

## Puzzle test: A tool for non-analytical clinical reasoning assessment

Alireza Monajemi<sup>1</sup>, Minoo Yaghmaei<sup>\*2</sup>

Received: 16 January 2016

Accepted: 26 June 2016

Published: 5 November 2016

### Abstract

Most contemporary clinical reasoning tests typically assess non-automatic thinking. Therefore, a test is needed to measure automatic reasoning or pattern recognition, which has been largely neglected in clinical reasoning tests. The Puzzle Test (PT) is dedicated to assess automatic clinical reasoning in routine situations. This test has been introduced first in 2009 by Monajemi et al in the Olympiad for Medical Sciences Students. PT is an item format that has gained acceptance in medical education, but no detailed guidelines exist for this test's format, construction and scoring. In this article, a format is described and the steps to prepare and administer valid and reliable PTs are presented. PT examines a specific clinical reasoning task: Pattern recognition. PT does not replace other clinical reasoning assessment tools. However, it complements them in strategies for assessing comprehensive clinical reasoning.

**Keywords:** Illness Script, Clinical Reasoning, Clinical Reasoning Assessment, Pattern Recognition.

*Cite this article as:* Monajemi A, Yaghmaei M. Puzzle test: A tool for non-analytical clinical reasoning assessment. *Med J Islam Repub Iran* 2016 (5 November). Vol. 30:438.

### Introduction

Contemporary clinical reasoning tests have been designed based on the assumption of single best instrument for each trait. Some researchers have criticized and questioned this approach, as clinical reasoning is multi-faceted trait that needs a battery of test to be comprehensively assessed (1,2). Most contemporary clinical reasoning tests typically assess non-automatic thinking; that is why all of them require providing differential diagnoses or listing consistent /contradicted findings with the hypothesis that it may activate non-automatic or analytical reasoning (3). Because clinical diagnostic thinking is based on rapid and unconscious thinking, its assessment is highly fundamental and crucial (4-7).

Therefore, a test is needed to assess automatic reasoning or pattern recognition, which has been largely neglected in clinical reasoning tests. The Puzzle Test (PT) has been designed to assess automatic clinical

reasoning in routine situations. This test has been introduced first in 2009 by Monajemi et al in the "Olympiad for Medical Sciences Students" (4). This paper in particular is focused on the need for describing rules to prepare and administer valid and reliable PT.

### Theoretical Background

All information that doctors have about diseases is organized in a structure called the illness script, which contains clinical presentation (i.e., consequences), pathophysiological mechanisms (i.e., fault), the risk factors and patients' background information (i.e., enabling condition), and patient management plan (i.e., management). According to this theory, scripts are made up of links between disease, clinical manifestations and management plans (1,3). Medical experts' researches have shown that knowledge structures available to doctors and medical students develop

<sup>1</sup>. Assistant Professor, Philosophy of Science Department, Institute for Humanities and Cultural Studies, Tehran, Iran. [monajemi@ihcs.ac.ir](mailto:monajemi@ihcs.ac.ir).

<sup>2</sup>. **(Corresponding author)** Associate Professor, Medical School, Shahid Beheshti University of Medical Sciences & School of Medical Education Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. [yaghmaei.minoo@gmail.com](mailto:yaghmaei.minoo@gmail.com)

and change through clinical practice (6,7). When facing common problems, expert doctors tend to use a rapid, largely automatic reasoning approach based on matching the patterns of the present case to the scripts (5). In this situation, automatic reasoning (i.e., pattern recognition) will be the rule, and reflective practice will make no difference in terms of diagnostic and management plan accuracy (6).

### Test Format

The format of Extended Matching Item (EMI), a kind of pattern recognition test, has been selected for PT (10,11). This test format is used to assess reasoning in routine situations that frequently take place in daily clinical practice, particularly in primary care settings.

In PT, students are presented with a series of brief case scenarios based on a single chief complaint (e.g., shortness of breath) and must select the most relevant medical history, physical examination, laboratory test results, and sometimes, they should choose related interventions from a menu of options. The test is case-based; and the cases are described as short typical scenarios, which are followed by the patient's information, presented in three or four parts. The first part contains the patient's pre-existing conditions such as past medical history; the second part presents clinical findings such as physical signs; the third part is the result of an imaging study or a laboratory test, and the fourth part is relevant diagnostic or management plans. The examinees should match these parts with the relevant scenario (9,12,13).

Cases, described as short scenarios, present challenging clinical situations in a few sentences, and simple cases represent problems frequently encountered by doctors. In these cases, the patient's data are compatible with only one diagnosis. Uncommon problems, rarely seen by doctors, patients

with comorbidities or a combination of different acute clinical conditions, or atypical presentation of diseases are not suitable cases for PT (6). Expert doctors can provide one single diagnosis because all the necessary data are available for the initial diagnosis, and scenarios are routine in clinical settings. If the experts cannot provide one single diagnosis after reading a scenario, the case is not appropriate. For example, when a doctor visits a 60-year-old man with cough and shortness of breath, it brings several diagnoses into mind, but the scenario should be a 60-year-old heavy smoker man, with the history of productive cough, especially in winters for many years, who complains about exacerbating of cough and dyspnea. The latter scenario easily activates COPD. It should be noted that when assessing undergraduate medical students and general practitioners, the typicality of scenarios have to be checked by subexperts (nonspecialists in the field; for example, a surgeon in internal medicine) or experienced GPs, as experts in each field have a tunnel vision and may ignore the diagnoses out of their field of expertise.

Another clinical reasoning assessment method in EMI format is the comprehensive integrative puzzle (CIP) (11). In this test, items are presented in the format of an extended matrix of rows and columns, in which examinees must insert the correct information. However, major differences exist between CIP and PT that are originated from their different theoretical perspectives. CIP scenarios are presented around a body system like cardiovascular system, whereas in PT they are arranged around a chief complaint. While diagnoses are written in CIP, they are not mentioned in PT. The pathophysiology of the diseases could be a part of exam in CIP. However, pathophysiology is preferred not to be a block in PT, as it does not play a role in automatic reasoning of the experts.

Table 1. The Structure of PT

Question	Part 1	Part 2	Part 3	Part 4
Scenario	Past medical Hx	Physical Examination	Lab data & imaging	Dx or Mx intervention

### Test Construction

The extent to which the test covers the area of clinical reasoning under consideration is called content validity. For instance, in internal medicine, situations are selected with respect to their frequency (e.g., congestive heart failure or chronic obstructive pulmonary disease), their severity (e.g., shock, torn aorta) and to the patients' sex ratio and age. This issue is often addressed using a specification table: All areas of the field are specified and cases are selected to ensure adequate sampling.

Another factor that affects test construction is the level of examinees: Clerkships, interns, residents, or practicing professionals. As the exam consists of a series of brief

case scenarios based on a single chief complaint (i.e., focal symptoms or signs), the selection of both the symptom /sign and the diagnoses of the scenarios is highly affected by the level of examinees. In addition, the contents of some parts of the questions may be changed in accordance to the level. For example, clerks prefer symptoms as a focal point, while signs could be the choice of more advanced examinees. Moreover, physical exams and simple diagnostic investigation could be part of the questions for less advanced medical students, while detailed imaging and paraclinical test results, management and follow up plans are selected as parts of the questions for more advanced groups.

Table 2. A sample of PT  
*Dyspnoea*

Scenario 1	Scenario 2	Scenario 3	Scenario 4
A 45-year-old previously healthy woman presented to the emergency department with a history of 8 hours acute onset dyspnea and chest pain. She had a history of traveling long distances by airplane in the last two weeks and returned yesterday.	22- year-old woman with a positive history of one-week respiratory infection is admitted to the emergency department with dyspnea attack. She had the same attack six months ago.	50- year- old man complains from severe chest pain and dyspnea since last night. He had a history of 40 pack/year smoking but quit smoking after dyspnea 5 years ago.	45- year-old man presented at ER with severe dyspnea and Pink, frothy sputum at 2 am at ER. The History of coronary bypass surgery 3 years ago is positive.
<i>Part A</i>			
1	2	3	4
RR= 24/m PR=140/m BP=110/65mm/Hg Lung clear Pain made worse by breathing	Thin and chacectic RR=16/m PR=70/m BP=90/60 mmHg T=37 C No sputum Expiratory rales	Agitated- restless- speaks one word-one word RR=30/ PR=140/m decreased breath sounds in both lungs	Orthopnea RR=30/m PR=140/m BP=180/95 mmHg T=37 C Heart auscultation S3, S4 rales in both lungs
<i>Part B</i>			
1	2	3	4
CXR: Cardiomegaly and infiltration in both lung specifically in hilums. ECG: sinus tachycardia ABG: PaO2:65mmHg PCo2:38 mmHg PH: 7.38	CXR reveals peripheral wedge of airspace opacity ECG: sinus tachycardia ABG: PaO2:65mmHg PCo2:30 mmHg PH: 7.45	CXR reveals hyperinflation. ECG: sinus tachycardia ABG: PaO2:45mmHg PCo2:34 mmHg PH: 7.42	CXR: severe hyperinflation ECG: Prominent P in lead 2 ST changes in inferior and right leads
<i>Part C</i>			
1	2	3	4
1- CCU admission 2-Swan-Ganz catheter 3- Fluid therapy	1- semi-seated position 2- ICU admission 3-O2 4-Lasix IV 5-Morphine IV	1- IV Heparin 2- CT angiography of lungs	1- relaxation 2-nasal O2 3- Salbutamol spary 2 puff every 10 minutes

Table 3. Key of the Sample PT

	Part A	Part B	Part C
Scenario 1	1	2	3
Scenario 2	3	3	4
Scenario 3	2	4	1
Scenario 4	4	1	2

With respect to the components of illness script, faults gradually become less prominent, while the differences between expertise levels are small in consequences. Therefore, these two components are not suitable to use in PT. In contrast, management strategies such as enabling conditions, monotonically increase with expertise level and are suitable to be a part of PT structure.

Moreover, Mx-knowledge not only grows, but also integrates with the Dx-knowledge into the illness script. That is, the workup of these two tasks (Dx and Mx) should usually be done together in every patient. Thus, these two types of knowledge become highly integrated and balanced in experts' illness scripts. The mutual relationship between Dx and Mx-knowledge is crucial for a successful and efficient patient encounter, and is assessed by PT (14-16).

The optimal number of scenarios to reach the reliability of 0.9 is 20 chief complaint groups, each containing four scenarios (11).

### Scoring

Here, answers were not weighted and a combination of items in four parts (patient's history, physical examination, para clinic and Dx and Mx plans) was considered as the correct answer, which was given a full score. When two or three pieces were matched, a part of the full score (four pieces matched) was allocated. For example, when referring to one, two and three correct pieces, the score would consecutively be 0.25, .5 and 0.75 of the full score.

### Conclusion

With PT, examinees are evaluated on the specific clinical reasoning task of pattern recognition for routine clinical cases, which is a crucial step within the clinical reasoning process. PT does not replace other clinical reasoning assessment tools, but it com-

plements them.

In general, other clinical reasoning exams are mostly focused on analytical reasoning. For example, KFP (Key Features Problems) not only thoroughly examines all parts of the clinical reasoning process in a deliberate manner, but also assesses each part in detail. On the other hand, CRP is generally dedicated to diagnostic reasoning, while both diagnostic and management reasoning is assessed in PT. Finally, PT differs from SCT (Script Concordance Test) as each scenario in SCT is assigned to several diagnoses, while each PT scenario has only one diagnosis.

PT format allows the examiners to explore a facet of clinical reasoning, which is usually excluded from traditional medical assessments but physicians frequently face them in daily clinical practice: Rapid decision making in routine situations. Studies have shown that students extensively accept PT because it is enjoyable for them to match the puzzle components (11,12).

### References

1. Norman GR, Brooks LR. The non-analytical basis of clinical reasoning, *Adv Health Sci Educ* 1997;2:173-184.
2. Schuwirth L. Is assessment of clinical reasoning still the Holy Grail? *Med Educ* 2009;43:298-299.
3. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses, *Med Educ* 2008;42:468-475.
4. 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. *J Edu Health Promot* 2012; 1:10.
5. Monajemi A, Rikers RM. The role of patient management knowledge in medical expertise development: Extending the contemporary theory. *Int J Pers Cent Med* 2011;1:109-14
6. Charlin B, Boshuizen HPA, Custers EJFM, Feltovich Paul J: Scripts and clinical reasoning. *Med Educ* 2007;41:1179-1185.

7. Norman G. Research in clinical reasoning: Past history and current trends. *Med Educ* 2005;39:418-27.
8. Schmidt HG, Rikers RM. How expertise develops in medicine: Knowledge encapsulation and illness script formation. *Med Educ* 2007;41:1133-9.
9. Custers EJFM. Thirty years of illness script: theoretical origins and practical application, *Med Teach* 2014.
10. Case SM, Swanson DB. Extended-matching items: A practical alternative to free-response questions, *Teach Learn in Med* 1993;5(2):107-115.
11. Ber R. The CIP (comprehensive Integrative puzzle) assessment method. *Medical Teacher* 2003; 25(2):171-176.
12. Amini M, Moghadami M, Kojuri J, Abbasi H, Arhami DoolatAbadi A, Molaee NA, 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.
13. 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);2 (9):222.
14. Monajemi A, Schmidt HG, Rikers RM. Assessing patient management plans of doctors and medical students: An Illness script perspective. *J Contin Educ Health Prof* 2012;32:4-9.
15. Monajemi A, Rikers RM, Schmidt HG. Clinical case processing: A diagnostic versus a management focus. *Med Educ* 2007;41:1166-72.
16. Monajemi A, Rostami ER, Savaj S, Rikers RM. How does Patient Management Knowledge Integrate into an Illness Script? *Educ Health* 2012;25:153-9.