ASSESSMENT IN MEDICAL EDUCATION IN PAKISTAN:
EVALUATING EVALUATION

JUNAID SARFRAZ KHAN, JOHN SG BIGGS, SAIMA TABASUM AND MARYAM IQBAL
Department of Examinations and *Adjunct Professor of Medical Education
University of Health Sciences, Lahore – Pakistan

ABSTRACT
Background: Over the past decade, Pakistan has seen phenomenal growth in Health Profession Education. As the products of these newer medical colleges join the industry, we have received anecdotal accounts of considerable variations in competency, knowledge and attitudes towards the profession, peers, patients and the industry.

Objective: The question addressed in this article is whether the assessment techniques used in medical education in Pakistan evaluate higher order critical thinking.

Methodology: A review of published literature in four international medical education journals, Medical Teacher, Medical Education, Journal of Pakistan Medical Association and Journal of Physicians and Surgeons Pakistan was conducted through systematically searching their databases using keywords. This review covers only the methods used for assessment in medical education in Pakistan at present and their contextual relationship to measurement of critical thinking.

Results: Multiple tools used to assess each of the three domains, cognitive, psychomotor and affective were identified. Each one of these tools in relation to the context can effectively evaluate critical thinking but requires careful planning and proper application. Tools used elsewhere (outside Pakistan) were holistic in their measurement with high contextual relevance.

Conclusion: Critical thinking sets higher education apart. Currently the tools of assessment employed to evaluate knowledge, skills and attitudes in medical education in Pakistan are sound but require a critical analysis and review in their construct and applicability in relation to the context. Better tools are also available that can be used to ‘teach’ as well as ‘assess’ critical thinking.

THE DILEMMA
Over the past decade, Pakistan has seen phenomenal growth in Health Profession Education especially in the Baccalaureate of Medicine; Baccalaureate of Surgery (M.B.B.S.) and Baccalaureate of Dental Surgery (B.D.S) both in the public (Government subsidized) and the private (with heavy economic implications) sectors.

As the products of these newer medical colleges join the industry, we have received anecdotal accounts of considerable variations in competency, knowledge and attitudes towards the profession, peers, patients and the industry. There is, therefore, a need to review the curriculum including the best practices in curriculum implementation and assessment. We need to consider how to adapt to these newer changes especially the mismatch between the teacher: taught ratio and how to adopt the Best Evidence Medical Education (BEME) practices blended to our local needs and culture.2

The objective of this article is to provide curriculum and assessment developers with a rationale for choosing their own approach to teaching / learning and assessment in changing times. We aim to provide an overview of the current techniques employed in assessment of cognitive, psychomotor and affective domains in Medical Education in the country and how best we can learn from the evidence to adopt different techniques to improve our product with the aim of improving healthcare delivery.3,4

ASSESSMENT
Medical Education in Pakistan generally follows the annual summative examination system with only 10% of the evaluation dedicated to the year-long continuous assessment. Neither the summative nor the continuous assessment is used to provide any formal feedback to the students, the teachers, administrators, industry or the public. However, these stakeholders can by and large, make educated guess/ inference on the quality of medical education – curriculum, training and assessment – based on the qualitative and quantitative results of the end – of – year evaluation.5

The adage that assessment drives education or, assessment drives learning and teaching has commonly been used in a negative sense. For decades, assessment has been considered the monster that
gobbles up all that is holy and good about education.\textsuperscript{6-8} Only recently, have we begun to realize the true implications of the fact that assessment indeed “drives” every aspect of educational activity; that it provides the impetus and force that gives direction and meaning to teaching and learning; defines teaching methodologies; molds attitudes and concepts related to education and fashions the product into what it is.

In the United Kingdom the first and foremost objective of Higher Education is to develop critical thinking amongst the students.\textsuperscript{9} Developing skills in life – long – learning, self – directed – learning, problem – solving – learning etc. are all hallmarks of a sound higher educational programme but eventually it is the ability to critically think, evaluate and synthesize knowledge, competencies and attitudes that define the essence of higher education.

Since assessment drives education, it is reasonable to expect that assessment and evaluation of the ability to critically think, reflect, evaluate, synthesize and “create” knowledge, competencies and attitudes will be included in testing the objectives of Higher Education and will be driving the educational strategies and methodologies to achieve these objectives.

**RESEARCH QUESTION**
The question addressed in this article is whether the assessment techniques used in medical education evaluate higher order critical thinking.

**METHODS**
In this study we chose to investigate the various tools currently used to assess cognitive, psychomotor and affective domains in medical education in the country and how this assessment correlates with the high stakes goals of developing critical thinking including reflection, life – long – learning and self – directed learning in our medical and dental graduates. The literature included in this review was obtained by searching the databases of four international journals, Medical Teacher, Medical Education, Journal of College of Physicians and Surgeons of Pakistan and Journal of Pakistan Medical Association. The key terms used in the search were ‘assessment techniques’, ‘methods of assessment’, ‘assessment of critical thinking’, ‘assessment of reflection’, ‘learning outcomes’, ‘assessment’, ‘medical education’, ‘assessment in higher education’ and ‘assessment of higher order thinking’. The references of all retrieved articles were scanned to identify additional articles. Over 110 articles related to the topic were found in this search. Before any research article was included in this review, its study design, sampling and methodology were assessed. Reliable and valid studies were included as these articles were identified as providing insight into the current practices related to assessment in the country and assessment of critical thinking norms elsewhere. (Table 1).

**RESULTS**
The review identified that currently in medical education assessment in Pakistan, the tools used are the following:

1. **Cognitive domain**
   a) Oral examination.
   b) Multiple Choice Questions.
   c) Structured Answer Questions.
   d) Long Essay Questions.

2. **Psychomotor domain**
   a) Direct observation of clinical skills (Long and Short cases).
   b) Objectively Structured Performance Evaluation (OSPE).
   c) Objective Structured Clinical Evaluation (OSCE).
   d) Task Oriented Assessment of Clinical Skills (TOACS).
   e) Practical examination.

<table>
<thead>
<tr>
<th>Sr. No.</th>
<th>Search Criteria / Key Words</th>
<th>Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Assessment techniques</td>
<td>2, 3, 16, 17, 38, 52</td>
</tr>
<tr>
<td>2.</td>
<td>Methods of assessment</td>
<td>7, 16, 38, 43, 44, 45, 46</td>
</tr>
<tr>
<td>3.</td>
<td>Assessment of critical thinking</td>
<td>4, 5, 9, 18, 19, 20, 28, 32, 35, 36, 40</td>
</tr>
<tr>
<td>4.</td>
<td>Assessment of reflection</td>
<td>21, 25, 34, 37, 39, 41, 46</td>
</tr>
<tr>
<td>5.</td>
<td>Learning outcomes</td>
<td>20, 26, 34, 37, 47, 48, 49, 50, 51</td>
</tr>
<tr>
<td>6.</td>
<td>Assessment</td>
<td>1, 6, 8, 22, 23, 29, 30, 40, 42, 53</td>
</tr>
<tr>
<td>7.</td>
<td>Medical education</td>
<td>36</td>
</tr>
<tr>
<td>8.</td>
<td>Assessment in higher education</td>
<td>27, 38</td>
</tr>
<tr>
<td>9.</td>
<td>Assessment of higher order thinking</td>
<td>4, 7, 24, 46</td>
</tr>
</tbody>
</table>
2. **Affective domain**
   a) Interviews.
   b) Direct observation of communication skill and behavior.

These are explained in further detail below:

**CRITICAL THINKING AND REFLECTION**

Critical thinking is defined as “the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing and / or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief or action”.

Central to the definition of critical thinking is learning through application, analysis, evaluation and synthesis (table 2). It focuses on learning through reasoning following observation, experience or reflection.

Critical thinking therefore, focuses on the higher level of Bloom’s taxonomy; takes the Kolb’s Model (Fig. 1) of experiential learning one step further by turning this cyclical process into an upward spiral whereby each turn of the spiral adds to the overall experience through reasoning, evaluation and synthesis.

Reflection, a term that most educationalists romanticize with is a kind of thinking that consists in turning a subject over in the mind and giving it serious thought. According to Race, reflection could be argued to be the essential stage where learning is integrated within the whole learner, and added to existing frames of reference, and internalized and personalized. Reflection is important to critical thinking because it allows one to learn from past experience, modify current practices based on reasoning and plan for the future by making conscious alterations to knowledge, skills and attitudes.

**Table 2: Definition of terms (Oxford Dictionary)**

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td>The action of putting something into operation.</td>
</tr>
<tr>
<td>Analysis</td>
<td>The detailed examination of the elements or structure of something.</td>
</tr>
<tr>
<td>Evaluation</td>
<td>Forming an idea of the amount, number, or value.</td>
</tr>
<tr>
<td>Synthesis</td>
<td>The combination of components or elements to form a connected whole.</td>
</tr>
</tbody>
</table>

**BLOOM’S TAXONOMY**

Bloom’s taxonomy of cognitive, psychomotor and affective domains is reproduced in table 3. Since assessment drives education, it can be argued that all the assessment techniques identified as being currently used in medical education can be employed to assess critical thinking and reasoning in order to drive the educational goals of fostering and developing critical thinking in the medical and dental graduates. Multiple Choice Questions are easy to
mark but very difficult to construct with high validity and fidelity especially at higher levels of cognitive assessment.\textsuperscript{16,17} Structured Answer Questions and Long Essay Questions may be easier to develop at higher levels. However, studies have shown that testing at any level with SAQs and LEQs is fraught with the dangers of lack of construct validity both in the question and the key (known as item writing flaws) and inter – rater bias.\textsuperscript{18} This subjectively when combined with the human and economic resource expenditures related to their construction and marking, allows a well constructed MCQ to be the first choice of cognitive assessment at all levels in comparison. However, Long Essay Questions may continue to retain their importance as an effective tool of assessment of evaluation and synthesis at the very top of Bloom’s taxonomy.\textsuperscript{16} Oral examination too, when carefully administered may very well do the same with the added benefit of measuring communication skills both verbal and body language and general attitudes and behavior.\textsuperscript{19–21} 

Direct observation of clinical and practical

<table>
<thead>
<tr>
<th>Cognitive Domain</th>
<th>Behavior Description</th>
<th>Examples</th>
<th>Keywords</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Knowledge</strong></td>
<td>Recall data or information</td>
<td>Multiple – choice test, recount facts or statistics, recall a process, etc.</td>
<td>Arrange, define, describe, label, list, recognize, relate, reproduce, select, state</td>
</tr>
<tr>
<td><strong>Comprehension</strong></td>
<td>Ability to grasp the meaning of material.</td>
<td>Explain or interpret meaning from a given scenario or statement, suggest treatment,</td>
<td>Explain, reiterate, classify, gives examples, illustrate, translate, review, report, discuss, re-write.</td>
</tr>
<tr>
<td><strong>Application</strong></td>
<td>Ability to use learned material in learned material in new and concrete situations.</td>
<td>Put a theory into practical effect, demonstrate, solve a problem.</td>
<td>Use, apply, discover, manage, execute, solve, produce, implement, construct, change, prepare</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td>Interpret elements, organizational principles, structure, construction</td>
<td>Identify constituent parts and functions of a process, or de-construct a methodology or process.</td>
<td>Analyze, break down, catalogue, compare, quantify, measure, test, examine, experiment, relate, graph, diagram, plot</td>
</tr>
<tr>
<td><strong>Synthesis</strong></td>
<td>Ability to put parts together to form a new whole.</td>
<td>Develop plans or procedures, integrate methods, resources, ideas.</td>
<td>Develop, plan, build, create, design, revise, formulate, propose, establish, assemble</td>
</tr>
<tr>
<td><strong>Evaluation</strong></td>
<td>Ability to judge the value of material for a given purpose.</td>
<td>Select the most effective solution. Hire the most qualified candidate.</td>
<td>Review, justify, assess, present a case for, defend, report on, investigate, direct, appraise, argue</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Affective Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Receiving</strong></td>
</tr>
<tr>
<td><strong>Responding</strong></td>
</tr>
<tr>
<td><strong>Valuing</strong></td>
</tr>
<tr>
<td><strong>Organization</strong></td>
</tr>
<tr>
<td><strong>Internalize or characterize values</strong></td>
</tr>
</tbody>
</table>
Psychomotor Domain

<table>
<thead>
<tr>
<th>Perception</th>
<th>The ability to use sensory cues to guide motor activity.</th>
<th>Detects non-verbal communication cues.</th>
<th>Recognize, distinguish, notice, touch, hear, feel, etc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Set</td>
<td>Readiness to act</td>
<td>Mental, physical or emotional preparation before experience</td>
<td>Arrange, prepare, get set, states, volunteers</td>
</tr>
<tr>
<td>Guided response</td>
<td>The early stages in learning a complex skill that includes imitation and trial and error.</td>
<td>Imitate or follow instruction, trial and error</td>
<td>Imitate, copy, follow, try</td>
</tr>
<tr>
<td>Mechanism</td>
<td>Basic proficiency</td>
<td>Competently respond to stimulus for action</td>
<td>Make, perform, shape, complete</td>
</tr>
<tr>
<td>Complex Overt Response</td>
<td>Skillful / expert proficiency</td>
<td>Execute a complex process with expertise</td>
<td>Coordinate, fix, demonstrate</td>
</tr>
<tr>
<td>Adaptation</td>
<td>Skills are well developed and the individual can modify movement patterns to fit special requirements.</td>
<td>Alter response to reliably meet varying challenges</td>
<td>Adapts, alters, changes, rearranges, reorganizes, revises, varies</td>
</tr>
<tr>
<td>Origination</td>
<td>Creating new movement patterns to fit a particular situation or specific problem.</td>
<td>Develop and execute new integrated responses and activities</td>
<td>Design, formulate, modify, re-design, trouble – shoot</td>
</tr>
</tbody>
</table>

skills, attitudes and behaviors using long and short cases, OSCE or any of its variants like TOACS and OSPE are indispensable tools of higher order skill and attitude measurement if applied correctly\(^\footnote{22,23}\). Long and Short cases can be standardized and made nearly as objective as OSCE.\(^\footnote{24,39}\)

The problem therefore, is not in the tools deployed for assessment in Pakistan but the way they are structured and the level of cognitive, psychomotor or affective domain they actually measure. Ineffectively used at lower levels of cognition, skill and attitudes, they deliver the wrong and potentially life-threatening (to the public) message of lower-order cramming and lack of professionalism, demeaning the very goals that set the higher education apart.

The need, therefore, is to understand assessment as a science and stop assessing for the sake of assessment. Unless concrete steps are taken to ensure that assessment techniques holistically measure higher levels, assessment shall do more damage than good.

**ASSESSMENT OF CRITICAL THINKING ELSEWHERE**

Being knowledgeable and the ability to critically think cannot be equated.\(^\footnote{31}\)

‘…… knowledge is no more a substitute for thinking than thinking is a substitute for knowledge …… There are too many brilliant academics whose brilliance in their own field and lack of it outside those fields shows the difference between knowledge and thinking.’

That critical thinking can be ‘taught’ has been supported by research of Coles and Robinson and De Bono.\(^\footnote{32,33}\) However critical thinking alone is not sufficient for problem solving.\(^\footnote{34-37}\) Nevertheless, that, critical thinking can be taught implies that it can be measured and assessed. Construction of Multiple Choice Questions, Structured Answer Questions / Short Essay Questions and Long Essay Questions in a fashion that they require the respondent to apply knowledge, critique it and analyze it, synthesize and create requires time and as is cognitively challenging but potentially rewarding.

Increasingly, students are being presented with problem-solving scenario in OSCE or its variants and in practical and clinical examinations. The use of reflective writing and its addition in assessment repertoire, 360° evaluations, work-place and work-based evaluation, Mini – CEX exercises, peer and self assessment techniques, all increase the content in assessment of critical thinking, reflection and problem solving.\(^\footnote{38-45}\)

Portfolios and eportfolios can also be innovatively used for assessing the course of professional and personnel development, management and organization, problem solving and critiquing existing concepts, knowledge and competencies and moving on to the realm of creativity.\(^\footnote{46,47}\)

Assessment in medical education addresses complex competencies and thus requires quantita-
tive and qualitative information from different sources as well as professional judgment. Adequate sampling across judges, instruments and contexts can ensure both validity and reliability. 48-51

Feedback can in turn promote critical thinking and when the assessments are combined with formal and informal feedback mechanisms, it allows the students to think laterally. 52-53

In conclusion critical thinking sets higher education apart. If critical thinking is to be taught in higher education, it needs to be assessed as well, since assessment drives education. Currently the tools of assessment employed to evaluate knowledge, skills and attitudes in medical education in Pakistan are sound but require a critical analysis and review in their construct and applicability in relation to the context. Better tools are also available that can be used to ‘teach’ as well as ‘assess’ critical thinking especially when feedback is made an essential component of all aspects of education including assessment.

ACKNOWLEDGEMENTS

The authors are thankful to the Vice Chancellor of UHS for allowing to prepare this report.

REFERENCES


26. Harden RM, Crosby J, Davis MH, Howie PW, Struthers AD. Task-based learning: the answer to integra-
Distinguish between assessment and evaluation processes in adult education. Apply assessment and evaluation techniques appropriate to the context in which they deliver their courses or programs. Use different methods to assess and evaluate individual learning throughout the course. Identify and describe a variety of written, performance and alternative assessment formats that link assessment to course goals. Specify and design criteria necessary to assess learning objectively. Describe the role technology plays in assessment and evaluation (i.e. tools for design and tracking). Identify ethical challenges/potential barriers in the design of and use of assessment and evaluation. Recent papers in Assessment in Medical Education. Papers. People. Development of a Brazilian Portuguese adapted version of the Gap-Kalamazoo Communication Skills Assessment Form. The literature has revealed that OSCE is an effective evaluation tool for evaluating the clinical skills of medical students. The objectives of this study were to investigate perception of the students and staff of OSCE as a clinical assessment tool during their clinical years at the Faculty of Medicine, Suez Canal University (FOM-SCU) as well as the perception of both the staff and medical interns regarding their ethical concerns of using medical interns as simulated patients in OSCE stations. Medical education in Pakistan. October 2011. Authors: Idrees Mohammed, Idrees Anwar. Rawalpindi Medical University, Rawalpindi, Pakistan. Download full-text PDF. Read full-text. Assessment design, pre exam and post exam analysis, standard setting, quality assurance, system of audit and feedback, program evaluation are some of neglected areas that need gross changes. I feel the time has come where we would start rising again as our eastern traditions of medical excellence and once again show the world what we are capable of achieving through our faith and hard work. Research studies on the evaluation of foreign aid in education have developed an argument on the effectiveness and usefulness of foreign aid. Some studies have even discussed the reasons for failure of foreign aid. USAID was evaluated and some key principles were underlined for evaluating foreign assistance. Evaluation of UK aid (DFID) has critically reflected on its procedures. The main objective of this study was to evaluate TEP to see the influence of the programme on education in Pakistan. The research has specifically explored the operations of TEP and its benefits to the education system of Pakistan. The following research questions were posed to conduct the research:- 1. Who were the recipients of TEP?