MEDICAL EDUCATION

ASSESSMENT DRIVEN INTEGRATED LEARNING (ADIL): ASSESSMENT DIRECTED MEDICAL EDUCATION (ADME) CURRICULUM

Junaid Sarfraz Khan

Department of Examinations, University of Health Sciences, Lahore, Pakistan

Assessment has been shown to drive learning. For assessment to drive learning in the right direction, it needs to fulfil certain criteria including reliability, validity, educational impact, feasibility, acceptability to stakeholders, cost-effectiveness etc. Assessment has the greatest effect as a driving force in learning if it is authentic, context relevant and samples widely. In this paper, we present a new curriculum design and a model for experiential learning through assessment. We call it Assessment Directed Medical Education (ADME) Curriculum using Assessment Driven Integrated Learning (ADIL). In this model, we propose a cent percent rightest shift in the curriculum design and learning strategy by using assessment as the sole tool to a self-directed learning process through supervised simulated and real-patient encounters testing competencies in cognitive, psychomotor and affective domains in an integrated approach utilising key resources for active learning. It is presumed that, this will, in the long term improve health care delivery by producing competent healthcare professionals that have learned through direct student-'patient' encounters from day one of their medical education and are competent in making professional judgments, exhibit appropriate skills and behaviours.

Keywords: Assessment, Learning technique, CME, Curriculum, Evaluation, Competency based

INTRODUCTION

It is generally accepted that assessment drives learning and influences students' learning strategies.[1–3] Apart from this extrinsic effect, assessment also has an intrinsic effect on the memory of studied material.[45] The act of remembering is based on two strengths: storage and retrieval. When students study they store the data or material and when they are tested or have to apply knowledge, skills and attitudes to real life problems, they retrieve the data and information.[6–8] This ability to retrieve data is tested by testing and by following Bloom's taxonomy[9] as tests become more complex, the application of knowledge, skills and attitudes; their integration, synthesis and synergy is tested by assessment as well. This in turn develops critical thinking, strengthens skills and establishes behavior.[10,11] This is the intrinsic effect of testing and the more frequently data and information is retrieved, the more complex the situation and more context relevant the assessment, more pronounced is this intrinsic effect.[12–14]

In 1996, Cees PM Van Der Vleuten presented a conceptual model for defining the utility of an assessment method[1]. The utility of an assessment method was judged based on it's validity, reliability, educational impact through extrinsic and intrinsic effects, acceptability to all stake-holders and cost-effectiveness in terms of time, structural and human resources and emotional cost.[15]

The utility of an assessment method depends on how well the criteria related to the method are weighted by the stake-holders. This basically implies that an assessment method can be of high utility in one situation and not so in another.[16] This contextual relevance of the assessment methods is an important factor in choosing the right method for the right context. To take the argument further, in order to strengthen the effect of assessment and to gather meaningful results out of it, we need to improve on its reliability, validity, context-relevance and cost-effectiveness.[17] To do so, we have to think outside the box. For too long, we have been 'trivialising' assessment; choosing one method for assessment of one competency[18]. The problem with it, is that competencies are interrelated. Problems in the real world do not present that require either, a cognitive, a psychomotor or an affective solution but a combination of these and retrieval of data, skills and attitudes at multiple levels with their proper applications.

If we accept that assessment drives learning, that students can be made to actively learn through assessment; assessment no longer remains a measurement tool but an instructional design problem including educational, implementation and resource aspects.[19] In such a case, any assessment method can be measured on a scale based on context-relevance and design that ranges from solely measurement to explicitly instructional. On such a scale, any assessment method's positioning shall be effected by the context it is used under. For example, Multiple Choice Questions used in the Medical Colleges Admission Test (MCAT) pose more of a measurement than instructional problem but, the same questions when used for formative assessment in a pre-test have a high instructional value when feedback motivates learning and improvement.

In this paper, we present a competency-based assessment curriculum model. However, in order to present the model more coherently it is important that we revisit important criteria of good assessment, namely, reliability, validity, authenticity and impact on learning.

Reliability
Reliability which is expressed as a co-efficient ranging from 0 (no reliability) to 1 (perfect reliability) refers to the reproducibility of the scores obtained from an assessment[15]. Simply stated, a test or an assessment method is said to be reliable if it is consistent and this consistency for a good assessment method should be represented by at least 0.80 as a co-efficient. However, reliability can be outset by bias.[20] To reduce bias, a number of protocols can be observed:

a) Use more then one method to assess similar competencies.
b) Assess on more than one occasion.
c) Use more patients, examiners, items, scenarios etc. to assess.
d) Use methods that are contextually relevant.

The greater the sampling the more reliable shall be the final judgment on the assessment.[21] A single assessment at a single moment in time provides one snapshot of the competency or range of competencies tested, whereas by following the above protocol a series of snapshots provides a better measure of the same. It has been observed that by applying the protocol above, reliability can be increased for both objective, standardized methods of testing like OSCE and subjective, judgmental methods like long cases and oral examinations.[15,22] The recent trend of portraying standardized objective assessment tools as the ultimate in assessment technology is nothing more than a fad, since more subjective tools can be used with equal if not better results provided they sample widely. As long as sampling is appropriate across conditions of measurement any method can be sufficiently reliable. A method in itself can not be termed reliable or unreliable.[23]

Validity
Validity refers to whether an assessment tool measures what it is supposed to measure[7,24]. The ability of a student to pass a nasogastric tube can not be assessed by a pen-and-paper test. A test that requires the student to write down the steps of nasogastric tube insertion, tests simple recall of the steps involved in passing the tube. If the purpose of the test is to determine the competence of the candidate in passing the tube, the test is not valid, it is context irrelevant.

Authenticity
Increasingly, stakeholders demand that competencies should be measured with authenticity, by offering candidates simulated real world challenges whether on paper as clinical scenarios for example, in clinical skills laboratories, or in computerized formats.[25,26] More recently, realistic simulations have been replaced by short and focused vignettes that require application of knowledge and synthesis of information and focus on the ability to make correct key decisions.[27]

Further still, we witness the movement of assessment from examination halls and clinical skills laboratories to workplace and wards in search of absolute authenticity with use of real patients in the day-to-day practices.[28,29] In fact, some assessment tools developed have not only taken the assessment to the work place but have incorporated judgment of patients and their experiences of the performance of candidates as part of final score of the measurement. This provides a more holistic, real and live assessment, a pinnacle of authenticity. This is because these methods do not test singular competencies in isolation but a range of competencies at multiple levels, thereby increasing both validity and reliability and being cost-effective at the same time. This has also shown to promote learning because the modern educational theory suggests that learning is facilitated when tasks are integrated.[30,31] Since a competency is the ability to handle a complex professional task by integrating the relevant cognitive, psychomotor and affective skills, curricula can be built around competencies to facilitate integrated learning and measurement of these complex professional tasks.[15]

What is to be avoided in this model is reducing competencies into parts that separately on their own mean nothing. Instead competencies need to be defined, learned and measured holistically to give meaning and form to professional tasks.[32] Fundamental to this shift will be a move towards testing these competencies through multiple tools and instruments thereby, increasing the reliability and validity of the final judgment. After all, it is the learning and assessment of professional judgment that is of the greatest importance in medical education.[33] This is not to say that general professional competencies like team-work, critical thinking, professional behaviours are in any way less important in medical education. On the contrary, tools and assessment instruments in medical education routinely incorporate assessment of general professional competencies while measuring competencies related to medical teachings in particular.[1]

Consequential validity
We do not doubt that assessment affects learning; it is perhaps one of the strongest driving forces steering learning of students. If it is not context relevant, holistic, reliable and valid, it can steer learning in all the wrong directions with disastrous effects in medical education in particular and the health care system in general. This impact of assessment on learning also called,

consequential validity, is therefore important and can not be overlooked.[15]

Having discussed some important issues related to assessment, we would like to present a new curriculum model centred around assessment called Assessment-Directed Medical Education (ADME) Curriculum.

Assessment Driven Integrated Learning (ADIL)

Currently medical education in Pakistan follows medical curricula that have the following common characteristics:

a) Medical education is divided into five academic years covering 14 subjects.
b) A subject may be taught for one, two or up to three years but shall be examined once in a summative examination.
c) In the syllabus provided, each subject is covered by a list of topics and subtopics, much like a table of contents or an index.
d) Cognitive, psychomotor and affective competencies are not delineated for any subject in general or its topics and subtopics.
e) The depth and breadth of teaching in each subject and its topics and subtopics is left for individual institutions, departments and faculty members to decide.
f) Teaching methodologies consist of lectures lasting up to an hour, tutorials lasting up to three hours, laboratory and clinical work which is mostly unsupervised and very occasional field trips for community orientation.
g) Assessment is 90% end-of-year and 10% continuous. Continuous assessment is not structured and the end-of-year assessment consists of a theory exam for cognitive assessment comprising of Multiple Choice Questions and Short Essay Questions; a practical (basic sciences) or clinical examination consisting of short and long cases; viva voce and an Objective Structured Performance Evaluation (OSPE) comprising of up to 20 stations in some subjects but mostly these stations are static and very few are interactive.

h) There is no systematic means of providing feedback to the students, the faculty or program managers.

The system is 'safe', because it is 'traditional' and 'age-old'. In the 'developed' world, the curricula are either called outcome-based, community-oriented, objective or competency-based. Whatever name may be given to these curricula they broadly:

a) Show structure to and objectivity in medical education in general.
b) Define the competencies that shall be gained by the candidate at the end of the program in all three domains.
c) Describe how these competencies shall be gained.
d) Classify the objectives of the course and the outcomes of the program.
e) Describe clearly the role of the program developers, candidates, faculty and community in 'achieving' those objectives, outcomes and competencies.
f) Provide directions for accessing resources to achieve the outlined goals.
g) Indicate educational strategies that can be used for a problem solving, critical thinking learning process.
h) Base on a system of feedback to all stakeholders for continuous improvements in the program at all levels for all concerned.
i) Fashion pass/fail judgments around formative, continuous and summative assessments measuring general and specific objectives and competencies.
j) Foster self-directed, life-long learning traits in the candidates.

These curricula are 'modern' and 'safe' for all stakeholders especially public, because they are based on authenticity of education and evaluation Figure-1.

ADIL is the ultimate step forward in the integrated competency-based self directed learning experience. In ADIL, we propose a paradigm shift from teaching, to learning through experience using resources available under expert supervision throughout the program. We propose a complete shift away from passive teaching in any form to active learning in every area of the competency. This will put the 'Does' at every level of the Miller's pyramid (Figure-2).
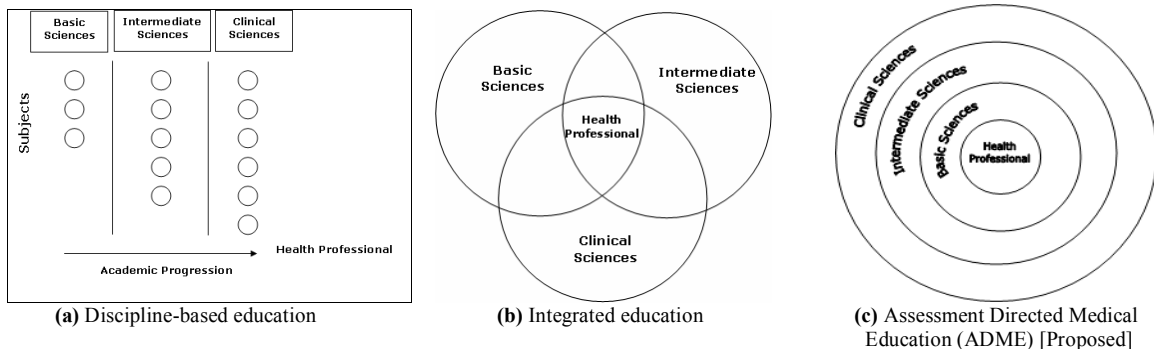

(a) Discipline-based education   (b) Integrated education   (c) Assessment Directed Medical Education (ADME) [Proposed]

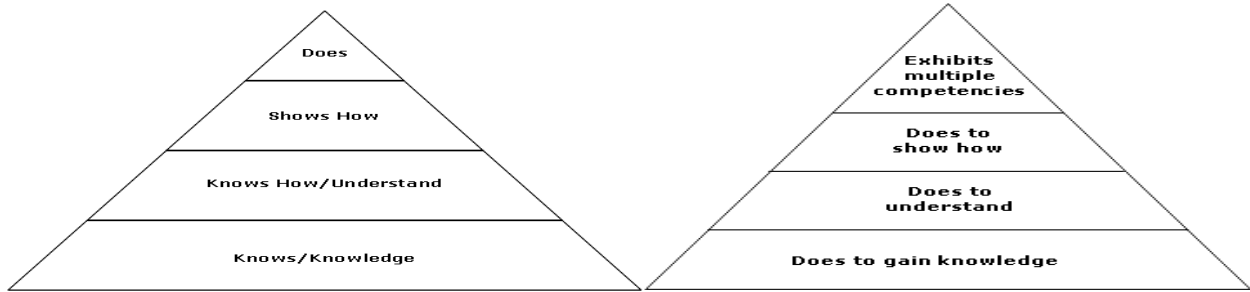**Figure-1: Different models of medical education currently in practice**
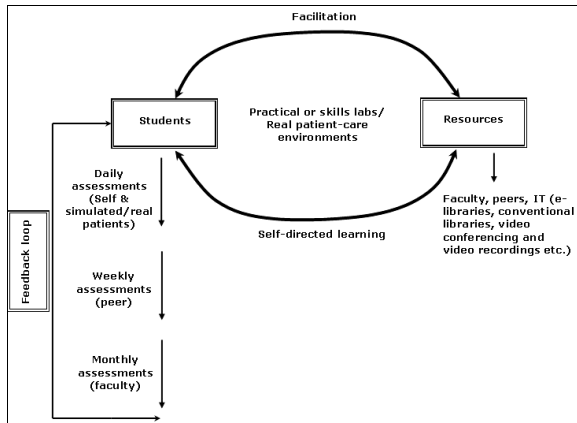
Figure-2: Miller's pyramid

Figure-3: Assessment Directed Medical Education

Key features of ADIL:

a) The curriculum is divided into five broad academic years.

b) Each year, a certain number of major and minor competencies are to be achieved by the candidate in a systematic structured fashion in the subjects delineated.

c) This requires that each subject is pronounced through listing of the core and associated cognitive, affective and psychomotor competencies required as minimum criteria for the candidate to be declared a safe practitioner.

d) Competencies are no longer addressed individually subject-wise but integrated horizontally across subjects in every academic year and vertically amongst subjects throughout the program.

e) By horizontally and vertically integrating competencies, education and evaluation is made authentic. i.e., education and evaluation takes place in real-time, based on real-life practices.

f) The final curriculum design is a horizontally and vertically integrated competency document in all medical and allied subjects and not in any way a list of topics to be taught/learned. General professional competencies like team-work and professional behaviour are integrated at all levels.

g) Teaching/training is taken from out of the lecture halls and tutorial classes into simulation labs and skills laboratories from day one of the program and progressively to workplace and day-to-day patient-encounters and patient-care areas.

h) Students are provided with clear time-line for gaining the competencies in each academic year.

i) Students are provided information about the nature and methods of accessing resources including libraries, e-libraries, information technology, community encounters, faculty support, peer and support groups, video-recordings of encounters of patients with faculty and students.

j) Students are exposed in groups of up to 10 to structured, supervised simulated-patient/real-patient encounters presenting common problems that require solutions through tapping into cognitive, psychomotor and affective domains at various levels and demonstrating competencies across subjects and disciplines.

k) The initial simulated encounters serve as a feedback to the students and supervisors on the deficiencies in domains and competencies.

l) The students tap into the resources individually and in groups to make up for those deficiencies.

m) The simulated patient encounters are repeated till a critical level of competence is observed and then the group moves on to the next competency according to the timeline.

n) As the years progress, the clinical content is increased, so that problems/encounters require solutions through and evaluation by higher order thinking, psychomotor and affective domains at a level at par with a Houseman in year 5 of medical education with exposure to real patients in real environments.

o) End-of-year pass/fail judgments are based on the progress shown by each candidate in achieving mastery in competencies which will be through continuous assessment as well as an end-of-year examination requiring the candidate to exhibit these competencies in a simulated or work-place real environment.

p) Throughout the program while groups rotate from one competence to the other, peer, self and faculty/supervisor simulated/real patient feedback, i.e.,

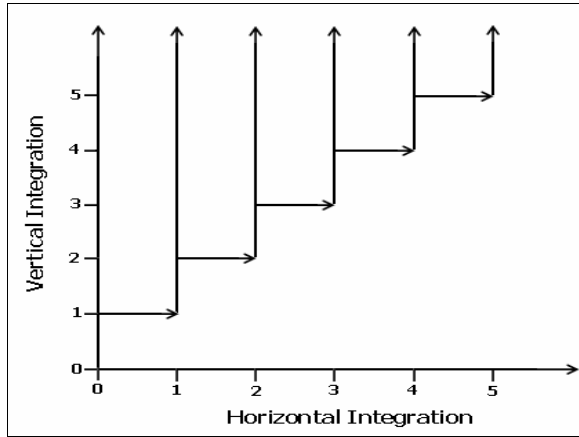360° feedback provides a percentage score to the continuous assessment and the impetus to learn.

Figure-4: Integration model in ADIL (Proposed)

ADME ensures that education is valid and reliable because it simulates real-life environment and exposure from day one of the program. For a health professional, education is not only solving problems but experiencing these problems[34]. By ensuring that the students experience the problems that they are supposed to encounter in their practice on graduation, ADIL transforms assessment into a learning strategy that is authentic and based on experiential learning. By requiring students to face problems head-on and tap into available resources, this instructional design shifts their attitude towards self-reliance and self-directed adult learning[35]. This shift is important if we are to produce competent health professionals.
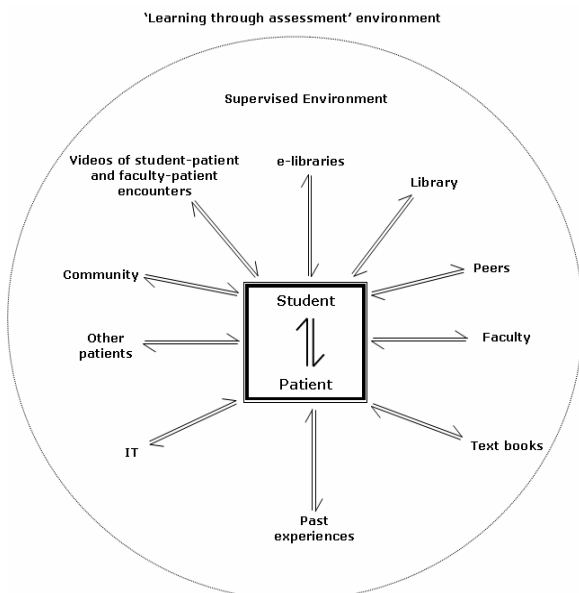

Figure-5: Competency in handling the patient and his problems in a controlled but practice-based environment
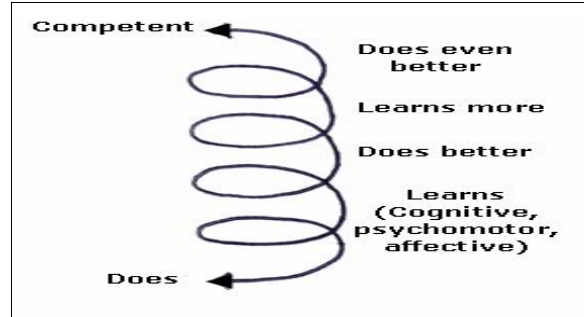

Figure-6: The road to competent Health Professional

CONCLUSION

Validity and reliability are important criteria for good assessment. By increasing sampling, a wider range of competencies in all domains can be tested at multiple levels by multiple examiners. For assessment to have any real impact on learning, it has to be context relevant. The greatest motivational force for students to learn is assessment. This is because assessment and the concept of being judged and 'loose' and 'gain' on the basis of that judgment is part of every aspect of our lives and is ingrained in the human psyche. If assessment is context-relevant and samples widely, it can drive learning in the right direction. ADME curriculum proposed is based on our knowledge that assessment drives learning. ADIL is a concept which is feasible as well as cost-effective and can prove to have lasting improvements in the health care delivery through producing competent self-directed adult-learning healthcare professionals. It does require a major paradigm shift in an approach to teaching and learning but once established, it can be shown to add the authenticity to medical education and evaluation that has been lacking so far. Resources in the form of skills laboratories and simulated patients and environments are widely available in the 'developed' world and a host of day-to-day practice environments and patients readily accessible in the 'less developed' world. We introduce this concept as the ultimate step forward to a self-directed, authentic, real time, live learning experience through direct action.

REFERENCES

1. Van der Vleuten CPM. The assessment of professional competence: developments, research and practical implications. Adv Health Sci Educ 1996;1(1):41–67.
2. Gibbs G, Simpson C. How Assessment Influences Student Learning – a Conceptual Overview. Centre for Higher Educ Prac; 2002.
3. McLachlan JC. The relationship between assessment and learning. Med Educ 2006;40:716–7.
4. Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves longterm retention. Psychol Sci 2006;17:249–55.
5. Woollard M, Whitfield R, Newcombe RG, Colquhoun M, Vetter N, Chamberlain D. Optimal refresher training intervals for AED and CPR skills: a randomized controlled trial. Resuscitation 2006;71:237–47.

6. Wheeler MA, Ewers M, Buonanno JF. Different rates of forgetting following study versus test trials. Memory 2003;11:571–80.
7. Downing SM. The metric of medical education. Validity: on the meaningful interpretation of assessment data. Med Educ 2003;37:830–7.
8. Arthur W Jr. Factors that influence skill decay and retention: a quantitative review and analysis. Hum Perform 1998;11:57–102.
9. Bloom BS. Taxonomy of educational objectives, Handbook 1: The cognitive domain. New York. David McKay Co Inc. 1956.
10. Druckman D, Bjork RA. Optimising longterm retention and transfer. In the Mind's Eye: Enhancing Human Perform. National Academy Press; 1991. p. 23–56.
11. Berden H, Willems F, Hendrick J, Knape J. How frequently should basic cardiopulmonary resuscitation training be repeated to maintain adequate skills? BMJ 1993;306:1576–7.
12. Kuo TM, Hirshman E. Investigations of the testing effect. Am J Psychol 1996;109:451–64.
13. Karpicke JD, Roediger HL. Repeated retrieval during learning is the key to enhancing later retention. J Mem Lang 2007;57:151–62.
14. Carpenter SK, DeLosh EL. Impoverished cue support enhances subsequent retention: support for the elaborative retrieval explanation of the testing effect. Mem Cognit 2006;34:268–76.
15. van der Vleuten CPM, Schuwirth WT, Lambert. Assessing professional competence: from methods to programmes. Med Educ 2005;39:309–17.
16. Swanson DB, Norman GR, Linn RL. Performance-based assessment: lessons from the health professions. Educ Res 1995;24(5):5–11.
17. Norcini JJ, Swanson DB, Grosso LJ, Webster GD. Reliability validity and efficiency of multiple choice question and patient management problem item formats in assessment of clinical competence. Med Educ 1985;19:238–47.
18. Rethans JJ, Norcini JJ, Baron-Maldonado M, Blackmore D, Jolly BC, LaDuca T, et al. The relationship between competence and performance: implications for assessing practice performance. Med Educ. 1985;36(10): 901–9.
19. Smith SR, Dollase RH, Boss JA. Assessing students' performance in a competency-based curriculum. Acad Med 2003;78:97–107.
20. Williams, RG, Klamen DA, McGaghie WC. Cognitive, social and environmental sources of bias in clinical performance ratings. Teach Learn Med 2003;15:270–92.
21. Miller GE. The assessment of clinical skills/competence/ performance. Acad Med 1990;65(9):63–7.
22. van der Vleuten CPM, Norman GR, de Graaff E. Pitfalls in the pursuit of objectivity: issues of reliability. Med Educ 1991;25:110–8.
23. Epstein RM. Comprehensive Assessment Manual. New York: Rochester; 2001.
24. Crossley J, Humphris G, Jolly B. Assessing health professionals. Med Educ 2002;36:800–4.
25. Epstein RM, Hundert EM. Defining and assessing professional competence. JAMA 2002;287:226–35.
26. Turnbull J, van Barneveld C. Assessment of clinical performance: In-training evaluation. In: Norman GR, van der Vleuten CPM, Newble DI, eds. International Handbook of Research in Medical Education. Dordrecht, Netherlands: Kluwer Academic Publishers; 2002. p. 793–810.
27. Issenberg SB, McGaghie WC, Petrusa ER, Gordon DL, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005;27:10–28.
28. Rosenbaum DA, Carlson RA, Gilmore RO. Acquisition of intellectual and perceptual motor skills. Annu Rev Psychol 2001;52:453–70.
29. Wass V, Jones R, van der Vleuten C. Standardized or real patients to test clinical competence? The long case revisited. Med Educ 2001;35:321–5.
30. Bangert-Drowns RL, Kulik J. Effects of frequent classroom testing. J Educ Res 1991;85(2):89–99.
31. Kromann BC, Jensen ML, Ringsted C. The effect of testing on skills learning. Med Educ 2009;43:21–7.
32. Schoonheim-Klein ME, Habets L, Aartman IHA, van der Vleuten CP, Hoogstraten J, van der Velden, U. Implementing an objective structured clinical examination (OSCE) in dental education: effects on students' learning strategies. Eur J Dent Educ 2006;10:226–35.
33. Coles C. Developing professional judgment. J Contin Educ Health Prof. 2002;22(1):3–10.
34. Issenberg SB, McGaghie WC, Hart IR, et al. Simulation technology for health care professional skills training and assessment. JAMA 1999;282:861–6.
35. Roediger HL, Karpicke JD. The power of testing memory: basic research and implications for educational practice. Perspect Psychol Sci 2006;1:181–276.

Address for Correspondence:
Dr. Junaid Sarfraz Khan, Controller of Examinations, University of Health Sciences, Khayaban-e-Jamia Punjab, Lahore.54600. **Tel:** +92-42-99231218, **Fax:** +92-42-99231857
Email: junaidsarfraz@hotmail.com