Assessment of Learning Outcomes

18

Mohammed Yahya Alnaami,
Hamza Mohammad Abdulghani, Salwa Elsobkey,
and Hazar Yacoub

18.1 Introduction

Recently, health professional curricula have emphasized competencies that are pertinent to each country based on population nature and needs. Hence, such local competencies (learning outcomes) were published as the Saudi Meds [1]. The Saudi Meds is a national competence framework that has been developed by medical schools in the Kingdom of Saudi Arabia. The framework has seven domains: (1) approach to daily practice, (2) doctor and patient relation, (3) doctor and community, (4) communication skills, (5) professionalism, (6) doctor and information technology, and (7) doctor and research. The framework will guide curriculum development and assessment existing in all health professional education to ensure its adaptation to changing needs. The following sections address assessment and evaluation of the aforementioned domains.

M. Y. Alnaami (✉)
Department of Surgery, College of Medicine, King Saud University, Riyadh, Saudi Arabia

H. M. Abdulghani
Department of Medical Education, College of Medicine, King Saud University, Riyadh, Saudi Arabia

S. Elsobkey
College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

H. Yacoub
College of Pharmacy, King Saud University, Riyadh, Saudi Arabia

18.2 Strategic Goal 10: To Develop a Comprehensive Approach to Students' Assessment that Addresses All Educational Domains Including Knowledge, Skills, and Attitudes/Values

This goal, along with other goals in this manual, was developed through several meetings and workshops shared by an elite group of faculty educators and students representing almost all HSCs at KSU. Assessment of learning outcomes encompasses several issues including understanding of the principals of assessment, appropriate use of assessment methods and tools, and the comprehensive approach of assessment that covers the full range of educational domains. Effective assessment must consider the psychometric properties of the examination, i.e., to be valid, reliable, and feasible and to have a measurable impact on learning through quality indicators. There are several assessment methods and tools to measure learning outcomes. However, each tool is appropriate for the context to be measured. Therefore, no one method is appropriate for all domains of learning outcomes. Assessment methods include written examinations, practical and clinical, observational, portfolios, peer assessment, and self-assessment. Learning outcomes include knowledge, skills, attitudes, and values. These learning outcomes differ from one institute to another and from an environment to another depending on community needs [2]. The Miller's Pyramid explains the learning outcome domains and specific tools and methods utilized to assess those domains (Fig. 18.1). The following initiatives are proposed to achieve this goal:
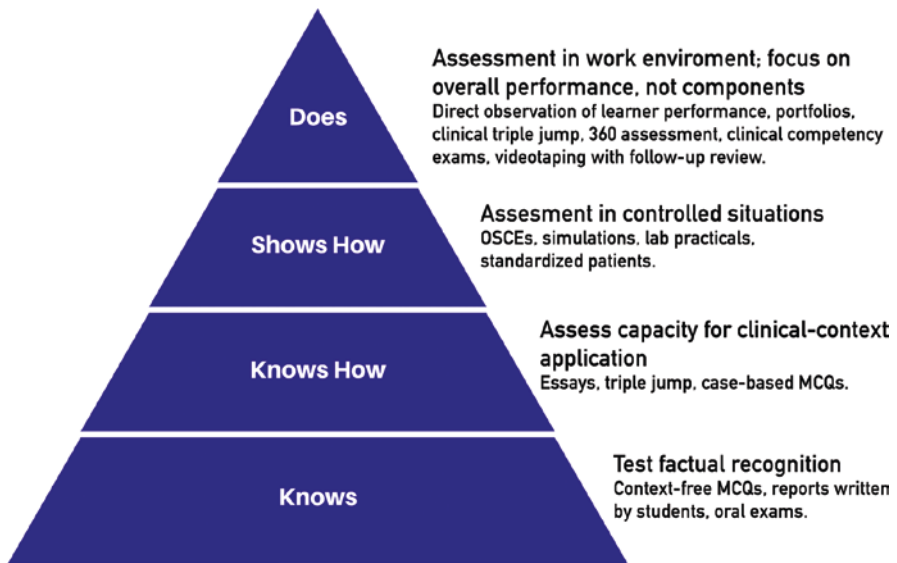

Fig. 18.1 Miller's pyramid to assess clinical competence

18.2.1 Objective (Initiative) 10.1: To Develop Comprehensive Assessment Approaches for Courses that Address All Learning Domains

This initiative is the core issue of assessment that addresses a product expressed as students' learning outcome. Most health professions educational programs adopt several methods that assess students' learning outcomes of their knowledge, skills, and attitudes/values. However, they vary in type and number of assessment methods used depending on staff experience in assessment methods, presence or absence of health professions educational departments or centers, and program accreditation requirements by higher authorities. Ideally, assessment methods should address the achievement of all learning objectives in a valid, reliable, and feasible manner with an impact on the learner and the educational program [3, 4]. *Validity* of an assessment tool is the degree to which the tool is measuring what it is supposed to measure. In other words, it assesses the validity of scores rather than the instrument itself. Validity can be further broken into three sub-types. The first type is *content validity*, which reflects the degree of sampling from different learning domains of the subject to be assessed. This is the most important type of validity, which could be achieved by appropriate assessment blueprint construction [5]. The second type is *criterion validity*, which compares test scores against a criterion or gold standard. The third type is *construct validity*, which is the ability of an instrument to measure what it purports to measure using additional information that supports this notion. *Reliability* refers to consistency, reproducibility, or stability of test scores upon repetition. When a group of experts agree or become in close agreement about an examinee is called *inter-rater reliability*. Therefore, reliability can be measured by test–re-test, equivalent forms, split-half, and item-to-total scores comparison *(internal consistency)*. Usually, reliability is measured by commercially available software, especially for multiple choice questions tests. *A reliability coefficient* value above 0.7 on regular exams or 0.8 and higher for high stakes exams is considered reliable. For a test to be valid, it has to be reliable as well, but not necessary the opposite (Fig. 18.2).

Feasibility refers to the practicality of the assessment method with regard to available resources and expertise and costs. The impact of assessment on the learner and educational programs varies from one system of assessment to another. Since assessment drives learning, learners will try to pass exams the way they are designed for, e.g., memorization, last moment studying, review of previous exams, and guessing or even cheating. In order to have positive impacts on learners, educators should ensure validity of the assessment content, the way it is conducted, what is asked (information is given), and the time and frequency of continuous assessment sessions. Once these concepts of psychometric properties for testing are understood, it has to be applied to all types of testing methods to ensure best assessment and learning outcomes. Learning domains include knowledge, skills, and attitudes/values.
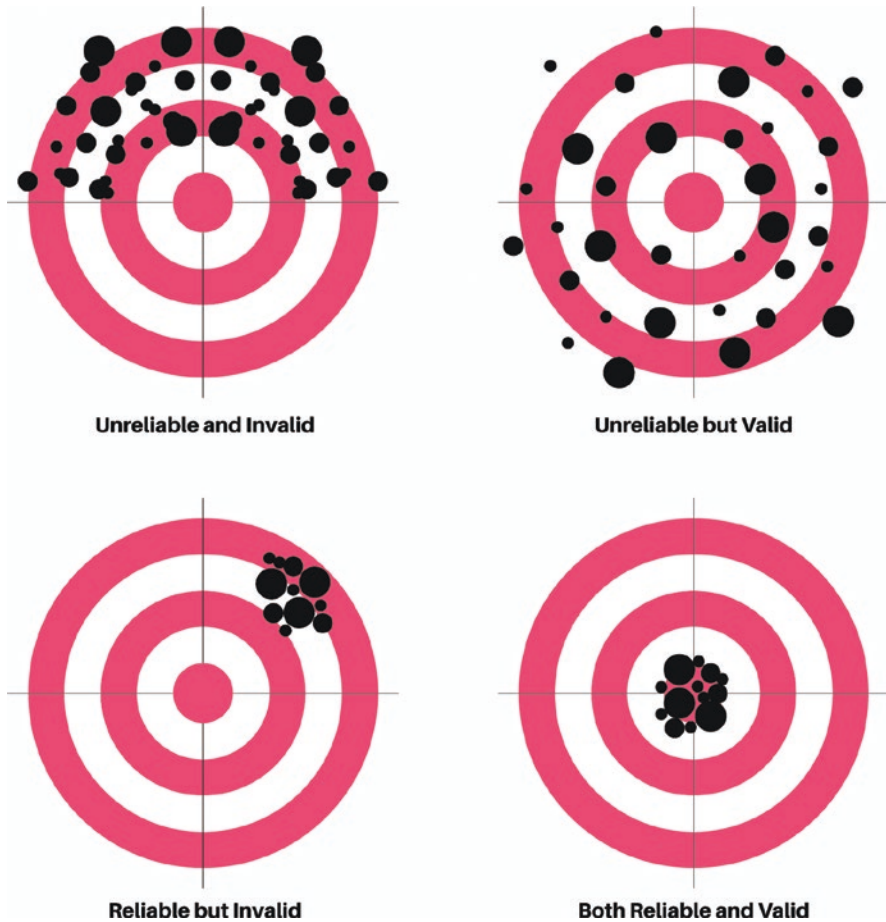
Unreliable and Invalid

Unreliable but Valid

Reliable but Invalid

Both Reliable and Valid

Fig. 18.2 Reliability and validity target

Details of this initiative are summarized in Table 18.1. Reports of the needs assessment from Goal 9, initiative 9.3 "*Develop a comprehensive approach looking at the entire students' process (input–process–outcomes) and document issues arising during studies*" are reviewed by the Assessment Steering Committee and areas for improvement in assessment systems are highlighted. Assessment systems are then benchmarked with best practices in assessment approaches in order to develop a comprehensive handbook for assessment approaches through several workshops and final workshop inclusive of all HSCs and stakeholders. The handbook assessment guide must be aligned with the NCAAA and similar international assessment guidelines. Once the final draft of the handbook is reviewed and approved, it can be implemented and monitored with supporting research studies to validate it.

Table 18.1 Strategic plan for developing comprehensive assessment approaches for courses

Goal 10: Developing a comprehensive approach to students' assessment that addresses all educational domains including knowledge, skills, and attitudes/values

Objective (10.1): To develop comprehensive assessment approaches for courses that address all learning domains

Initiative (10.1)	Responsible	Accountable	Partners
Developing comprehensive assessment approaches for courses that address all learning domains	Assessment Steering Committee and Assessment Units in HSCs	VRHS, Leadership Committee, and Deans of HSCs	Deanship of Development and Quality, Vice-Deans of Academic Affairs in HSCs, and Vice-Deans of Development and Quality

Initiative description

Reviewing report of priorities for improvement, benchmarking best practices in assessment approaches, and developing workshop and handbook for assessment approaches

Requirements and interdependencies	Stakeholders
1. Workshop arrangements (place, agenda, and moderators) and members invited to the workshop are: (a) Members of the teaching, learning, and assessment unit from each HSC (b) Vice-dean for academic affairs from each HSC (c) Vice-dean for quality and development from each HSC (d) Senior students\interns (e) Representatives from faculty members (f) Alumni representatives 2. Letters for facilitation and cooperation with HSCs 3. Survey forms 4. Data collectors and data entry personnel 5. Statistician 6. Computer with internet access, printer, copier, A4 paper, the SPSS program	Faculty staff and undergraduate and postgraduate students

Action plan	Estimated time
1. Reviewing the report of 9.1.3 2. Benchmarking best practices in assessment approach for HSCs courses 3. Developing a draft of Comprehensive Assessment Approaches Handbook for HSCs courses, considering national and international accreditation requirements, and the matrix should combine NCAAA learning domains with professional bodies domains 4. Organizing a workshop for all stakeholders to review the drafted handbook in assessment approaches for HSCs courses 5. Distributing drafted handbooks for comprehensive assessment approaches for HSCs courses to all deans for internal review and feedback 6. Getting the approval of handbook for assessment approaches from the vice-rector for educational affairs	To be decided later as phase II

Table 18.1 (continued)

Goal 10: Developing a comprehensive approach to students' assessment that addresses all educational domains including knowledge, skills, and attitudes/values	
KPIs	**Estimated budget**
1. Reviewing at least five top international assessment centers in HSE for benchmarking, e.g., Dundee UK, Singapore, Harvard USA, and Maastricht	To be studied and decided later as phase II
2. Attendance of the workshop should be >80% of nominated members	
3. >75% satisfaction of the effectiveness of the workshop	
4. Developing a handbook for comprehensive assessment approaches for HSCs courses	
5. Collecting internal feedback from each HSCs on the drafted handbook	

HSCs health sciences colleges, *VRHS* vice-rector for health specialties, *SPSS* statistical package for the social sciences, *NCAAA* National Commission for Academic Accreditation and Assessment, *HSE* health sciences education, *KPIs* key performance indicators

18.2.2 Objective (Initiative) 10.2: To Develop Guidelines for Comprehensive Assessment Across All Learning Domains

This initiative (Table 18.2) discusses how to implement the assessment handbook and guidelines by all HSCs. This requires tremendous efforts by responsible parties and partners of this initiative to train faculty and administration of HSCs in every step and detail of the assessment handbook in order to be ready for successful implementation. Once the approved guidelines manual is ready, it will be published by KSU press and distributed to all HSCs teaching and learning (medical education) units/departments as a reference. At this stage, the assessment guidelines manual will be ready to be implemented at all levels of health professional education. Before implementation, however, knowledgeable and experienced educators at all HSCs will start making a strategic plan to implement the assessment guidelines manual at all HSCs through a one-day workshop led by the assessment steering committee at VRHSs. Then, each T&L unit/medical education department at the corresponding HSC will conduct a half-day seminar introducing the assessment manual to all faculty, students, and administrative representatives who are actively involved in courses coordination and students' assessment. When the assessment manual needs to be implemented, the assessment, teaching, and learning unit (the medical education department) in each HSC needs to conduct short workshops for the various assessment tools along with course coordinators, faculty, students, and administrative representatives in each department. Monitoring and measuring the effectiveness of the implementation process is also of paramount importance to assure best learning outcomes. Assessment of the assessment methods is another issue that deserves more attention from educators and statisticians to analyze exam

Table 18.2 Strategic plan for developing guidelines for comprehensive assessment across all learning domains

Goal 10: Developing a comprehensive approach to students' assessment that addresses all educational domains including knowledge, skills, and attitudes/values			
Objective (10.2): To develop guidelines for comprehensive assessment across all learning domains			
Initiative (10.2) Providing HSCs with the handbook of comprehensive assessment approaches for implementation	**Responsible** Assessment Steering Committee and Assessment Units in HSCs	**Accountable** VRHS, Leadership Committee, and Deans of HSCs	**Partners** CELT, IT Deanship, and IT units in HSCs
Initiative description Providing HSCs with the handbook of the comprehensive assessment approaches for implementation			
Requirements and interdependencies 1. The HSCs comprehensive assessment approaches handbook that was developed and approved in initiative 10.3.1 2. Cover letter from the MEU/D of the HSCs to the colleges' deans. 3. Secretary 4. Computer with internet access, a printer, a copier, A4 paper, and a cartilage 5. An allocated assessment steering group to conduct the orientation workshops 6. An allocated group to moderate the orientation workshops 7. Proper time, place, audiovisual facilities, and catering 8. Rewards for trainers of the workshops 9. Budget for conducting workshops 10. Workshops equipment/materials 11. Feedback surveys to measure the satisfaction of workshops' participants 12. Internal and/or external experts to carry on the requested training workshops 13. An allocated group to moderate the requested training workshops 14. Proper time, place, audiovisual facilities, and catering for the requested training workshops 15. Rewards for the requested training workshops experts 16. Budget for the requested training and conducting workshops 17. Feedback survey to measure the effectiveness of implementation of comprehensive assessment approaches handbook on improving assessment strategies 18. An allocated assessment steering group to handle the survey 19. Data entry and statistician 20. SPSS program 21. Allocated assessment steering group to report on the statistical results 22. Independent opinion of expert on the report 23. Budget	**Stakeholders** Faculty staff, students, and HSCs		

Table 18.2 (continued)

Goal 10: Developing a comprehensive approach to students' assessment that addresses all educational domains including knowledge, skills, and attitudes/values	
Action plan	**Estimated time**
10.2.1. Providing HSCs with the approved handbook of comprehensive assessment approaches	Phase II
10.2.2. Conducting workshops for the HSCs representatives for orientation about the Comprehensive Assessment Approaches Handbook	
10.2.3. Facilitating and supporting training workshops (as needed) for HSCs programs to enhance and support the implementation of the Comprehensive Assessment Approaches Handbook	
10.2.4. Measuring the effectiveness of implementing the Comprehensive Assessment Approaches Handbook on improving assessment strategies	
KPIs	**Estimated budget**
1. At least 75% of HSCs will receive a copy of the comprehensive assessment approaches handbook	Phase II
2. At least 75% of HSCs representatives attend orientation workshops	
3. HSCs representatives attended orientation workshops show 75% overall satisfaction of organized workshops	
4. At least 75% of HSCs programs members attend different training workshops	
5. HSCs members attended training workshops show 75% overall satisfaction of the organized workshops	
6. At least 75% of the HSCs members respond to the survey	
7. Statistical analysis of the survey shows overall 75% effectiveness of implementing the Comprehensive Assessment Approaches Handbook	

HSCs health sciences colleges, *VRHS* vice-rector for health specialties, *CELT* Center of Excellence in Learning and Teaching, *IT* information technology, *MEU/D* medical education units/departments, *SSPS* statistical package for the social sciences, *KPIs* key performance indicators

results and assure their reliability and validity. This will improve examination items writing, conduct, and results. Moreover, it will assure justice and equity among students who always look for evidence of their performance in examinations. Most health science programs have assessment centers where assessment training and workshops are conducted and examinations are revised before their conduct. In addition, assessment centers can monitor and supervise examinations, analyze and assess examination results, give feedback reports to various departments, and publish research on assessment and learning outcomes. The process of implementation may take one full academic year time depending on the degree of authority and support of the VRHSs, knowledge and skills of the personnel involved, and enthusiasm and cooperation of faculty. An estimated budget of 26,000 USD will be needed for the initiative resources, personnel incentives, and rewards for those who cooperate and compete for excellence. The KPIs of this initiative indicate that not all HSCs will be ready for the implementation of the assessment manual. Therefore, about (75%) of the HSCs are expected to participate in this initiative.

18.3   Discussion

Health professional education has undergone dramatic changes over the last century through four stages of innovation. In a significant shift from the scientific approach to health professional education commonplace in Europe at the end of the nineteenth century, the Flexner Era at the beginning of the twentieth century [6] was noted for the idea of teaching basic sciences as the basis of clinical sciences and practice. In the 1970s, problem-based learning was strongly promoted in an attempt to integrate basic, clinical, and social sciences through the use of problem scenarios [7]. Competency/outcome-based curricula became popular around the turn of this century. Learning outcomes vary from country to another and from an institute to another depending on societal and political needs [8]. In North America, broad learning objectives and learning outcomes were recommended by the American Association of Medical Colleges (AAMC) and the Accreditation Council for Graduate Medical Education (ACGME), respectively [9, 10]. Canadians also developed their own (CanMEDS) competencies [11]. The WHO—International Institute of Medical Education has produced a consensus of learning outcomes as minimal essential requirements for medical school graduate [12]. The Scottish Deans Medical Curriculum Group [13] adopted a framework of outcomes based on a three-circle model: what the doctor is able to do, his/her approaches to practice, and professional attributes (Table 18.3).

Table 18.3 Recommended assessment methods for the 12 learning outcomes in order of importance

What the doctor is able to do	
Learning outcomes	Assessment methods
1. Clinical skills OSCE	Observation; logbooks; mini-CEX
2. Practical procedures	OSPE; portfolios and logbooks; observation; DOPS
3. Patient investigations	Written examinations OSCE; observation; portfolios
4. Patient management	Written examinations OSCE; observation; portfolios
5. Health promotion and disease prevention	OSCE; portfolios and logbooks; observation; and written examinations
6. Communication	OSCEs; observation; peer/self-assessment; portfolios
7. Information management	Skills; portfolios; OSCE; observation; written exam
How doctors approach their practice	
1. Principles of social, basic, and clinical sciences	Written examinations; portfolios; observation; and OSCE
2. Attitudes, ethics, and legal responsibilities	Observation; portfolio; OSCE; peer/self-assessment; and written examinations
3. Decision making; clinical reasoning and judgment	Portfolio; observation; written assessment; OSCE; and peer/self-assessment
Doctors as professionals	
1. Role as a Professional	Observation; peer/self-assessment; OSCE; written
2. Personal development	Portfolio; observation; peer/self-assessment; OSCE; and written assessment

In the Kingdom of Saudi Arabia, the Executive Committee for SaudiMED Framework developed six learning domains or themes with seventeen learning outcomes which are adopted by all medical schools: scientific approach to practice; patient care; community-oriented practice; communication and collaboration; professionalism; and research and scholarship [14]. Assessment of such learning competencies/outcomes worldwide, however, did not develop well along with these innovations in curricular development [10]. Assessment of learning outcomes in fact encompasses several issues including an understanding of the principals of assessment, appropriate use of assessment methods and tools against the desired competency/outcome, and the comprehensive approach of assessment that covers the full range of educational domains. Effective assessment must consider the psychometric properties of the examination that is to be valid, reliable, feasible and has a measurable impact on learning outcomes through quality indicators. These metrics are important measures that guard appropriateness and quality of examination methods; otherwise, examinations will be of low quality and products are usually weak. Low performance in common placement testing such as the progressive testing [15] may indicate indirectly poor performance at health professional schools and/or low-quality examinations. There are several assessment methods and tools to measure learning outcomes; however, each tool is appropriate for the context to be measured. Therefore, no one method is appropriate for all domains of learning outcomes. Investment in good assessment is also an investment in teaching and learning [16]. Shumway and Harden [2] summarized the assessment tools against each assessment category in "AMEE's Assessment Guide No. 25" (Table 18.4).

Written tests such as Long Essay Questions (LEQs) were a common assessment tool in health profession education at the begging of the nineteenth century. The

Table 18.4 Assessment tools (instruments) against each assessment category

Assessment category	Assessment tools
Written assessments	Essay: Short answer questions; completion questions; multiple choice questions (MCQs); extended matching items (EMIs); modified essay questions (MEQs); script concordance; key features; patient management problems (PMPs); and dissertation report
Practical assessments	Spot examination; objective structured practical examination (OSPE); practical examination
Clinical assessments	Long and short cases examination; objective structured clinical examination (OSCE); objective structured long examination record (OSLER); group objective structured clinical examination (GOSCE); mini clinical evaluation exercise (Mini-CEX), direct observation of practical skills (DOPS), etc.
Direct observation	Tutors report; checklists; rating scales; patient report; reflective and diary
Portfolios and other records	Logbooks; portfolios report; and procedural logs
Peer- and self-assessment	Peer report and self report

LEQs are reliable for in-depth assessment of a knowledge segment (e.g., Describe the process of fat digestion and absorption in the gut?); however, they are not content valid tool that can explore the knowledge domain of the gastrointestinal tract, for example. Therefore, long essays might be appropriate assessment tool for in-depth knowledge. LEQs are very easy to construct but time consuming to correct, and teachers will lose concentration and interest while reading many texts, which may affect concentration and compromise fairness in grading. This, of course, will affect its practicality and validity to some extent. To avoid these disadvantages of long essays, the modified essay questions (MEQs), completion questions, and short answer questions have emerged at the middle of the past century as reliable, valid, and practical assessment tools, which have replaced most LEQs in health professions education [17]. Over the last three decades or so, there has been a general move to MCQs over all types of essay questions as objective, reliable, content valid, practical, easy to administer/share/correct and analyze assessment tools with good impact on learning outcome in health professions education [2]. Not only that, MCQs nowadays are widely used in admissions, progress testing, promotion from one level to a higher level, licensing, and in high stakes postgraduate board examinations. MCQs, however, cannot assess in-depth knowledge like essays, difficult and costly to construct, and have some cuing and guessing effects. Patient management problems (PMPs) and Extended Matching Items (EMIs) are not popular assessment tools nowadays as they used to be at the end of the past century because of difficulties with question setting, marking, and standardization.

Practical and clinical examinations are very important tools to assess practical and clinical skills and attitude domains of the clinical practice. Practical assessment includes spot examination and practical observation. These are easy to construct and administer, but lack content validity, i.e., they cannot sample enough from the skills and attitudinal domains. In order to solve this problem, the Objective Structured Practical Examination (OSPE) came to improve this issue by increasing the number of practical encounters through multiple stations and by standardizing answer checklists for all stations. Similarly, the clinical assessment tools are used to include long and short cases, which lack content validity and fairness of distribution among students (the luck of the draw!). Therefore, Objective Structured Clinical Examination (OSCE), Objective Structured Long Examination Record (OSLER), and Group Objective Structured Clinical Examination (GOSCE) have emerged as more valid and reliable tools to solve these issues and drawbacks. The OSCE, however, gained popularity over other methods during the last three decades as a reliable, valid, and feasible tool to assess clinical competence [18]. For a reliable and valid OSCE, a minimum of 20 stations are required, with the use of checklist and standardized patients (SPs) [10, 19]. SPs need to be well trained to portray real patients' role in order to increase OSCEs' validity [16, 20]. Feasibility of OSCEs varies from one academy to another depending on available resources, SPs, and experienced educators. Also, cost of OSCEs varies from center to center [21]. These costs increase with recruitment and training of SPs, training of examiners, and maintenance of exam security. Positive impacts of OSCEs on students include increased learning, satisfaction due to fairness in evaluation, and increases of their

experience for future OSCEs. However, OSCE has some drawbacks including fragmented learning, no time for in-depth assessment, and students know most OSCE stations beforehand. However, lots of modification in OSCE constructions have modified these drawbacks.

Other forms of learning outcome assessment include direct observation during attachment (global rating), reviewing written reports (portfolios), logbooks, and self/peer/360° feedback reports. These are best used for communication, interpersonal, and other attitudinal skills. Reliability of these assessment tools increases if done by a committee of expert faculty/examiners and decreases if done by a biased faculty. For positive impacts on learning, these forms of assessment are best used for formative feedback, improving communication and interpersonal skills, and must be revealed to the student as early as possible. Negative impacts on learning happen when students are informed late and/or if done by inexperienced or biased faculty. More research to assess the validity and reliability of these forms of assessment is needed to encourage educators and faculty to use them more frequently. Another important area in any assessment system is the practice of post-assessment test items analysis, which gives the function of each tool and gives valid and reliable results.

18.4 Summary

To develop a comprehensive approach to students' assessment requires faculty's knowledge and skills on assessment principles, use of appropriate assessment method(s) that matches the appropriate domain (i.e. knowledge, skills, attitude), analysis of the results, and their interpretation. The strategy to achieve this goal involves two initiatives. First, to develop comprehensive assessment approaches for courses that address all learning domains. Second, to develop guidelines for these domains. The strategic details for each initiative, the recommended assessment methods, and tools were outlined. The estimated time needed to complete each initiative and its budgetary details depends on studies, meetings, and discussions by relevant stakeholders involved during the implementation process.

References

1. Zaini RG, Bin Abdulrahman KA, Al-Khotani AA, Al-Hayani AM, Al-Alwan IA, Jastaniah SD. Saudi Meds: a competence specification for Saudi medical graduates. Med Teach. 2011;33:582–4.
2. Shumway JM, Harden RM. AMEE education guide No 25: the assessment of learning outcomes for the competent and reflective physician. Med Teach. 2003;25(6):569–84.
3. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. Med Teach. 2011;33(6):478–85.
4. Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Epidemiol Serv Saude. 2017;26(3):649–59.

5. National Board of Medical Examiners. Test blueprinting II: creating a test blueprint. 2019. https://www.nbme.org/sites/default/files/2020-01/Test-Blueprinting-Lesson-2.pdf.

6. Flexner A. Medical education in the United States and Canada: a report to the Carnegie Foundation for the Advancement of Teaching. New York: The Carnegie Foundation for the Advancement of Teaching; 2003.

7. Schmidt HG. Problem-based learning: rationale and description. Med Educ. 1983;17:11–6.

8. Alderson A, Martin M. Outcomes based education: where has it come from and where is it going? Issues Educ Res. 2007;17(2):161–82.

9. AAMC. Report I: learning objectives for medical student education. Medical Schools Objectives Project. AAMC Washington; 1998.

10. ACGME & ABMS. Toolbox of Assessment Methods. A product of the Joint Initiative. Accreditation Council for Graduate Medical Education (ACGME) Outcomes Project. American Board of Medical Specialties (ABMS). Version 1.1. September 2000. Accreditation Council for Graduate Medical Education and American Board of Medical Specialties, 2000.

11. CanMEDS. Extract from the CanMEDS 2000 Project Societal Needs Working Group Report. Med Teach. 2000;22(6):549–54.

12. IIME Core Committee. Global minimum essential requirements in medical education. Med Teach. 2002;24(2):117–20.

13. Harden RM, Crosby JR, Davis MH. An introduction to outcome-based education. Med Teach. 1999;21(2):7–14.

14. SaudiMED Framework. 2016. https://etec.gov.sa/en/productsandservices/NCAAA/AccreditationProgrammatic/Pages/Medical-Colleges.aspx.

15. René A, Schutte B, Meiboom A, Greidanus J, Dubois E, Bremers A. The progress test of medicine: the Dutch experience. Perspect Med Educ. 2016;5(1):51–5.

16. van der Vleuten CPM, Swanson DB. Assessment of clinical skills with standardized patients: state of the art. Teach Learn Med. 1990;2(2):58–76.

17. Feletti GI. Reliability and validity studies on modified essay questions. J Med Educ. 1980;55(11):933–41.

18. Norman G. Research in medical education: three decades of progress. Br Med J. 2002;342:1560–2.

19. Cushing A. Assessment of non-cognitive factors. 2002. https://doi.org/10.1007/978-94-010-0462-6_27.

20. Tamblyn R, Klass DJ, Schnabl GK, Kopelow M. Sources of unreliability and bias in standardized-patient rating. Teach Learn Med. 1991;3(2):74–85.

21. Carraccio C, Englander R. The objective structured clinical examination: a step in the direction of competency-based evaluation. Arch Pediatr Adolesc Med. 2000;154(7):736–41.