

*Panov S.*

## APPLICATION OF STANDART ASSESMENT METHODS IN STUDENT TUTORING IS CONTEMPORARY QUESTION

*The article is devoted to the consideration of concept assessment of student academic achievements. The author of it demonstrates a typical methods in gaining teaching goals. The most common misconceptions and training method-oriented activities for teachers are considered in the article.*

**Key words:** *Assessment, curriculum, formative, summative, quality, target, standards, scoring criteria, communicative method, master, knowledge, skills.*

**Problem stating.** Assessment measures the breadth and depth of learning. It has been criticized as being inaccurate and unreliable, and for distorting both teaching and the curriculum; it is also true that assessment results are notoriously poor at predicting future performance. And yet society and teachers are unable to manage without it. In the right hands, assessment can inspire, motivate and provide the feedback which is essential for targeting prompt corrective help. But it can also lead us to ignore what cannot easily be measured. Assessment serves many different purposes. It can grade the attainment of learners, help to select candidates for jobs or future courses, contribute to evidence on the effectiveness of courses and teachers, and provide a long-term goal for learners. But this applies mainly to the *final* or *summative assessment* of a course, which aims to sum up the learners' achievements. The main use of assessment for teachers is the ongoing or *formative assessment*. This is used throughout the course to form judgments on whether, and to what extent, learning has been successful; and to pinpoint difficulties so that remedial action can be taken. Initial and diagnostic assessments have similar 'formative' roles [1;45].

There are two keys to assessment quality. The first key is to know at the outset how we intend to use assessment results. Sometimes we can use them to promote learning (assessment *for* learning) and other times to check to see if learning has occurred—that is, for purposes of accountability (assessment *of* learning). As the second key to quality, we have established that assessments must be designed to reflect the variety of achievement targets that underpin standards: mastery of content knowledge, the ability to use knowledge to reason, demonstration of performance skills and product development capabilities. Now we consider the third key to classroom assessment quality—how to design assessments that cover our targets and serve our purposes. In this article we describe different communicative assessment methods representing the range of assessment options, explain how to choose which method to use for any given learning target, and outline the steps in assessment planning and development [3;90].

**Basic material** Any assessment, it depends on the day in which it's given, on the, on the environment which the student is coming from. It may be that the student can do the same assessment a week apart and get totally different results. And, what is more, a lot more about the person giving the assessment than the person being assessed, so no, I wouldn't agree [5;67].

Throughout university careers, both as students and as teachers, thousands of different assessments were accounted. Although the variations are endless, all of the assessments we have experienced and give today fall into one of four basic categories of methods:

1. Selected response and short answer;
2. Extended written response;
3. Performance assessment;
4. Personal communication [4;98].

All four methods are legitimate options when their use correlates highly with the learning target and the intended use of the information. Selected response and short answer methods consist of those in which students select the correct or best response from a list provided. Formats include multiple choice, true/false, matching, short answer, and fill-in questions. (Although short answer and fill-in-the-blank do require students to generate an answer, they call for a very brief answer that is counted right or wrong, so we include these options in the selected response category.) For all selected response assessments, students' scores are figured as the number or proportion of questions answered correctly.

Extended written response assessment requires students to construct a written answer in response to a question or task rather than to select one from a list. An *extended* written response is one that is at least several sentences in length. Examples include the following:

- Compare pieces of literature, solutions to environmental problems, or economic events.
- Analyze artwork, forms of government, or solutions to problems.
- Interpret music, scientific information, or polling data.
- Solve a mathematics problem and show and explain all work.
- Describe in detail a scientific, mathematical, or economics process or principle, such as how supply and demand works [7;32].

We judge correctness of extended written responses by applying one of two types of predetermined scoring criteria. One type gives points for specific pieces of information that are present. For example, when students in a biology class are asked to describe the Krebs cycle, points might be awarded for noting that the cycle describes the sequence of reactions by which cells generate energy, takes place in the mitochondria, consumes oxygen, produces carbon dioxide and water as waste products, and converts ADP to energy-rich ATP. The second type of criteria can take the form of a rubric, such as a general rubric for making comparisons, which can be applied to any exercise calling for comparison. Scores therefore also take one of two forms: number or percentage of points attained, or rubric scores [7;32].

Performance assessment is assessment based on observation and judgment; we look at a performance or product and make a judgment as to its quality. Examples include the following:

- Complex performances such as playing a musical instrument, carrying out the steps in a scientific experiment, speaking a foreign language, reading aloud with fluency, repairing an engine, or working productively in a group. In these cases it is the doing—the process—that is important.
- Creating complex products such as a term paper, a lab report, or a work of art. In these cases what counts is not so much the process of creation (although that may be evaluated, too), but the level of quality of the product itself.

As with extended written response assessments, performance assessments have two parts: a performance task or exercise and a scoring guide. Again, the scoring guide can award points for specific features of a performance or product that are present, or it can take the form of a rubric, in which levels of quality are described. For example, to assess the ability to do a simple process, such as threading a sewing machine, doing long division, or safely operating a band saw, points might be awarded for each step done in the correct order. Or, for more complex processes or products, you might have a rubric for judging quality that has several dimensions, such as ideas, organization, voice, word choice, sentence fluency and conventions in writing, or content, organization, presentation, and use of language in an oral presentation. Again, scores could be reported in number or percent of points earned, or in terms of a rubric score. Gathering information about students through personal communication is just what it sounds like—we find out what students have learned through interacting with them.

Examples include the following:

- looking at and responding to students' comments in journals and logs;
- asking questions during instruction;
- interviewing students in conferences;
- listening to students as they participate in class;
- giving examinations orally.

We usually think of this as informal, rather than formal assessment (in which results are recorded for later use). Often it is. However, as long as the learning target and criteria for judging response quality are clear, information gathered via personal communication can be used to provide descriptive feedback to students, for instructional planning, and for student self-reflection and goal setting. If planned well and recorded systematically, information from personal communication can be used as the basis for assessments *of learning* [8;12].

Student responses are evaluated in one of two ways. Sometimes the questions we ask require students to provide a simple, short answer, and all we're looking for is whether the answer is correct or incorrect. This is parallel to scoring for written selected response questions. Questions during instruction usually call for these short answer oral responses.

Other times, student oral responses are longer and more complex, parallel to extended written response questions. Just as with extended written response, we evaluate the quality of oral responses using a rubric or scoring guide. Longer, more complicated responses would occur, for example, during oral examination or oral presentations.

Much "lore" exists about assessment methods, and we address a few of the most common misconceptions here [4;34].

1. *Shouldn't we only be using "authentic" assessments—performance assessments—to judge student progress?* None of these methods is inherently superior to any other, and all are viable if used well. Good assessment means clearly knowing what it is you want to assess and then choosing the best method to get the job done, which, as we will show, depends on the purpose and the learning targets being assessed.

2. *I can see how to involve students in assessment when using a performance assessment, but how do you do it with other methods? Doesn't student involvement require performance assessment?* Although many of our strongest examples of student involvement in the past have come from performance assessment applications, there is a vast, untapped reservoir of student-involvement practices leading to higher learning within each assessment method.

3. *What about portfolios? I notice they aren't listed as a method. Where do they fit in?*

Portfolios are a wonderful idea and we devote an entire chapter to their use later in the book. However, they are not an assessment method, but a vehicle for collecting evidence of, tracking, and communicating about student learning. Portfolios offer a way to involve students deeply in the overall process—self-assessment, tracking progress, reflecting on work, goal setting, and communicating about learning. In this sense, portfolios play a valuable role in creating assessment *for learning* in the classroom.

4. *What about presentations, group projects, worksheets, observations, exhibitions of mastery, posters, and the other ways that teachers gather information?* All of these artifacts and procedures can be classified within the four basic assessment methods described. Presentations and observations are examples of performance assessment. Exhibitions of mastery and group projects can take the form of extended written response, performance assessment, or personal communication depending on how they are carried out. Worksheets are not a method at all because they can contain various types of questions. (Usually worksheets consist of selected response or extended written response questions.) Likewise, posters can be considered either extended written response or performance assessment depending on the assignment and the learning targets being assessed [5;65].

One of the values in classifying assessments according to method is that we can think clearly about how to assess what we are teaching. The heart of accuracy in classroom assessment revolves around matching different kinds of achievement targets, with all the forms and nuances of each, to the appropriate assessment method. This is easily done and can save time in the long run.

To begin thinking about the match between kind of learning target and assessment method, please complete the following two activities. You may want to discuss possible answers with colleagues. Here are activities for training method finding skills.

**Activity 1. Which Method?** Let's say you need to assess student achievement on each of the following learning targets. Which assessment method—selected response/short answer, extended written response, performance assessment, or personal communication—would you choose?

Please jot down your answers and save them for later reference.

1. Ability to write clearly and coherently.
2. Group discussion proficiency.
3. Reading comprehension.
4. Proficiency using specified mathematical procedures.
5. Proficiency conducting investigations in science.

### **Activity 2. Target–Method Match.**

For this activity, you will determine which assessment method is the best match for each of the four kinds of learning targets: knowledge, reasoning, skill, and product targets. To do this you will need to read through the following four scenarios and record your answer to each question by marking an “X” in the appropriate box on Figure 4.2 when your answer is “Yes.” (A printable version of the figure is on the accompanying CD in the file, “Target–Method Match Chart.”) You can put an “X” in more than one box. You can use capital “X” to denote really good matches, and a checkmark to denote an acceptable match under certain conditions (or whatever easily distinguishable marks you wish). On a separate sheet of paper, write your justifications for each answer, as requested. To make your choices, think about accuracy and efficiency: which methods will provide the most accurate information with the highest degree of efficiency? If you are working with a learning team, consider discussing your responses as a group.

#### ***Scenario 1: Assessing Student Mastery of Content Knowledge.***

*Scenario:* You want your students to master specific subject matter knowledge because it represents an important foundation for later work. You plan a series of instructional activities to help your students reach this goal. Now you want to assess to be sure they’ve got it. In this particular case, you want them to know the material outright, not through the use of reference materials.

*Question 1:* Should you assess mastery of this material using selected response or short answer modes of assessment, such as multiple choice, true/false, or matching exercises? Briefly explain your response.

*Question 2:* Should you assess your students’ mastery of this material using an extended written response form of assessment? Defend your answer.

*Question 3:* Should you use a performance assessment to assess students’ mastery of this content knowledge? Defend your answer.

*Question 4:* Do you think the personal oral communication form of assessment—by oral exam, interview, conference, or discussion—could viably assess your students’ mastery of this content knowledge? Why or why not?

### **Activity 2. (Continued).**

#### ***Scenario 2: Assessing Reasoning Proficiency.***

*Scenario:* You are a teacher who has seen to it that your students are able to access important knowledge when required. Now you want to see if they can use that knowledge productively to solve relevant problems. You want to see if they can reason analytically (think about the parts of things) and comparatively (think in terms of similarities and differences), draw inferences, and think critically (take and defend a position on an issue, for example).

*Question 1:* Can you get at these things with selected response or short answer assessments? Why or why not?

*Question 2:* Does extended written response assessment work in contexts where we seek to assess reasoning proficiency? Why or why not?

*Question 3:* Is performance assessment a viable alternative? Why or why not?

*Question 4:* Can we use personal oral communication as an assessment method to probe a student’s ability to reason effectively and solve problems? Defend your response.

#### ***Scenario 3: Assessing Mastery of Skills.***

*Scenario:* You teach French and wish to assess your students’ skill at communicating in that language in a conversational situation. So the skill of *oral language proficiency* is your target.

*Question 1:* Can you assess oral language proficiency in a conversational context using a selected response or short answer mode of assessment? Defend your answer.

*Question 2:* Can you assess these skills using extended written response assessment? Why or why not?

*Question 3:* Will performance assessment work as a basis for assessing the foreign language speaking proficiency of your students? Why or why not?

*Question 4:* Can you use personal oral communication as a basis for assessing conversational skill in a second language? Defend your response.

*Question 5:* Would your responses also apply to other skills such as operating a sewing machine, dribbling a basketball, or reading aloud fluently?

### **Activity 2. (Continued)**

#### ***Scenario 4: Assessing the Ability to Create Quality Products***

*Scenario:* You want your students to be able to create quality products that meet certain specified standards. They might be samples of writing, term papers, technology products, craft products, artistic creations, or others. Your instruction has centered on helping students learn the differences between products that are of high and low quality. You have provided practice in developing products that meet your standards. Now it is time to assess the students' achievement to see if your instruction was effective.

*Question 1:* Can you assess the ability to create these kinds of products using selected response or short answer modes of assessment? Why or why not?

*Question 2:* Will extended written response assessment work for evaluating this kind of achievement? Explain.

*Question 3:* Can performance assessment provide the evidence of proficiency needed to evaluate this kind of achievement target? Defend your response.

*Question 4:* Is personal oral communication a viable way to assess when creation of a product is the target? Why or why not?

**Conclusion** No single assessment method is superior to any other. Selected response, extended written response, performance assessment, and personal communication are all viable options depending on the learning targets to be assessed, the purpose of the assessment, and special student characteristics such as age, English proficiency, or specific learning disabilities [3;40-67].

All assessment development proceeds through the same five stages: (1) identify the purpose, specify the targets, select appropriate methods, decide on relative importance of the targets and sample well; (2) write the questions using guidelines for quality; (3) eliminate as many potential sources of bias and distortion as possible; (4) administer the assessment; and (5) examine the results for areas needing fine tuning. By doing the work at each stage, we can have confidence that our assessments are yielding accurate results. Because only the measurable can be reliably assessed, much of importance is usually ignored by the assessment process – and therefore, all too often, by the teaching process. Both teachers and students tend to the pragmatic view: 'If it's not assessed – ignore it'. And so the assessment tail is rightfully accused of wagging the dog. At least a third of young people emerge from school branded as failures. The emotional damage inflicted on our children and young people by this process can only be guessed at by people like you and me, who for the most part have succeeded in our learning.

Some of these 'failures' go on to reject the norms of the society which has rejected them, and pass into a twilight world of G-roads, drugs, petty crime and imprisonment. It is no accident that over 50% of those in prison are functionally illiterate, in many cases as a result of dyslexia that was not adequately diagnosed or attended to. Failure also has its economic consequences. Advanced economies like ours cannot compete on the world market with cheap labor, but only with the skills passed on by education and training. In 1993, the Audit Commission reported that less than 50% of 17-year-olds were in full-time education. They found that one-third of those in education either dropped out of their courses or failed them. The situation has improved slightly since, but the introduction of new vocational qualifications has not been as helpful as you might think.

These qualifications are vocational by name, but academic by nature, and so offer little to students in search of an alternative to 'book and biro'-based education. This social, psychological and economic damage is due in large part to a curriculum which is heavily academic (see pages 125–6); to norm- rather than criterion referenced assessment, with a consequent bias towards the achieve-

ments of the able; and to a tendency not to recognize and reward qualities which are difficult to measure. Try not to mirror these mistakes in your own assessment. Whatever the summative assessment of a course, for formative assessment consider using competence-based systems, profiles, graded tests and other mastery methods.

These reward the effort and successes of every learner, and encourage the self belief on which future learning relies. Remember that formative assessment has much more impact than summative on learning.

*'Not everything that counts is countable, and not everything that is countable counts.'*  
Albert Einstein.

#### LIST OF THE REFERENCE

1. Bloom, B. S., Madaus, G. F. and Hastings, J. T. (1981) Evaluation to Improve Learning, New York and London: McGraw-Hill.
2. Ericsson, K. et al. (1993) 'The role of deliberate practice in the acquisition of expert performance', Psychological Review, 100, 3: 363–406.
3. Gipp, C. and Stobart, G. (1993) Assessment: A Teacher's Guide to Issues (2nd edition), London: Hodder and Stoughton.
4. Rowntree, D. (1987) Assessing Students: How Shall We Know Them? (2nd edition), London: Kogan Page.
5. Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. psychological Reports, 99, 166-17.
6. Brown, G. T. L., Irving, S. E., & Keegan, P. J. (2007). An Introduction to Educational Assessment, Measurement, and Evaluation: Improving the Quality of Teacher-Based Assessment. Auckland, NZ: Pearson Education NZ.
7. Harris, L. R., & Brown, G. T. L. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. Assessment in Education: Principles, Policy and Practice, 16(3), 379-395.
8. Smith, M. L. (2006). Multiple methods in education research. In J. Green, G. Camilli & P. Elmore (Eds.), Handbook of complementary methods in educational research (pp. 457-475).

**Панов С.Ф., Василичина Ю.М., Василичина Н.М.**

#### **СОВРЕМЕННАЯ ПРОБЛЕМА ПРИМЕНЕНИЕ МЕТОДОВ СТАНДАРТНОГО ОЦЕНИВАНИЯ В ПРОЦЕССЕ ОБУЧЕНИЯ СТУДЕНТОВ**

*Статья посвящена концепции оценки учебных достижений студентов. Автор демонстрирует типичные методы преподавания для достижения учебных целей. В статье рассматриваются самые типичные недопонимания и деятельность ориентированная на практическую деятельность.*

**Ключевые слова:** оценивание, учебный план, словообразовательный, общее оценивание, качество, цель, стандарты, коэффициент оценивания, коммуникативный метод, магистр, знания, навыки.

**Панов С.Ф., Василичина Ю.М., Василичина Н.М.**

#### **СУЧАСНА ПРОБЛЕМА ВИКОРИСТАННЯ МЕТОДІВ СТАНДАРТНОГО ОЦІНЮВАННЯ ДОСЯГНЕНЬ СТУДЕНТІВ**

*Стаття присвячена концепції оцінки учбових досягнень студентів. Автор демонструє типові методи викладання в досягненні учбових цілей. в статті розглядаються найпоширеніші непорозуміння та практично спрямована діяльність студентів.*

**Ключові слова:** оцінювання, навчальний план, словотворчий, загальне оцінювання, якість, мета, стандарти, коефіцієнт оцінювання, комунікативний метод, магистр, знання, навички.