know what the requirements, and therefore possible measures of quality are, it is much easier to manage performance. Further, there is no quality management without quality measurement. All processes within an organisation have to be measured in order to determine the degree to which the requirements are met.

The assessment process aims at identifying and discriminating between different levels of student

[1] For example, IT programs are accredited by C.I.P.S. – Canada's Association of IT Professionals, Engineering programs are accredited by ABET – Accreditation Board for Engineering and Technology, Business programs are accredited by ACBSP – Accreditation Council for Business Schools and Programs and Education programs are validated by the University of Melbourne in Australia.

achievement. The quality, meaning consistency, of the assessment process has not only a profound effect on students' future careers, as it enables selection for further study and employment, but it also provides evidence of institutional accountability (Bloxham, 2008). Students should be assessed based on clear requirements, procedures and criteria for marking and, where possible, the assessment should not rely only on the judgment of a single examiner (ENQA, 2009). To assure that assessments are reliable, fair and fit for purpose, the assessment process should be periodically reviewed (QAA, 2006). Any discrepancy (e.g. grade inflation) may have an impact on the overall quality of education. It is empirically proven that learning is negatively correlated with lenient grading (Johnson, 2003). Therefore adherence to the assessment requirements should minimise variation of the assessment process, and as a result, improve grading consistency – the consistency between the student's real accomplishment and grade given.

Grade inflation as a measure of grading inconsistency

Grade inflation can be seen as an indicator of quality relating to grading consistency. It may be a result of either the lack of the requirements relating to the assessment process or the lack of awareness of what the requirements are.

In his research, Millet (2010) finds grade inflation as one of the major causes of grading inconsistency and uses Grade Lift to measure it. Grade Lift is the difference between average class grade for a particular course (course GPA) and average cumulative class GPA for all courses the same students take within a semester (Millet, 2010). Positive Grade Lift means that grades assigned by an instructor for a particular course are on average higher than peers' for the same group of students, indicating possible grade inflation. Negative Grade Lift means exactly the opposite. Generally speaking, grade inflation issues exist when the same GPA is obtained over time by students with poorer academic skills than previous cohorts with higher skills (Schiming, 2009). It can also be understood as content deflation where students receive the same grades as students in the past but with less work required and less learning (Cohen, 1984). According to Pressman (2007), it is very similar to price inflation where more money is spent to buy the same product. This corresponds with the situation where higher grades are given for the same achievement. Nevertheless, there is a significant difference between price and grade inflation. In case of price inflation, the price can still be used as an indicator of quality – the higher the price, the better the quality. As far as grade inflation is concerned, unfortunately it is not possible to distinguish between the quality of students' achievements because all of them have the same – high 'price'.

There are many perspectives to take into account while analysing possible effects of grade inflation. One of them is the student's perspective: the extent to which grade inflation has an impact on the student's engagement in learning. There is no doubt that the engagement of a student in the learning process will depend on how the grades reflect his actual academic performance (Asante et al., 2012). Will the student engage further in the learning process if he is already awarded an 'A' for putting in less effort than required for the highest grade? What about the best students in the class? What 'morale impact' will it have if they receive the same grades as those who simply don't deserve it? The sense of being challenged is lost – the student is not distinguished enough in the labor market, because high grades could be either for better performance or due to general problems with grade inflation. As a result, such a situation impacts prospective employers and graduate schools (Pressman, 2007).

HCT's requirements relating to the assessment process

Having understood that 'quality is as a result – the result of being aware', spreading awareness of the

requirements is the first and necessary step to minimize variation of processes and at the same time to identify more clearly the parameters for quality improvement. There is a constant need to find such approaches and tools which enable an institution to strike a balance between accountability for quality (awareness of and meeting the requirements) and quality improvement (Wosik, 2009, Stensaker, 2003).

HCT has used a variety of requirements and tools to support performance management of the assessment process at the institutional level. The most important document in this matter is the HCT Assessment Policy. It aims at ensuring 'that assessment is conducted fairly, equitably and consistently across the system, in line with the principles of the HCT Learning Model, in order to measure and provide feedback to stakeholders on student achievement of learning outcomes, thereby informing strategies for continual improvement of teaching and learning at the HCT' (HCT Assessment Policy, 2013; HCT Learning Model, 2006). A number of documents related to the HCT Assessment Policy specify requirements relating to such areas as:

- Responsibilities for Assessment;
- Course-Level Assessment;
- Program Level Assessment;
- Make-up and Supplemental Assessments;
- Challenge Assessments;
- Moderation of Assessments;
- Assessment Methods and Accommodation.

Another important set of criteria in terms of managing consistency of the assessment process is the HCT grading system (Figure 1.). The grading system is standardised and used by faculty across all HCT campuses and programs.

Grade	Range	Grade Point	Descriptor
A	90 – 100	4	Achievement that is outstanding relative to the course and GPA requirements.
A-	85 – 89	3.7	
B+	80 – 84	3.3	Achievement that is significantly above the course and GPA requirements.
B	75 – 79	3	
C+	70 – 74	2.3	Achievement that satisfactorily meets the course and GPA requirements.
C	65 – 69	2	
D	60 – 64	1	Achievement that minimally meets the course requirements but may not meet the GPA requirements.
F	0 – 59	0	Achievement that does not meet requirements.

Figure 1. HCT grading system (Source: HCT Grading System Policy, 2012).

Grades' descriptions help to distinguish between students' levels of achievement. If a normal curve was applied to the existing grading system, most of grades would range between B's and C's (Figure 2.). Although using the normally distributed 'bell' curve to set the expectations regarding an estimated distribution of grades might be helpful to strike a balance between different qualities of students' performance, it is not a mandatory standard to meet in the HCT system.

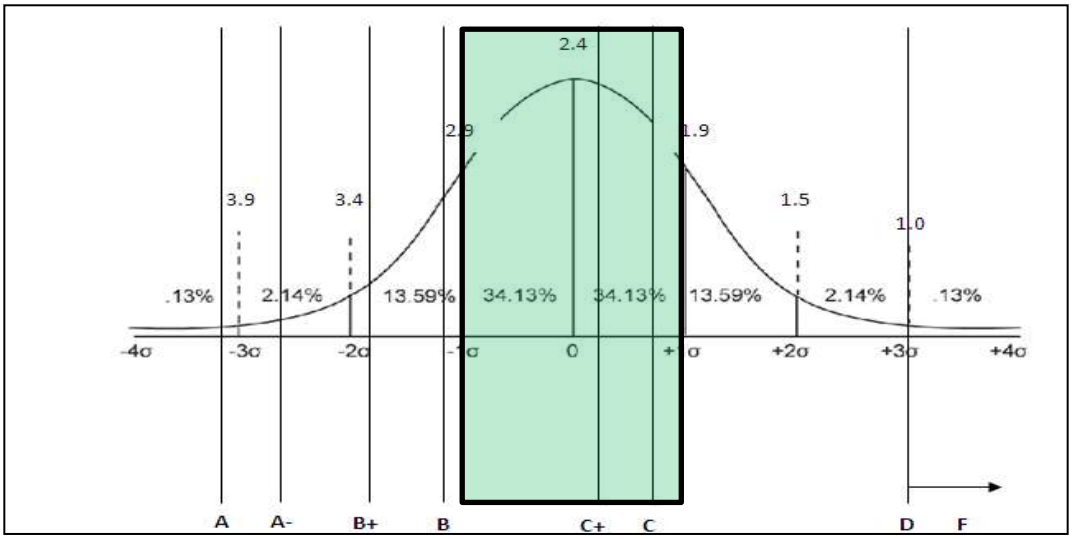

Figure 2. HCT grading system in the light of normal distribution (Source: Wosik: own elaboration).

Grade distribution and common examinations have been used at HCT to measure quality of the assessment process as well as to support an ongoing discussion and analysis of possible 'assignable' causes of its variation.

Grade distribution

Grade distribution is accessible throughout the academic year across the system. All faculty and academic chairs have access to such data allowing them to take essential corrective actions, if needed. Below is an example of a specific – structured report used at Fujairah Men's and Fujairah Women's Colleges as part of a weekly QA scheme (Figure 3.). It is important to highlight that such an approach is specific to Fujairah Colleges' QA processes and it is not a system-wide requirement.

FWC	As of week 14 (2013-06-02)				
Row Labels	A	B	C	D	F
⊞AB	24.49%	40.65%	24.91%	5.06%	4.90%
⊞BP	44.13%	48.60%	5.59%	0.56%	1.12%
⊞Diploma Year 2	3.57%	50.00%	46.43%	0.00%	0.00%
⊟HD/BAS Year 3	34.69%	50.94%	13.13%	0.63%	0.63%
⊞10S5BUHR01	24.53%	56.60%	16.98%	0.94%	0.94%
⊞10S6APMS01	82.14%	17.86%	0.00%	0.00%	0.00%
⊟10S6BUAC01	12.82%	64.10%	20.51%	0.00%	2.56%
BMAC-N350	5.26%	47.37%	42.11%	0.00%	5.26%
BMGN-N350	21.05%	78.95%	0.00%	0.00%	0.00%
ECON-N260	0.00%	100.00%	0.00%	0.00%	0.00%
⊟10S6BUAD01	48.72%	38.46%	12.82%	0.00%	0.00%
BMGN-N350	0.00%	72.22%	27.78%	0.00%	0.00%
BMHR-N360	100.00%	0.00%	0.00%	0.00%	0.00%
ECON-N260	33.33%	66.67%	0.00%	0.00%	0.00%
⊟10S6BUHR01	22.03%	64.41%	13.56%	0.00%	0.00%
BMAC-N250	50.00%	50.00%	0.00%	0.00%	0.00%
BMGN-N350	0.00%	69.23%	30.77%	0.00%	0.00%
BMHR-N310	42.31%	57.69%	0.00%	0.00%	0.00%
Grand Total	27.72%	37.58%	22.21%	5.92%	6.57%

Figure 3. Coursework grade distribution: weekly snapshot of HCT-Fujairah Women's College data (Source: Wosik: own elaboration).

Weekly reports on grade distribution give one perspective of looking at grading consistency by analysing a relative number of A's, B's, C's etc. The analysis is broken down by program, section and a particular course, hence enabling comparisons and stimulating discussion between faculty and academic chairs. For example, in the above snapshot the BMGN-N350 course is taken by three different sections of students – 10S6BUAC01, 10S6BUAD01 and 10S6BUHR01. In the first section, 21% of students received an A and 79% received a B. In the second section there were no As awarded and 72% of students received a B while 28% of them were given a C. A similar situation took place in the last section where 69% and 31% of students received B's and C's respectively.

Moreover, an average coursework grade of a chosen subject can be analysed from a system perspective. A report showing *as-of-now* individual courses' average grades in light of the system-wide data is accessible by both faculty and chairs. It stimulates not only informative discussions with peers from the same college but also contributes to system-wide improvements (Figure 4.). For example, a faculty member from Fujairah campus can refer to system-wide data and compare an average grade of his students with an average grade of all the students taking the same course across all campuses.

Course	Course Name	CRN	SWA Course	Faculty ID Students	CW Entry % Completed	CW Entry % Completed (System)	Men's College CW Average	Women's College CW Average	Average Course.. 0 50 100
BMAC-N250	Financial and..	46980	No	2	100.00	39.50	Null	72.17	84
BMAC-N350	Management..	45439	No	19	25.00	53.80	79.49	83.95	75
BMGN-N310	Commercial ..	45390	No	17	40.00	39.00	66.79	75.38	77
BMGN-N320	LeaderShip a..	45393	No	17	20.00	52.80	71.82	76.73	89
BMGN-N350	Business Tactics	45432	Yes	26	28.60	43.00	75.45	75.88	78
		45436	Yes	18	28.60	43.00	75.45	75.88	75
		45437	Yes	19	14.30	43.00	75.45	75.88	79
BMHR-N310	Global Cop..	45433	No	26	50.00	49.80	78.71	81.37	84
BMHR-N320	Managing Or..	45396	No	17	100.00	97.50	76.41	76.30	81
BMHR-N360	Training and Development	45395	No	17	73.50	55.90	80.09	80.30	79
		45438	No	18	25.00	55.90	80.09	80.30	95
BMHR-N400	HR Planning and Recruitm..	45448	No	23	25.00	53.70	Null	81.17	80
		45971	No	4	50.00	53.70	Null	81.17	86

Figure 4. Grade distribution: weekly snapshot of HCT-Fujairah Women's College coursework grades compared to system-wide data (Source: Wosik: own elaboration based on HCT's BI Analytics report).

The data shows that two courses: BMGN-N350 and BMHR-N360 are taught by three and two faculty respectively. It presents an average coursework (CW) grade for each course along with the average value for all men's and women's colleges in the HCT system. It also gives some additional information about the percentage of grades completed for an individual course/faculty and the system. The report indicates whether or not a course has a common examination (SWA Course indicator).

Common examinations

HCT common requirements with regard to academic programs provide an opportunity to administer common examinations across all divisions. What it means is that a chosen course taken by students at different colleges can have the same final exam, aiming at assuring consistency of teaching, learning and assessing across the entire system. Every semester there are courses identified to have a system-wide assessment (SWA). This is 'a centrally-managed, comprehensive assessment of student achievement of intended learning outcomes, taken by all students registered on the same course across the HCT system' (HCT Assessment Policy, 2013). The purpose of a SWA is to:

- 'Audit the consistency of academic standards, student learning, and curriculum delivery across colleges offering the same course;
- Measure student achievement of intended learning outcomes so as to identify areas for the improvement of student learning' (HCT Assessment Policy, 2012).

HCT uses SWAs in key performance indicators (KPI's) related to the assessment process. These KPIs are reported and analyzed on a semester basis:

- 'CW – SWA' – the difference between coursework grade and the final exam grade;
- 'Non SWA Course Grade Point Average (GPA) – SWA Course GPA' – the difference between an average GPA of courses without and with a common examination.

Spreading awareness of the requirements – the process

The most important reason for establishing the preceding practices is to provide faculty and management with data regarding the distribution of grades and therefore the quality of the assessment process. Spreading awareness of different sets of data aims at initiating discussion on possible improvements to the assessments to ensure better validity, reliability, consistency and accuracy of the process. It may result in amending assessment strategies, revising mechanisms for validating assessment instruments and moderating the marks awarded.

At Fujairah campuses new faculty are introduced to all requirements relating to the assessment process during an orientation session with the Institutional Research and Quality Improvement Co-ordinator. The session aims at spreading awareness of the HCT Grading System, grade distribution data as well as the SWA process. The Academic Chair of each department is expected to organise a meeting within the first five weeks of a semester which would focus on grading consistency. In addition, faculty are encouraged to discuss grading issues during their team meetings at least three times throughout a semester.

Weekly grades distribution reports highlight all courses where grade distributions deviate from the program average. At the end of each semester a SWA report identifies courses where differences between CW and SWA are greater than 10%. The report is sent to the Associate Director and Academic Chairs and requires a campus response and a proposed way forward.

Conclusion

Quality standards can't be conceptualised as having an existence or relevance separately from the context (Sadler, 2005). Depending on the circumstances an institution operates in, different measures will be used in order to determine various aspects of quality.

The HCT system consisting of 17 colleges provides an ideal setting to administer common examinations as one of the ways to monitor quality of the assessment process. The use of common assessments enables the introduction of various performance indicators aimed at signaling deviations in grading, and hence initiating quality improvement.

Reporting the distribution of grades system-wide stimulates discussion, and often leads to campus-specific methods of dealing with grading inconsistency (e.g. offering professional development sessions on the assessment process, moderating assessments or organising department meetings focused on grading practices). Reporting on grade distribution has the advantage that it can be applied even when common assessments are not feasible. Trying different approaches, disseminating data and spreading awareness of requirements helps to understand and define quality from many different perspectives. Consequently, quality awareness becomes essential to establish performance measures to effectively address the grading inconsistency issues.

References

Asante, C., Al-Mahrooqi, R., Abrar-ul-Hassan, S. (2012) The Effects of Three Teachers Variables on the Use of Motivational Strategies in EFL Instruction in Oman, *TESOL Arabia Perspectives,* 19(1), 12-22.

Bloxham, S. (2008) Assessment in teacher education: stakeholder conflict and its resolution, *Practitioner Research in Higher Education,* 2(1), 13-21.

Brookhart, S. (1999). *The art and science of classroom assessment: The missing part of pedagogy*. Washington. DC: ERIC Clearinghouse on Higher Education.

Burrows, A., Harvey, L., Green, D. (1992) *Concepts of Quality in Higher Education: A review of the literature*. Birmingham: Birmingham Polytechnic. Available at: http://www.qualityresearchinternational.com/Harvey%20papers/Burrows%20Harvey%20and%20Green%20Concepts%20of%20Quality%20in%20Higher%20Education%20A%20review%20of%20the%20literature%20.pdf (Accessed: September 2013).

Cohen, H. (1984) Inflated Grades, Deflated Courses: How Insecurity Induced Compromise, *Change: The Magazine of Higher Learning,* 16(4), 8-10.

European Association for Quality Assurance in Higher Education (ENQA) (2009) *Standards and Guidelines for Quality Assurance in the European Higher Education Area.* 3rd edn. Helsinki: European Association for Quality Assurance in Higher Education (ENQA).

Higher Colleges of Technology (2006) *HCT Learning Model*. Available at: http://www.hct.ac.ae/about/learning-model/ (Accessed: September 2013).

Higher Colleges of Technology (2012) *HCT Grading System Policy.* Available at: http://portal.hct.ac.ae/sites/pnp/cass/Pages/lp209.aspx (Accessed 07 January 2014).

Higher Colleges of Technology (2013) *HCT Assessment Policy.* Available at: http://portal.hct.ac.ae/sites/pnp/cass/Pages/lp220.aspx (Accessed: 07 January 2014).

International Organization for Standardization (ISO) (2005) *ISO 9000 Quality management systems – Fundamentals and vocabulary.* Switzerland: International Organization for Standardization (ISO).

Johnson, V. (2003) *Grade Inflation: A Crisis in College Education*. New York: Springer-Verlag.

Millet I. (2010) Improving Grading Consistency through Grade Lift Reporting, *Practical Assessment Research & Evaluation,* 15(4), 1-8.

Pande, P.S., Neuman, R.P., Cavanagh, R.R. (2000) *The Six Sigma way: How GE, Motorola, and other top companies are honing their performance.* New York: McGraw Hill.

Pressman, S. (2007) The Economics of Grade Inflation, *Challenge,* 50(5), 93-102.

Sadler D. R. (2005) Interpretations of criteria-based assessment and grading in higher education, *Assessment & Evaluation in Higher Education,* 30(2), 175–194.

Schiming, R.C. (2009) *Grade Inflation*. Available at: http://www.mnsu.edu/cet1/teachingresources/articles/gradeinflation.html (Accessed: September 2013)

Stensaker, B. (2003) Trance, transparency, transformation: the impact of external quality monitoring in higher education, *Quality in Higher Education,* 9(2), 151-159.

The Quality Assurance Agency for Higher Education (QAA) (2006) *Code of practice for the assurance of academic quality and standards in higher education: Section 6: Assessment of students*. Mansfield: The Quality Assurance Agency for Higher Education (QAA).

Vlăsceanu, L., Grünberg, L., and Pârlea, D. (2007) *Quality Assurance and Accreditation: A Glossary of Basic Terms and Definitions.* Bucharest: UNESCO-CEPES.

Vroeijenstijn, A.I. (1995). *Improvement and Accountability: Navigating Between Scylla and Charybdis Guide for External Quality Assessment in Higher Education*. London: Jessica Kingsley Publishers.

Vroeijenstijn, T. (1991). External Quality Assessment: servant of two masters? the Netherland University Perspective, in Craft, A. *Quality Assurance in Higher Education in Hong Kong Proceedings of an International Conference, Hong Kong, 1991*. London: Falmer Press, 109-132.

Wosik D. (2009) *Towards Excellence in Higher Education through Quality Awareness*. Available at: http://www.inqaahe.org/admin/files/assets/subsites/1/documenten/1241709204_26-wosik-towards-excellence-in-he-trough-quality-awareness.pdf (Accessed: September 2013).