Review Article
http://mjiri.iums.ac.ir

Medical Journal of the Islamic Republic of Iran (MJIRI)
Iran University of Medical Sciences

Clinical reasoning assessment through medical expertise theories: past, present and future directions

Elham Boushehri[1], Kamran Soltani Arabshahi[2], Alireza Monajemi*[3]

Received: 4 June 2014 Accepted: 20 October 2014 Published: 15 June 2015

Abstract

Exploration into the concept of "medical expert" dates back to more than 50 years ago, yet yielding three leading theories in the area of clinical reasoning, namely, knowledge structure, hypothetic-deductive, and dual process. Each theory defines "medical expert" in a dissimilar way. Therefore, the methods of assessment through which the experts are identified have been changed during the time. In this paper, we tried to categorize and introduce some widely used tests for identification of experts within the framework of existing main theories. Implementation of the proposed categorization for providing future assessment tools is discussed.

Keywords: Clinical reasoning, Assessment, Analytical approach, Non-analytical approach, Experts.

Cite this article as: Boushehri E, Soltani Arabshahi K, Monajemi A. Clinical reasoning assessment through medical expertise theories: past, present and future directions. *Med J Islam Repub Iran* 2015 (15 June). Vol. 29:222.

Introduction

Clinical reasoning is an essential competency for every physician in their everyday practice. The attempt to clarify this concept of which longed for more than 50 years begun with this question: who is a medical expert?

To answer this question, we introduce three leading theories introduced in the area of clinical reasoning, known as: knowledge structure, hypothetic- deductive, and dual process (1-4). Being or becoming an expert are defined differently in each theory. During the course of time, methods of assessment for identification of experts were changed, and in each period various clinical reasoning tests were suggested to distinguish experts from novices, all inspired by the expert concept dominated in that time.

Providing new clinical reasoning tests has adopted various new criteria regardless of underpinning theory of expertise. While the main purpose of clinical reasoning tests is discriminating expert from novice, the definition of expert provided by the existing theories must be also taken into account. Giving a precise definition of medical expert has always been a major concern for clinical reasoning researchers.

In this paper, we tried to categorize and introduce the most applied clinical reasoning tests within the framework of the main existing theories. This categorization was performed according to the similarity and distinction criteria for expert and novice within the existing theories. This categorization could help researchers recognize tests with more conformity with the concept of expertise.

It is also supposed that this study will provide educators with the opportunity to expand contemporary clinical reasoning assessment tools, apply tests in their proper situation, and to raise considerations for

[1]. PhD candidate of Medical Education at Iran University of Medical Science, Faculty of Health, Hormozgan University of Medical Sciences, Bandar Abbas, Iran. elhamboushehri@yahoo.com
[2]. MD, Professor, Center for Educational Research in Medical Sciences (CERMS), Iran University of Medical Science, Tehran, Iran. soltarab34@gmail.com
[3]. (**Corresponding author**) MD, PhD, Assistant Professor, Department of Philosophy of Science, Institute for Humanities and Cultural Studies, Tehran, Iran. monajemi@med.mui.ac.ir

future research in the field.

Hypothetico-deductive paradigm

Early attempts on assessing clinical reasoning skill was based on the idea that considered "clinical reasoning ability" as an individual trait or special personal characteristics (3,5). It was also believed that experts have a different mental strategies or heuristics to solve problem which is "domain (content)-independent"(4,6). Thus, an expert was distinguished for his/her special mental ability he/she had for reasoning.

Therefore, many attempts were made to create some procedural instruments by which it was possible to know to what extent performance of physicians (as examinee) and experts (as reference group) are matched during the reasoning process. In this regard, multi-stage assessment tests were created. It means that test providers emphasized on the intermediate phases of problem solving while encountering a simulated clinical situation and not merely final diagnosis or management plan (3,5,7).

As such, Rimoldi in1961 advised his method of knowing how a medical student proceeds when posing a clinical situation (8). The tool itself consisted of some two-sided cards. One side of each card contained some questions that a participant may ask of a patient. On the other side, he/she receives related information like X-ray findings, laboratory results, etc. An examiner was responsible for writing the number of selected cards in the same order in which they are selected. Scoring was based on the number of questions asked, their usefulness for the final diagnoses, and the order in which they are asked. It provided some opportunities to study the process decided by a subject when diagnosing a case. Rimoldi created a tool by which it could possibly clear some parts of diagnosis remained uncovered through ordinary true-false or multiple choices tests. He succeeded to make a new approach (maybe qualitative one) to analyze the pattern used by both experts/novices to reach diagnoses (8). The results showed that junior students used fewer questions than seniors and the number of irrelevant and redundant questions asked by juniors was considerably high compared with senior groups (8).

There was also a kind of less known test known as programmed test, used by the National Board of Medical Examiners in the U.S., consisted of three parts: primary clinical information about a patient, some probable diagnoses, and investigations a participant may select to confirm the diagnosis. Participants were scored for the number of correct decisions they made (5,9,10).

In the early 1970, "hypothetic-deductive" general thinking model was derived from cognitive psychology. The model explored the cognitive process of problem solving in experts in two phases: hypothesis generation and hypothesis testing (2,11-13). Hypothetic-deductive model seemed to be the best answer to the question of how experts think. On that idea, some tools like Branched Management Problems, Diagnostic Management Problems and Sequential Management Problems were developed from1965 to 1975. Patient problem Management (PMP) was the improved version of these kinds of tests (9, 14).

PMP was known as the best paper-based simulation method. During some few years, this method became very popular and many improved versions were suggested (15). It was claimed that PMP can offer a practical means of assessing both process and product of clinical reasoning (16). There were booklets containing short case reports and some data on history, diagnoses and therapeutic works. Participants were asked to sequentially match them with the given cases. Some changes to the method like invisible ink, tabs and latent print image made it much more energetic and truthful (17-19). Participants were scored based on the data of correct selections and taken pathways. Except for many problems existed in using the test (to be expensive in time and running, to be difficult having experts consensus on the correct answer and way), there was also low correlations between the

physicians' performance and the problems. This occurred frequently in PMP tests with all kinds of problems and scoring methods (11,14,20,21). The phenomenon was named "content specificity" which contradicted with the beliefs that introduced reasoning as the skill independent from "knowledge structure" (11).

Extended Matching Questions (EMQs) proposed in 1980s are multiple choice questions; each consisted of a problematic clinical case with just one correct answer within a list of at least 7 options. They were assumed as a reasonable alternative to either multiple choice questions or free response tests (23).

EMQ was supposed to have the ability to discriminate expert from novice. It has been used by National Board of Medical Examiners (NBME) and U.S Medical Licensing Examination board (USLME) (23). Scoring of EMQ is based on the number of right diagnoses and the type of reasoning that participants verbally describe through the process. There were some critics who believed the EMQ is highly dependent on student's memory with a little attention to the process of hypothesis testing. However, number of studies that emphasized on its ability to assess problem solving skill is considerable. One of the most important limitations in using EMQ is the increased work for faculty members when adapting to the EMQs tests. There was also low correlation between physicians' performance and sequential problems in an equal domain. In addition, they often are designed for assessing diagnostic reasoning not all types of clinical situations like treatment and following (23-29).

Triple-Jump (TJ) test was an assessment method originated from McMaster curriculum aimed to estimate student's clinical reasoning skill. TJ is a three step examination that begins with a written clinical scenario. It makes student to improve his/her hypothesis. The first step follows by a 2 hour session for self- directed learning based on new issues a student learned from step one. Then, there is a 30 minute session for giving report and feed-back. The first version of the examination was in oral format and graded (pass or fail) subjectively (30,14) by faculties. The next versions were improved by adding a written part to the oral test which allowed the faculties to grade it objectively (31,32,11). TJ examination was an actual method to assess student's competency in the setting of problem based curriculum. It was a kind of assessment method which its proper effects on learning were proved (14). However, the tool had many problems of validity and reliability for standardization (30).

Clinical Reasoning Exercise (CRE) was proposed to problem–based learning (PBL) medical schools as a multiple question test which had more reliability compared to TJ. They were some oral or written tests which could evaluate different aspect of student competency through engaging them in problem solving process. However, it's inter-rater and inter- case correlations, concurrent validity and reliability is yet open to challenge. Some authors also believed that CREs could assess student's knowledge not their clinical reasoning skill (9,31-34).

Clinical Reasoning Problems (CRPs) introduced as the best tool that can purely assess the process of reasoning (35). It has a patient's history and some clinical findings. During the test, examinees are asked to nominate two from the most probable diagnoses among a list of diagnoses suggested for a short scenario. Then, examinees should mark some given findings in the next stage and determine which of them support or reject his/her selected diagnoses respectively.

Researchers have showed the usefulness of CRPs as an accurate indicator of diagnostic reasoning. However there are also some controversial findings on the construct validity of CRP method and its ability to determine the process that leads to the outcomes (9,36).

Hypothetico-deductive paradigm's major emphasis in providing the tests was on a set of reasoning strategies or heuristics through which experts could solve clinical problems

successfully. Thus, related assessment tests were multi stage- largely domain-independent- and were assumed as having the capability of differentiates experts from novices. Elstein et al showed that problem solving heuristics is strongly dependent on acquired knowledge (6). Therefore, content- specificity phenomenon casted a big doubt on definition of experts and the ways they are identified according to hypothetico-deductive paradigm.

Knowledge structure paradigm

Given evidence presented above, it was believed that experts should use hypothetico-deductive model to possibly solve any problem encountered. However, it was found that clinical reasoning skill is specific to the problem encountered as was shown by the low inter-case correlations in clinicians' performance (content specificity) (4). This caused a shift from general to content of knowledge paradigm. Early cognitive psychological research also emphasized on related knowledge as an essential determinant of successful problem solving. Besides, it was expected that experts may possess more knowledge than novices. However, some PMPs results demonstrated that intermediate students' scores of PMPs were upper than those of experienced experts. This fact surprised those who believed experts may have considerably more explicit knowledge than novices (14). Again, cognitive psychology presented an explanation: experts did not necessarily have more "knowledge" but compared to students, they have the ability to organize their specific knowledge more efficiently (2). This changed researchers approach toward "knowledge structure", "expert process" and "knowledge itself". This kind of mental network was known as "script" (4). Scripts are like schema, but have some orderly actions or events which are essential to achieve a goal. Script theory was first introduced to the field of medical education to explain the differences between experts and novices in their practice (37). Also known as semantic networks, illness scripts or instance scripts (38,39), these kinds of networks help experts to remember relevant knowledge easier than students, thus, they solve a clinical problem more successfully. In other words, problem solving appeared to involve the way a person organizes his/her knowledge and experience in the brain (40-42). Thus, attempts started to create an instrument to estimate variation and quality of illness scripts in physicians' mind. According to the script theory, while diagnosis clinicians take cues from relevant hypothesis in their minds. This activated knowledge is then deductively applied to the situation. This psychological view helped investigators to use a new approach in assessing clinical reasoning naming scrip concordance tests (SCTs) (43).

SCTs were designed to evaluate the richness of mental networks and the capacity of data interpretation in decision making situations. It directly evaluated knowledge organization. The test approach consisted of short scenarios followed by some questions categorized in three sections. The first two sections consisted of some diagnostic or managerial decisions. Section three was a five-point Likert scale by which examiners were able to evaluate participants' decisions. The clinical reasoning competency of each examinee was evaluated according to the percent of concordance between examinee decisions and expert panel decisions. SCT let examinees make decisions for to a problematic situation through analysis of available data (43-50). Some studies have showed that the reliability and validity of SCT are good and its reliability directly related to the experience of physicians. However, some research findings on its differential and predictive powers are not considerable (51-53).

Key features (KF) test was suggested in 1987 and implemented in 1992 as one section in MCC Qualifying Examination (MCCQE) (54). KF test was based on the concept of "case specificity". It was a good replacement for PMP test because of the increased reliability due to wider sampling of cases and more attention on evaluating

the important steps (key features) in determination of a problem (55). It seems that the ability of a person to gather critical data during a clinical decision depends on the related script exists in her/his mind. Thus, KF test could be also considered as a tool for evaluating the adequacy (richness) of scripts (56-59).

Another tool to evaluate diagnostic thinking and clinical reasoning is comprehensive integrative puzzle (CIP) (60). CIP is an 'extended matching' crossword puzzle of various disciplinary elements of more than 4 clinical situations.

Its answer sheet is a matrix with rows and columns and examinees are asked to fill its cells through matching all related parts of findings and create a complete picture of a clinical problem. When the puzzle is completed each horizontal row shows a complete medical scenario. Each completed column shows the student's proficiency in analyzing medical history, physical examination, laboratory test results and, sometimes, some related interventions. It is shown that CIP is extensively accepted by students because there is too much fun in matching puzzle components (1,61). One of the main limitations of utilizing CIP is its low reliability due to lack of a complete agreement among experts' panel. Thus, test providers should have scant attention to transparency and enough details of tests.

A combination of Comprehensive Integrative Puzzles and Extended Matching Questions (EMQ) has been proposed to assess most aspects of clinical reasoning, particularly for the computer based assessment approaches (60).

Scenario Formation test (SF) is one of the newest assessment methods in the field of clinical reasoning (61). SF reports the accuracy of all components of an illness script (1,62). An examinee is required to write two separate scenarios according to some given signs and symptoms and determine a diagnosis for each of them. Scenarios are viewed by two expert members and scored according to a standard checklist.

Disagreement points, if any, are discussed to reach final agreement. Correct diagnosis and appropriateness of using signs and symptoms to complete scenario are considered as main criteria in scoring test (1,61).

Therefore, compared to tests with process approach, it seemed that tests with knowledge structure had a better approach toward defining the expert, resulted in a shift from multi-stage, single-question to single-stage multi-question. To prove this claim we have to compare the PMP test with other tests such as scenario and puzzle. The knowledge structure paradigm claims that the contents and richness of every script correspond to a physician's expertise in a way that he/she might qualify for the rich scripts in his/her expertise, while, on the other hand, he/she lacks any scripts or owns weak scripts in clinical situations not appropriate for his/her expertise. In this regard, the results of some studies have shown that experts when are faced with clinical tests unrelated to their expertise acquire lower scores. Moreover, these studies have also revealed that the evaluation scores of experts in areas unrelated to their expertise are higher than those of novices. The question is that, based on what kind of reasoning the experts can manage clinical situations unrelated to their expertise? And why non-knowledge structure approach is unable to distinguish them from novices? Other studies showed that, in managing clinical situations unrelated to their field of expertise, experts use reasoning process approach instead of knowledge structure approach. In other words, the expert performs better than the novice in both areas of process approach and knowledge structure approach. Therefore, despite the fact that the tests of knowledge structure approach significantly corresponded to the characteristics of experts, they were not able to evaluate the expert in situations other than his/her field of expertise. Thus, the mere use of knowledge structure approach tests does not seem to properly evaluate individual's expertise. Therefore, an expert is a person more successful than a novice in using both reasoning approaches; this defi-

nition for expert has not been offered in the mentioned two clinical reasoning approaches.

Some scholars claim that the hypothetico-deductive approach and knowledge structure approach denote analytical reasoning and non-analytical reasoning, respectively. They believe that general model of reasoning could apply to analyze an unfamiliar situation in which there is no related script. In contrast, knowledge structure reasoning is applied in handling every familiar problem in which there is no need to analyze situation (63,64). Therefore, the significance of expertise evaluation was determined based on individual's skills in both analytical and non-analytical areas as it will be discussed in the next part.

Dual process paradigm

Further studies have suggested a more comprehensive model of clinical reasoning in which analytical and non- analytical approaches are combined. The combined model shows that reasoning could proceed in both directions of analytical and non-analytical, while encountering a clinical problem (3). The two domains should be considered interactive instead of two separate parts or lying along a continuum. That is, they keep influencing each other. Effective medical diagnosis and management are affected by analytical and non-analytical thinking modes which work together to handle all types of clinical practice situations (63,64). Thus, greater awareness on the part of combined model has underlined the need to provide examinee with some new clinical reasoning tests that could better evaluate them in situations in need of bi-directional reasoning. The nature of dual process theory which defines an expert as a person with the flexibility to switch between two modes of thinking with high performance in both types, confirms the idea that expert could not be identified by a single test emerged from analytic or non-analytic thinking approach. It means that the methods of clinical reasoning evaluation in future should be able to test both skills of clinical reasoning, i.e., analytical and non-analytical (64).

Conclusion

In this paper, more than 50 years of effort made by researchers to identify clinical experts' characteristics were investigated. The earliest theory of clinical reasoning emphasized on the special mental stages an expert proceeds to solve a clinical problem. It was hypothetico-deductive reasoning through which an expert creates some clinical hypothesis and evaluates them deductively. It led in creating multi-stages, single-question assessment tests like PMP, CRE, TJ and CRP. Using these kinds of assessment tools showed the extent of similarity between the steps a participant and an expert take to solve a clinical problem (1-36).

Knowledge structure reasoning suggested illness script as a cognitive framework to explain how medical knowledge is organized and applied to a clinical situation. It revealed that more experienced physicians have much of the rich illness scripts which enable them to handle many clinical situations in their field of expertise successfully. Therefore, the related tests were created in a way that could investigate the similarity of illness script of participants with those of expert panel.

Accordingly, medical expert is a person whose illness scripts are as rich as those of expert panel. Thus, multi-stages, single-question assessment tools with focus on the cognitive process of an expert were shifted to multi-questions, one- stage assessment tests like SCT, KF, scenario, puzzle with an emphasis over richness of illness scripts an expert has (4,14,37-64). In the newest theory of clinical reasoning, known as the dual process, the expert is defined as a person who has both analytical and non-analytical thinking simultaneously (3). Based on the theory, Clinical reasoning is not a completely analytical or non-analytical mental process. The physicians' reasoning can be intuitive and/or analytical based on their experience and the kind of clinical condition they encounter. In other words, the

clinical reasoning is on a continuum, with one extreme being intuitive thinking and the other analytical thinking (63,64). This understanding of clinical reasoning concept shows that every physician owns both types of reasoning and the extent to which he employs each depends upon the clinical situation he/she encounters. The expert physician is more successful than a novice in using any of these types of reasoning. That is because of the accuracy of decisions he/she makes in the two dimensions of diagnosing and managing the illness, apart from the type of reasoning he/she adopts (63,64). It seems that accepting the concept of expertise in the combined theory of clinical reasoning has two considerable impacts on the tests differentiating expert from novice:

1. Evaluating analytical clinical reasoning skill in experts is as important as the non-analytical clinical reasoning skill in novices.

2. The clinical reasoning tests must be designed in a way that can evaluate both the accuracy of test answers and the individuals' ability in appropriate switching between the two types of analytical and non-analytical reasoning.

One way to accomplish this claim might be designing tests which can evaluate both kinds of reasoning. Accordingly, instead of designing one single test with twofold application, a suggestion is to select a set of tests covering both analytical and non-analytical reasoning. Therefore, each individual's clinical reasoning skill can be evaluated by this set of tests at the same time. This suggestion has been also supported by recently published studies in the field of evaluation which have highlighted the importance of reaching a holistic picture of an expert's skills (65-67). This has been a shift in clinical reasoning assessment from "one instrument for a skill" to "multi-instrument for all dimensions of a skill (3,31,34,68-70). Along the same line, early studies showed that designing a battery of tests may be useful to assess all aspects of clinical reasoning skill. However it has not yet received much attention in the medical education literature except few ones (1, 61,71-73).

Multi instrument approach and dual process both demonstrate that clinical reasoning skill should be evaluated from multiple sources, because neither single instrument approach nor each of mentioned approaches are able to provide an appropriate framework for "multiple biopsy" in evaluation (3). This leads to an approach in which tries to propose a method for evaluating all aspects of clinical reasoning skill rather than developing one single best instrument. However, administering such tests requires an appropriate study to determine which existing test is analytic-induced and which one is not.

It should be noted that the presupposition here is that we agree that the clinical reasoning test format determines whether a test is analytical or non-analytical induced, while in some studies on reflective practice, the type of assignment given to the individuals in a simulated clinical situation can determine the dimension of their reasoning (74,75).

References

1. Monajemi A, Arabshahi KS, Soltani A, Arbabi F, Akbari R, Custers E, et al. A comprehensive test of clinical reasoning for medical students: An Olympiad experience in Iran. Journal of Education and Health Promotion 2012;1:10.

2. Norman G. Research in clinical reasoning: past history and current trends. Medical Education 2005; 39: 418–427

3. Eva KW. What every teacher needs to know about clinical Reasoning. Medical Education 2004; 39: 98–106.

4. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. Academic Medicine 1990; 65(10): 611-21.

5. van der Vleuten C, Newble DI. How can we test clinical reasoning? The Lancet 1995;345(8956): 1032-1034.

6. Elstein Arthur SE, Schwarz A. Clinical problem solving and diagnostic decision making: selective review of the cognitive literature. BMJ 2002; 324(23):729-732.

7. McGuire CH, Babbott D. Simulation technique in the measurement of problem solving skill. Journal

of Educational Measurement 1967;4(1):1-10.
8. Rimoldi HJ. The test of diagnostic skill. Journal of Medical Education 1961;36:73-9.
9. Anderson KJ. Factors affecting the development of undergraduate medical students' clinical reasoning ability. Unpublished PhD Thesis. University of Adelaide. 2006.
10. Anderson K, Peterson R, Tonkin A, Cleary E. The assessment of student reasoning in the context of a clinically oriented PBL program. Medical Teacher 2008;30(8):787-94.
11. Elstein AS, Sprafka SA, Shulman LS. Medical problem solving, an analysis of clinical reasoning. Cambridge: Harvard University Press, 1978.
12. Pelaccia T, Tardif J, Triby E, Charlin B. Analysis of clinical reasoning through a recent and comprehensive approach: the dual-process theory. Medical Education Online 2011,16: 5890.
13. Neufeld VR, Norman GR, Barrows HS, Feightner JW. Clinical problem-solving by medical students: a longitudinal and cross-sectional analysis. Medical Education 1981;15:315–22.
14. van Der Vleuten CPM. The assessment of professional competence: Developments, research and practical implications. Advances in Health Sciences Education 1996;1(1):41-67.
15. McGuire CH, Babbott D. Simulation technique in the measurement of problem solving skills. Journal of Educational Measurement 1967;4:1-10.
16. Williamson JW. Assessing Clinical Judgment. Journal of Medical Education 1965;40(2):180-187.
17. Marshall J. Assessment of problem-solving ability. Medical Education 1977;11(5): 329–334.
18. Goran MJ, Williamson JW, Gonnella JS. The validity of patient management problems, The Journal of Medical Education 1973;(48):171-177.
19. Hubbard JP. Measuring Medical Education (2nd Ed) Philadelphia, PA: Lea and Febiger, 1978.
20. 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 Medical Education1985;19(3):238-47.
21. Norman GR, Feightner JW. A comparison of behavior on simulated patients and patient management problems. Medical Education 1981;15(1): 26–32.
22. Eva KW, Norman GR. Is thinking aloud equivalent to post hoc explaining. Present at the 1998 Research in Medical Education meeting. New Orleans, LA. 1998.
23. Case SM, Swanson DB. Extended-matching items: a practical alternative to free response questions. Teaching and Learning in Medicine 1993; 5(2):107–115.
24. Beullens J, Struyf E, van Damme B. Do extended matching multiple-choice questions measure clinical reasoning? Medical Education 2005; 39(4):410–417.
25. Fenderson BA, Damjanov I, Robeson MR, Veloski JJ, Rubin E. The virtues of extended matching and uncued tests as alternatives to multiple choice questions. Human Pathology 1997;28:526-32.
26. Wass V, van der Vleuten CPM. Shatzer J and Jones R: Assessment in clinical competence. The Lancet 2001, 357:945-9.
27. Bhakta B, Tennant A, Horton M, Lawton G, Andrich D. Using item response theory to explore the psychometric properties. BMC Medical Education 2005; 5:9.
28. Wilson RB, Case SM. Extended matching questions: an alternative to multiple-choice or free-response questions. Journal of veterinary medical education 1993;20(3).
29. Feletti G, Saunders NA, Smith AJ. Comprehensive assessment of final-year medical student performance based on undergraduate program objectives. The Lancet 1983; 322(8340): 34-37.
30. Smith R M. The triple-jump examination as an assessment tool in the problem based medical curriculum at the University of Hawaii. Academic Medicine 1993; 68(5):366-72.
31. Neville AJ, Cunnington J P.W, Norman GR. Development of clinical reasoning exercises in a problem-based curriculum. Advances in Medical Education1997:377-379.
32. Wood TJ, Cunnington J PW, Norman GR. Assessing the measurement properties of a clinical reasoning exercise. Teaching and Learning in Medicine 2000;12(4):196-200.
33. Boshuizen HPA, Vander Vleuten CPM, Schmidt HG, Machiels-Bongaerts M. Measuring knowledge and clinical reasoning skills in a problem based learning curriculum. Medical Education.1997; 31:115–121.
34. Croskerry, P.A universal model for diagnostic reasoning. Academic Medicine 2009; 84:1022–1028.
35. Groves M, Scott I, Alexander H. Assessing clinical reasoning: A method to monitor its development in a PBL curriculum. Medical Teacher 2002; 24(5):507–515.
36. Groves M, O'Rourke P, Alexander H. The clinical reasoning characteristics of diagnostic experts. Medical Teacher 2003; 25(3):308–313.
37. Barrows HS, Feltovich PJ. The clinical reasoning process. Medical Education1987;21:86–91.
38. Charlin B, Tardif J, Boshuizen HPA. Scripts and medical diagnostic knowledge: theory and applications for clinical reasoning instruction and research. Academic Medicine 2000;75:182–190.
39. Charlin B. Boshuizen HPA, Custer EJ, Feltovich PJ. Scripts and clinical reasoning. Medical Education 2007;41:1178–118.
40. Durning SJ, Artino AR, Schuwirth L, Van der Vleuten C. Assumptions to enhance our understanding and assessment of clinical reasoning. Academic Medicine 2013; 88(4).
41. Wiley J. Expertise as mental set: the effects of domain knowledge in creative problem solving.

Memory & Cognition 1998;26(4).716-30.
42. Norman G. The epistemology of clinical reasoning: perspectives from philosophy, psychology, and neuroscience. Academic Medicine 2000; 75(10): 127-133.
43. Charlin B, Brailovsky CA, Leduc C, Blouin D. The diagnostic script questionnaire: a new tool to assess a specific dimension of clinical competence. Advances in Health Sciences Education 1998; 3:51–8.
44. Brazeau-Lamontagne L, Charlin B, Gagnon R, Samson L, van der Vleuten C. Measurement of perception and interpretation skills during radiology training: utility of the script concordance approach. Medical Teacher 2004;26:326–32.
45. Cohen LJ, Fitzgerald SG, Lane S, Boninger ML. Development of the seating and mobility script concordance test for spinal cord injury: obtaining content validity evidence. Assistive Technology 2005; 17:122–32.
46. Khonputsa P, Besinque K, Fisher D, Gong W. Use of script concordance test to assess pharmaceutical diabetic care: a pilot study in Thailand. Medical Teacher 2006;28 (6):570–3.
47. Meterissian S, Zabolotny B, Gagnon R, Charlin B. Is the script concordance test a valid instrument for assessment of intraoperative decision-making skills? American Journal of Surgery 2007;193:248–51.
48. Brownell AK. The script concordance test. Canadian Journal of Neurological Sciences 2009;36:272–3.
49. Charlin B, van der Vleuten C. Standardised assessment of reasoning in contexts of uncertainty: the script concordance approach. Evaluation & the Health Professions 2004;27 (3):304–19.
50. Shbert L, Charlin B, Corcos J, Gagnon R, Grise P, van der Vluten C. Stability of clinical reasoning assessment results with the script concordance test across two different linguistic, cultural and learning environments. Medical Teacher. 2002; 24(5):522–527.
51. Charlin B, Gagnon R, Lubarsky S, Lambert C,Meterissian S, Chalk C, Goudreau J, van der Vleuten CPM. Assessment in the context of uncertainty using the script concordance test: more meaning for scores. Teaching and Learning in Medicine 2010;22 (3):180–6.
52. Lubarsky S, Charlin B, Cook David AC, Chalk C, van der Vleuten CPM. Script concordance testing: a review of published validity evidence. Medical Education 2011; 45: 329–338.
53. Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The script concordance test: A tool to assess the reflective clinician. Teaching and Learning in Medicine 2000; 12:189–95.
54. Bordage G, Page G. An alternate approach to PMPs, the key feature concept. In: Hart I, Harden R, et al. Further developments in assessing clinical competence. montreal: Can-Heal Publications 1987; 57–75.
55. Page G, Bordage B. The medical council of Canada's key features project: a more valid written examination of clinical decision–making skills. Academic Medicine 1995;70(2):104-110.
56. Page G, Bordage G, Allen T. Developing key-feature problems and examinations to assess clinical decision making skills. Academic Medicine 1995; 70:194–201.
57. Bordage G, Brailovsky C, Carretier H, Page G. Content validation of key features on a national examination of clinical decision-making skills. Academic Medicine 1995; 70:276-81.
58. Ali SK, Bordage G. Validity of key features for a family medicine pilot exam at the College of Physicians and Surgeons of Pakistan. Journal of the College of Physicians and Surgeons of Pakistan1995;5(6):256-60.
59. Fischer MR, Kopp V, Holzer M, Ruderich F, Jünger J.A modified electronic key feature examination for undergraduate medical students: validation threats and opportunities. Medical Teacher 2005; 27(5):450-5.
60. Ber R. The CIP (comprehensive integrative puzzle) assessment method. Medical Teacher 2003; 25:171–6.
61. Amini M, Moghadami M, Kojuri J, Abbasi H, ArhamiDoolatAbadi A, et al. An innovative method to assess clinical reasoning skills: clinical reasoning tests in the second national medical science Olympiad in Iran. BMC Research Notes 2011; 4:418.
62. Devlin J, Marquis F, Riker R, Robbins T, Garpestad E, Fong J, Didomenico D, Skrobik Y. Combined didactic and scenario-based education improves the ability of intensive care unit staff to recognize delirium at the bedside. Critical Care 2008; 12(1):19.
63. Norman G. Dual processing and diagnostic errors. Advance in Health Science Education 2009; 14:37–49.
64. Custers E. Medical education and cognitive continuum theory: an alternative perspective on medical problem solving and clinical reasoning. Academic Medicine 2013;88: 1074–1080.
65. McGuire C. Written methods for assessing clinical competence: Hart IR, Harden RM, Eds. Further development in assessing clinical competence. Montreal: Heal Publication 1987:46-58.
66. Newble D, Dawson B, Dauphinee D. Guidelines for the assessment of clinical competence. Teaching and learning in Medicine1994;6:213-20.
67. Coderre SP, Harasym P, Mandin H, Fick G. The impact of two multiple-choice question formats on the problem-solving strategies used by novices and experts. BMC Medical Education 2004;4:23.
68. van der Vleuten C, Schuwirth LW. Assessing professional competence: from methods to programs. Medical Education 2005;39:309–17
69. Goulet F, Jacques A, Gagnon R, Carlin C, Shahab A. Poorly Performing physicians: does the

script concordance test detect bad clinical reasoning? Journal of Continuing Education in the Health Promotion 2010;30(3):161–166,

70. Marshall J. Assessment of problem-solving ability. Medical Education 1977;11:329-34.

71. Monajemi A, Adibi P, Soltani-Arabshahi K, Arbabi F, Akbari R, Custers E, et al. The battery for assessment of clinical reasoning in the Olympiad for medical sciences students. Iranian Journal of Medical Education 2011;10(5):1056-1067.

72. Hadizadeh F, Yazdani Sh, Ferdosi M, Haghdoost A, Rashidian A, Hadadgar A, et al. The first national Olympiad on reasoning and decision making in health system management; an experience report. Iranian Journal of Medical Education 2011;10(5):1018-1032.

73. Adibi P, Hadadgar A, Hadizadeh F, Monajemi A, Eftekhari H, Haghjoo-Javanmard Sh, et al. Implementation of the first medical science Olympiad in Iran: A report. Iranian Journal of Medical Education 2011;10(5):1006-1017.

74. Mamede S, Schmidt H. The structure of reflective practice in medicine Medical Education 2004; 38: 1302–1308.

75. Mamede S, Schmidt H. Correlates of Reflective Practice in Medicine Advances in Health Sciences Education 2005; 10:327–337.