7th Research/Expert Conference with International Participations
"QUALITY 2011", Neum, B&H, June 01 – 04, 2011

CONTEMPORARY TRENDS IN SETTING STANDARDS FOR EDUCATIONAL ASSESSMENT; APPLICATION TO STATISTICAL EDUCATION

Velid Draganovic
International University of Sarajevo
Sarajevo
Bosnia and Herzegovina

SUMMARY
This paper deals with issues of developing standards in educational assessment. The central issue is outcome-based assessment, which is discussed from the viewpoint of paradigm shift from traditional methods based on standards of reliability and validity to newly developed outcome based assessment methods. Finally, discussion is applied to the field of teaching statistics and applicability of this paradigm shift to teaching this discipline including real world problems approach and individualized assessment.

Keywords: Assessment, learning outcomes, standards, statistics

1. INTRODUCTION

The notion of quality started in the productive industry in the middle of previous century, but since few decades it is taking roots as well in service based activities like health care and education. This paper will address the issue of assessment, as one of the important aspects of quality in higher education. Here, we are trying to go into details, choosing assessment as a focus of discussion, and going into issue of developing methods of assessment, which are going to serve the purpose of improving quality of education. For fulfilling the task, we will start with defining assessment, continue with analyses of the issue of outcomes, shift of assessment methods, and pointing out most important methods of assessment in statistical education.

2. ASSESSMENT AND ITS IMPORTANCE

European Higher Education Area (continuation of the Bologna process) is promoting standard and guidelines developed by ENQA - the European Association for Quality Assurance in Higher Education. In their publication ENQA [6] developed the quality assurance standards at various levels from internal to external quality assurance. An assessment of students is covered by one standard which is defined as: "Students should be assessed using published criteria, regulations and procedures which are applied consistently." [6] This standard is clarified through guidelines, out of which we have taken two into detailed analysis in this paper: one says that student assessment procedures are expected to be designed to measure the achievement of the intended learning outcomes and other programme objectives; the other says that procedures should be appropriate for their purpose, whether diagnostic, formative or summative. The first mentioned guideline makes direct connection between assessment and learning outcomes, and the second one indicates different assessment procedures, and

indirectly in some way assessment methods. Discussion about importance of outcomes would follow, but appropriate definitions of concept and form of assessment are milestones of our discussion; therefore we start by defining them.

Assessment is the systematic collection of information about student learning, using the time, knowledge, expertise, and resources available, in order to inform decisions that affect student learning. [7] This definition could be applied to the institution, program or course. The scope of this paper will limit discussion to the classroom assessment.

The chosen definition defines what assessment is and at the same time indicates its purpose. Actually, definitions points to main difference between grade and assessment. A grade in class is limited in scope and does not provide much information about how to adjust course design and increase students learning. While major adjustments are subject to approval at the department (or programme) level and usually are made outside the classroom, same minor adjustments are possible within the classroom itself. They are based on continuous assessment, while concepts are built sequentially throughout the course.

An insight into the classification of the forms of classroom assessment can give a glimpse into possibilities of adjustment within the duration of the course. Three most common terms regarding the classroom assessment are diagnostic, formative, and summative assessment. Diagnostic assessment (sometimes called pre-assessment) takes place before starting a proposed learning activity, and its main purpose is to provide class instructor with the insight into student's prior knowledge and ability to follow that activity. It can be a very short assessment focused on the key concepts, or it can be the summative assessment of the previous learning activity.

Formative assessments are performed during a learning activity. They are specifically convenient for observing process of producing term papers or projects, case studies or similar activities, and beneficial for students who have not done well in high school. They can provide evidence for evaluating or grading students, but their chief purpose is to improve the quality of student learning.

Finally, there is an assessment performed after learning activity has been completed. It is called summative assessment because the information is intended to sum up what has been learned to that point in time, and its' primary purpose is to assign grades to students. However, it can serve as diagnostic assessment in the cases when summative assessment is made for specific learning activity, and gives the direction of further steps. For example, results of a summative assessment indicate that majority of the students did not learn key concepts they need to proceed to presumably successful performance of coming learning activity or to start of their term project. In such a case, spending some time to improve acquiring those concepts, is a wise decision made on basis of summative assessment, which actually plays role of diagnostic assessment for next learning activity, or formative assessment when the whole course is concerned.

In order to further illustrate the meaning of assessment we will point to three important differences between classroom assessment and traditional or normal evaluation of student learning. These differences concern purpose, unit of analysis, and the criteria on which the analysis is based. First, classroom assessment is motivated by teacher's goal to learn more about student learning instead of merely providing institution the simple information about achievement (grade). Second, the focus of collecting data is the whole class instead of individual student. Third, evaluation is performed in a way which would provide in depth inquiry into learning process and improvement possibilities instead of mere classifications of students. [2]

Assessment is commonly perceived as an important tool of educational reform from traditional education to outcome based education. Assessment movement is a notion that higher education should examine what students have learned, not just what the institution or

department did that supposedly resulted in learning. [7]. The importance of assessment is stressed by accreditors and state legislatures. Demand that institutions of higher learning should examine what students have learned is actually indicating how assessment got its importance. It is related to dissatisfaction with the perceived level of knowledge of college and university graduates by society. Hence, the look to education process changed its direction. Starting point was now student, graduate profile, and learning instead of institution, study program, and teaching. Learning outcomes became the main goal and criteria setter for development of assessment methods. It is the quality of giving information for improvement of learning methods and learning outcomes that made assessment methods a prime target of numerous studies. On the other hand, it makes assessment and outcomes interrelated, so that it is quite natural to address the issue of learning outcomes immediately after discussion about assessment.

3. LEARNING OUTCOMES AS AN AGENT OF EDUCATIONAL REFORM

Learning outcomes are described as written statements of what a learner is expected to know, understand and/or be able to do at the end of a period of learning. [1] Therefore, outcomes are stated expectations of what someone will learn, or what student would be able to demonstrate, represent or produce when they graduate. They are key component of outcome based education, which is gradually replacing traditional education system.

The shift of primary concern from teaching to learning and from inputs to outcomes started in USA in 80s and in EU in the beginning of the 90s. Nowadays, student-centered learning and learning outcomes will be at the core of 'implementing Bologna' following support gained by Berlin Communiqué, the adoption of the European Standards and Guidelines (ESG), and outcomes-focused Qualifications Framework of the European Higher Education Area (QFEHEA) (Bergen Communiqué).

The importance of learning outcomes is mainly evident from couple of indicators: they make qualifications more transparent for students, then, the range of graduates is becoming wider, potential employers may have a better understanding of the acquired knowledge, skills and competences of job candidates, and learning outcomes are also valuable in terms of course design, providing benefits for course instructors and educational institutions. All those reasons influenced shift from traditional education system to outcome based education.

The practical meaning of this shift is expressed in diversity and complexity of stated outcomes. In traditional classroom evaluation, the criteria for assessing student work are intimately linked to grading. Such criteria are usually very simple, sometimes binary like right or wrong, acceptable or not, pass or fail, sometimes numerical values from 1 to 100, sometimes letter grades from A to F. As such, they can achieve the main objective of ranking students or rating them, and they can be universally used in various disciplines. On the other hand, various disciplines can produce various outcomes making their number so huge, that it is even possible to consider virtual universe of learning outcomes, which could be classified into four dimensions: knowledge outcomes, skills outcomes, attitudes and values outcomes, and behavioural outcomes. [4]

4. ASSESMENT STANDARDS

Learning outcomes, assessment methods, and standards are closely connected, so that paradigm shift to outcome based education necessarily affects standards of assessment. Further, it makes no sense of discussing standards without reference to criteria. In an attempt to define meaning of standards and criteria it must be clarified that world standard is not going to be used in confined scope of standardized test and exams. The term standard would rather refer to the different levels of fulfilling criteria as a quality that is expected to be revealed by students in their demonstration of desired outcomes. Criteria can even be

considered as a description of outcome, describing skills, knowledge, kind of thinking, comprehension, attitudes of students. Here again, like with outcomes we are confronted with the diversity of criteria. In order to provide clear picture by confining number of criteria, a good idea is to demonstrate analysis of criteria and standards on one scientific or artistic discipline. For this purpose scope of the paper is limited to application of those definitions to statistical education only. Hence, the illustration of relationship between outcome and criteria we can achieve by looking into examples. For instance, instructor of statistics course defined that one of the intended learning outcomes in statistics course is students' ability to define variables and choose scales of measurement. The assessment criteria for this learning outcome would be appropriateness and analytical thinking.

The question that appears now is: "how to measure analytical thinking or appropriateness"? The answer to this question is related to the shift from single evaluation approach to more complex, or in other words from measuring on numerical scale to rather descriptive or qualitative measurement. The assessment standards would take role of describing different levels of criteria. If we refer back to our example, we need standards to define and describe precise meaning of the mentioned criteria. For both of them standards would describe criteria for three different levels of achievement: excellent, satisfactory, and unsatisfactory.

Standards for criteria of analytical thinking could be summarized as follows: Excellent: The student consistently analyses context in which variable is operational; consistently reviews relationships between different variables; consistently considers possibilities of modification of operational definition of variable; Satisfactory: The student analyses context in which variable is operational in most of the cases; reviews relationships between different variables; considers possibilities of modification of operational definition of variable without major difficulties; Unsatisfactory: The student occasionally analyses context in which variable is operational; occasionally reviews relationships between different variables; occasionally considers possibilities of modification of operational definition of variable expressing difficulties in that process.

Standards for criteria of appropriateness could be summarized as follows: Excellent: The student accurately and consistently chooses relevant variables; accurately and consistently recognizes scale of measurement; Satisfactory: The student accurately chooses relevant variables in most of the cases; accurately recognizes scale of measurement in most of the cases; Unsatisfactory: The student occasionally accurately chooses relevant variables; occasionally accurately recognizes scale of measurement.

Presented examples illustrate how helpful criteria and standards can be to students as they learn and demonstrate their learning. On the other hand they illustrate trends in shift towards assessment culture, which can be listed in seven aspects of change. [5] The first aspect is a change from decontextualized, atomic tests to authentic contextualized tests. In other words, it refers to the shift from the so called objective tests with item formats such as short answer, fill-in blank, multiple-choice and true/false to the use of portfolio assessment, project based assessment, performance assessment, etc. The second aspect is a tendency from describing a student's competence with one single measure (a mark) towards portraying a student's competence by a student's profile based on multiple measures. The third aspect is expressed as change from low levels of competence towards higher levels of competence. The fourth aspect refers to the multidimensionality of intelligence acknowledging that intelligence in not IQ only which refers to cognition; it implies meta-cognition, emotional and social intelligence as well. The fifth aspect is a movement of integrating assessment into the learning process. The sixth aspect points to the change in responsibilities, not only in the learning process but also in the assessment process. The direction of this change is from teacher to student. The seventh very important aspect refers to the shift from the assessment of learning towards an equilibrium between assessment of learning and assessment for learning, where expression

assessment for learning actually means using results of assessment to improve learning and its outcomes.

Implications of this paradigm shift for statistical education could be observed starting from learning outcomes. They may be defined by course instructor, but usually are related to the learning outcome of the study programme or to the value defined by institution of higher learning. The attitudes and abilities developed in statistic courses are contributing to the profile of the graduate. However, truly perceived contribution could be attained only when skills are connected to the real world and the field of expertise. In order to achieve such an important goal statistics education is shifting from traditional approaches to teaching statistics focused on procedures and computations to modern education with focus on statistical literacy, reasoning and thinking. The shift in desired outcomes caused the shift from testing culture to assessment culture as already pointed. Testing culture refers to extensive reliance on edumetric indicators like reliability and validity. The validity question refers to the extent to which assessment measures what it is intended to measure, while reliability was traditionally defined as the extent to which a test measures consistently. On the other hand, consistency in test results was proving objectivity in scoring, because the same results were obtained if the test was judged by another person or by the same person at another time. The consequence of striving for objectivity as a main indicator of fairness in testing is the extensive use of standardized testing forms, like multiple-choice tests. Ongoing shift to assessment culture implies an adjustment in existing assessment methods, and innovating new methods and techniques of assessment, like Applications Cards, Minute Paper, .Problem Recognition Tasks and Productive Study-Time Logs [2]. Finally, by taking in account individual differences among students and goal of providing quality education to all, statistical education is increasingly relying on individualized assessment forms like small-scale surveys. [3].

5. CONCLUSION

Development of quality assurance systems in education generated paradigm shift from traditional to outcome based education, which caused development of assessment culture from previous testing culture. Consequently, new assessment methods based on criteria and standards are replacing previous models. In statistical education it is evident mainly in abandonment of objective tests and acceptance of new innovative assessment techniques and individualized assessment forms.

6. REFERENCES

[1] Adamson, L.; Becerro, M.; Cullen, P.; González-Vega, L.; Sobrino, J.; Ryan, N.: Quality Assurance and Learning Outcomes, European Association for Quality Assurance in Higher Education, Helsinki, 2010

[2] Angelo, T.; Cross, P.: Classroom Assessment Techniques A Handbook for College Teacher, Second Edition, Jossey-Bass, San Francisco, 2003

[3] Bidgood, P.; Hunt, N.; Jolliffe, F.: Assessment Methods in Statistical Education An International Perspective, Wiley, Chichester, 2010

[4] Driscoll, A.; Wood, S.: Developing Outcomes-Based Assessment For Learner-Centered Education A Faculty Introduction, Stylus Publishing , Sterling, Virginia, 2007

[5] Segers, M.; Dochy, F.; Cascallar, E.: Optimising New Modes of Assessment In Search of Qualities and Standards, Kluwer Academic Publishers, Dordrecht, 2003

[6] Standards and Guidelines for Quality Assurance in the European Higher Education Area, 3rd edition, European Association for Quality Assurance in Higher Education, Helsinki, 2009

[7] Walvoord, B.: Assessment Clear and Simple A Practical Guide for Institutions, Departments, and General Education, Second Edition, Jossey-Bass, San Francisco, 2010