HOW DO MEDICAL STUDENTS ACTUALLY THINK WHILE SOLVING PROBLEMS IN THREE DIFFERENT TYPES OF CLINICAL ASSESSMENTS: OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE), MULTIMEDIA CASE-BASED ASSESSMENT (CBA), AND MODIFIED ESSAY QUESTION (MEQ)

by

SEJIN KIM

(Under the Direction of Ikseon Choi)

ABSTRACT

In the context of medical education, clinical assessments, such as the objective structured clinical examination (OSCE) and modified essay question (MEQ), have been widely used. Although both have numerous advantages, minimizing the limitations of each assessment and assessing medical students' diagnostic reasoning using a reasonable and affordable method is needed. This study was conducted as an exploration of an alternative way to conduct clinical assessments in the form of a multimedia case-based assessment (CBA).

There are various research studies on investigating correlations between the OSCE and MEQ; however, little attention has been given to the types of thinking that medical students actually engage in during assessments. The purpose of the study was to identify medical students' cognitive processes in solving diagnostic problems and to compare how they think differently in three different types of clinical assessments.

A cross-case study was employed for this research. The study involved two 4th year medical students who had been videotaped taking the OSCE, CBA, and MEQ. Data were collected through one-on-one stimulated recall interviews where students were shown a video of themselves taking each assessment and asked to elaborate what they were thinking during each of the 20 partitioned clips of each video. Data were prepared with the smallest phrases or sentences representing a meaningful cognitive occurrence and coded using hypothetico-deductive reasoning (HDR) as representative of clinical reasoning. Any uncoded data were categorized as other cognitive occurrences, and then all data were reconstructed according to the chronology of the participant's actual performances in the assessment.

The study revealed that both research participants exhibited similar proportional frequencies for all types of cognitive occurrences; however, each type of clinical assessment presented different patterns of proportional frequencies of clinical reasoning process. Moreover, other cognitive occurrences that distract students' clinical reasoning were also detected, such as test-taking strategy, point-seeking/hunting, and unnecessary constraints. As a result, suggestions for future research are provided. This study's research design may be used to validate clinical assessments for diagnostic reasoning, and the results of this study can inform the redesign of clinical assessments, including multimedia case-based assessment, for medical education.

INDEX WORDS: Clinical reasoning, Cognitive occurrences, Real-world problem-solving, Cross-case study, Clinical assessment, Medical education, Cognitive affordances

HOW DO MEDICAL STUDENTS ACTUALLY THINK WHILE SOLVING PROBLEMS IN THREE DIFFERENT TYPES OF CLINICAL ASSESSMENTS: OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE), MULTIMEDIA CASE-BASED ASSESSMENT (CBA), AND MODIFIED ESSAY QUESTION (MEQ)

by

SEJIN KIM

B.A., Chung-Ang University, Republic of Korea, 2008

M.Ed., Chung-Ang University, Republic of Korea, 2010

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2018

HOW DO MEDICAL STUDENTS ACTUALLY THINK WHILE SOLVING PROBLEMS IN THREE DIFFERENT TYPES OF CLINICAL ASSESSMENTS: OBJECTIVE STRUCTURED CLINICAL EXAMINATION (OSCE), MULTIMEDIA CASE-BASED ASSESSMENT (CBA), AND MODIFIED ESSAY QUESTION (MEQ)

by

SEJIN KIM

Major Professor: Ikseon Choi
Committee: Allan S. Cohen
Janette R. Hill
Bo Young Yoon

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
May 2018

DEDICATION

This dissertation is dedicated to my mother, Meeyoung Jung, and my father, Jaehack Kim. You have always been there for me and supported me with your prayers on my academic journey. Without your unconditional love and encouragement, I could not have overcome the many challenges of my life, let alone have completed this dissertation. This is for you.

ACKNOWLEDGEMENTS

I would like to express deepest appreciation to my advisor, Dr. Ikseon Choi, for his academic support and advice. I am grateful for his encouragement and insightful guidance for my dissertation. I still remember his presentation at GLDI summer program in 2010, which was one of the reasons I decided to begin my doctoral study with him. I would also like to thank my committee members: Dr. Allan S. Cohen, who provided invaluable feedback and guidance in the area of educational assessment; Dr. Janette R. Hill, who was so caring and generous with her time to support and guide my research; and Dr. Bo Young Yoon, who suggested inspirational research directions and helped me collect and analyze data. Each of my committee members provided unique and valuable advice for my research.

My special appreciation goes to faculty members at Inje University College of Medicine: Dr. Byung Doo Rhee, Dr. Jong-tae Lee, Dr. Seok-jin Choi, Dr. Min Jeong Kwon, Dr. Sang-hyun Kim. I was able to begin my research with their unsparing support, including their time, effort, and resources, and they provided suggestions and feedback at every step of my study. I also would like to thank staff members at IUCM and my research participants for their participation.

I want to express my gratitude to my former advisor at Chung-Ang University, Dr. Heesu Lee, who showed me how to study by the way he lives; and Dr. Soonok Jo, who taught me all things research related. I sincerely appreciate them for giving me all the opportunities I have had.

Another special thanks go to JAMA/GLDI members: Dr. John C. Kim, Sarah S. Kim, Cathy Rabb, and all staff members, who provided eye-opening experiences. They are my great supporters, who were extremely generous to me in helping me adjust to the cultural differences in the United States, and they have become like family members to me who support me spiritually with their prayers.

Lastly, I would like to acknowledge the support from my friends: My academic colleagues from Research and Innovation in Learning and CBeL teams, UGA KSA members, Jasmine Choi, Dongho Kim, Meehyun Yoon, Minho Kwak, and Hyojung Han, lifelong buddies in Korea, and Elliot Kim, my dinner mate and great supporter for my English writing. Their support in each different part was tremendously helpful for me to stay on this journey.

Above all, from the bottom of my heart, I would like to attribute all the glory to God, who gave me life and strength. I am nothing without You.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

One of the primary goals in higher education is to develop learners' problem-solving abilities needed in real-world contexts (Choi & Lee, 2009; Feldman, 2009; Jonassen, 1997). To help learners develop and enhance their problem-solving skills and abilities most instructional models in schools and universities deal with well-structured problems that contain clear problem statements and correct answers with complete information (Jonassen, 1997). However, most problems in the real-world are related not to well-structured problems but ill-structured problems, which represent everyday human reasoning that has undefined elements in the problems, goals, and information needed to solve the problems (Chi & Glaser, 1985; Jonassen, 1997; Voss, 1988). According to research, a classroom environment framed around well-structured problems does not help learners think and reason to solve ill-structured real-world problems (Choi & Lee, 2009; Grotzer & Perkins, 2000). As alternatives of the traditional instructional methods, various methods, such as case-based learning, problem-based learning, and project-based learning have been developed for enhancing real-world problem-solving (Eseryel, Ifenthaler, & Ge, 2013; Jonassen, 2011; Spector, 2006).

Problem Statement

Although several instructional methods have been used for developing real-world problem-solving abilities, the efforts of promoting these approaches are often limited by a lack of reliable and affordable methods for assessing both intended learning outcomes and real-world problem-solving abilities (Eseryel et al., 2013; Jonassen, 2011; Spector, 2006). As indicated above, well-structured problem-solving and ill-structured problem-solving feature fundamentally different approaches to solving problems. However, well-structured problems with clear problem statements and correct answers based on complete information are the ones most commonly encountered in exams (Jonassen, 1997). Although assessing ill-structured problem-solving requires a variety of different tasks and measurements, recall or simple domain tests have been fragmentarily used instead (Jonassen, 1997; Spector, 2006). Generally, the amount of domain knowledge that learners can recall is the primary method of assessment (Jonassen, 2011). Although established approaches such as think-aloud protocol and cognitive task analysis are already in use, these processes are difficult to be used in assessment because ill-structured problems may have multiple answers, and these approaches must be accompanied with time consuming work and effort that include knowledge elicitation, data analysis, and knowledge representation (Eseryel et al., 2013; Crandall, Klein, & Hoffman, 2006). Moreover, because current educational goals focus on developing problem-solving abilities rather than on simple knowledge acquisition (Choi & Lee, 2009; Feldman, 2009; Jonassen, 1997), an alignment between instruction, learning and assessment (ILA) is essential and therefore an alternative assessment to close the gap between the way of teaching and learning and the way of assessing knowledge is required (Biggs, 1996; Gulikers, Bastiaens, & Kirschner, 2004; Segers. Dochy, & Cascallar, 2003).

These aforementioned challenges suggest the need for more research in order to develop an alternative way to assess learners' abilities in solving ill-structured problems in a real-world context. Characteristically, ill-structured problems can have multiple possible answers and various ways of solving processes (Kitchener, 1983). Jonassen (1997) writes that "Ill-structured problem-solving becomes a process of iteratively restricting alternatives and refining arguments before selecting a solution" (p. 81). Because the perspective of thinking like experts treats learning as a process (Spector, 2006), the fastest and most efficient way to enhance learners' problem-solving skills and abilities is to assess their thinking process during problem-solving and provide feedback regarding their thought process. Through the feedback process, learners will be able to realize the importance of this method in solving the problems, rather than the traditional methods of memorizing domain knowledge or guessing the correct answer (Jonassen, 2011).

With efficiency and accountability for assessment, multiple-choice types of questions as traditional standardized tests have been considered as a powerful solution with several benefits for measuring student achievement (Hart, 1994). The multiple-choice question is one of the most flexible types of objective tests (Thorndike & Thorndike-Christ, 2010) and assesses a variety of content (Osterlind, 1998). Because the multiple-choice questions are simple to use, easy to grade, and highly reliable (Meng, Kang, & Lee, 1994), it is typically used for high-stakes tests involving a number of examinees at the same time (Hogan, 2007). Collecting aggregated test scores makes it possible to compare an individual's scores and school's average scores and identify the strengths and limitations of individuals in order to help them improve (ACGME & ABMS, 2000). Despite these several benefits, however, traditional multiple-choice questions have limitations. In particular, multiple-choice questions are an inappropriate assessment method for assessing real-world problem-solving abilities because multiple-choice questions are often used for measuring simple knowledge acquisition and are based on students' recall abilities of a certain set of knowledge (Hart, 1994). Students may be misled in thinking that there is a single right answer to solve a problem and be encouraged not to solve a problem but to find only a right answer (Hart, 1994).

In the context of medical education, medical knowledge regarding basic and clinical science (knows) and competence levels (knows how) can be measured easily through a written examination with multiple-choice questions (ACGME & ABMS, 2000; Kim & Hur, 2005; Miller, 1990). However, the purpose of medical education is to help students develop essential competencies needed for medical treatments and foster students' competencies to solve problems (Small, Stevens, & Duerson, 1993; Park, 2004; Park, 2008). In order to provide appropriate treatments for patients, medical doctors must have medical knowledge as well as clinical performance abilities to solve problems (Hwang & Jeong, 2011; Park, 2004; Han et al., 2004). Medical students have gained medical knowledge in order to perform it and have been trained to develop medical competencies (Choi & Sunwoo, 2009). Therefore, assessing medical students' core competencies is one of the major areas in medical education (Cho, Kim, Park, & Hwang, 2011), and it is important to assess not only their medical knowledge as well as medical performance (Hwang & Jeong, 2011).

In order to assess students' clinical competencies, a written examination was traditionally used for assessing medical knowledge, although clinical assessment should include all areas of medical treatment, knowledge (knows), competence (knows how), performance (shows how), and action (does) (Cho et al., 2011; Hwang & Jeong, 2011; Miller, 1990; Park, 2008). There have been many attempts to apply the framework that includes knowledge, competence, performance, and action for clinical assessment (Miller, 1990). For example, a written examination or interview using graphic slides or pictures tried to assess the level of performance and action (Cho et al., 2011). However, it remains another type of knowledge test, and these attempts have a limit in assessing actual areas of students' performance and action (Cho et al., 2011; Park, 2008). Simple knowledge tests have several limitations. First, most of the questions from multiple-choice items are not used for assessing students' clinical problem-solving but their memorization skills (Hur, Kim, & Park, 2007; Kim, 2009; Meng et al., 1994). This type of questions assesses only a single dimension of clinical competence based on content knowledge (Sloan, Donnelly, Schwartz, & Strodel, 1995). According to studies investigating the relationship between the result of the written examination and clinical performance abilities, conflicting results have shown that the clinical performance abilities cannot be predicted by written examinations (Hur et al., 2007; Hwang & Jeong, 2011; Kim, Lee, Choi, & Lee, 2004; Kramer et al., 2002; Remmen et al., 2001). Second, multiple-choice questions limit students' reasoning and thinking processes regarding problem-solving by reducing the represented knowledge into simple statements (Osterlind, 1998). Although the multiple-choice format provides flexibilities and conveniences for developing and distributing test items, these advantages are construed as negative characteristics due to the simplification of knowledge. Moreover, some distractors within test items may reduce important knowledge for problem-solving and allow the creation of a guessing factor, where students are distracted from actual problem-solving (Meng et al., 1994; Osterlind, 1998).

To overcome the limitations of written knowledge tests and to close the gap between medical education and medical practice, the Korean medical education field is currently using the objective structured clinical examination (OSCE) and clinical performance examination (CPX) to assess performance and action areas, and these examinations have now been widely used in medical assessment (Cho et al., 2011; Han et al., 2004; Hur et al., 2007; Hwang & Jeong, 2011). Harden et al (1975) introduced the OSCE in medical education, which can measure knowledge, behavior, and clinical skills at the same time. The OSCE has been implemented in Korean medical education as a method of teaching and assessment since 1994 (Choi & Sunwoo, 2009; Han et al., 2004; Park, 2004). Through the OSCE, students' clinical performance based on clinical knowledge can be assessed using the OSCE as a standardized way for assessing clinical competence (Choi & Koh, 2008; Tervo et al., 1997; Hwang & Jeong, 2011). According to ACGME and ABMS (2000), "the OSCE format provides a standardized means to assess: physical examination and history taking skills; communication skills with patients and family members, breadth and depth of knowledge; ability to summarize and document findings; ability to make a differential diagnosis, or plan treatment; and clinical judgment based upon patient notes" (p. 7). Because a procedure of problem-solving in clinical practices can be assessed through the OSCE, it offers a significant advantage in improving the quality of physicians' clinical competence (Koh & Park, 2009). Although clinical performance examinations have several advantages for assessing students' actual reasoning process, there are also some drawbacks (ACGME & ABMS, 2000; Kramer et al., 2002; Park, 2004; Smee, 2003). First, creating a test set and administering the test with planning the examination, training the standardized patients, implementing the examination, and grading the scores require considerable costs, time, and effort (ACGME & ABMS, 2000; Carpenter, 1995). In order to create the test, the test objectives and difficulty of the test should be discussed in a preparation stage, and the number of stations and staff affects the costs and the quality of assessment (Carpenter, 1995; Kramer et al., 2002, Park, 2004). Second, using stations with limited time forces examinees to perform isolated aspects of the clinical encounter (Smee, 2003). This might be solved by increasing the number of stations to provide reliability; however, it causes the same problem mentioned above (Kramer et al., 2002; Park, 2004; Smee, 2003). Third, because the OSCE uses a simulation-based approach with standard patients and scenarios, there are limits to what can be simulated and what can be sampled for patient problems (Smee, 2003). Due to these limitations, administering the OSCE method on a large-scale is restricted (Kramer et al., 2002).

Although both types of examinations have numerous advantages, minimizing the limitations from each examination and assessing medical students' clinical problem-solving abilities with a reasonable and affordable method is needed. Therefore, a new approach of assessment that complements both types of assessments explained above – a clinical knowledge and reasoning examination and a clinical competence examination in medical education – was introduced and examined in this study.

Research Focus

This study was conducted to understand how medical students think differently in three different types of clinical assessments. The purpose of the study is to identify medical students' cognitive processes while they encounter clinical diagnostic problems in three different types of clinical assessments, including: a comprehensive clinical competence examination (Objective Structured Clinical Examination), a multimedia case-based assessment, and a comprehensive clinical knowledge and reasoning examination (Modified Essay Question). This study discovered what clinical reasoning and other cognitive occurrences were afforded by the different assessment methods.

Traditionally, the information processing theory of knowing and learning posits that symbolic representations are processed in a person's head (Young, Kulikowich, & Barab, 1997). This approach has led to an objectivist view, with the concept of knowledge as a product (Fredreiksen, 1994), where one best solution in a certain context can be applied to another context, and that there is only one solution or correct answer in an umbrella of well-structured problems, each containing clear problem statements and correct answers with completed information (Jonassen, 1997; Young et al., 1997). On the other hand, from the ecological psychology perspective, knowing and learning occur in the interactions between a learner and an environment through direct perception (Gibson, 1979; Young et al., 1997), and most real-world problems fit with this perspective. From this approach, knowledge can be viewed as a process, and therefore the traditional assessments based on information processing theory with the concept of knowledge as a product may not be suitable for dealing with problem-solving processes, such as problem identification, solution generation and integration, and falsification and evaluation (Young et al., 1997). Characteristically, ill-structured problems can have multiple possible answers and various ways of solving processes (Kitchener, 1983). Therefore, in this study, medical students' problem-solving abilities were assessed by identifying their cognitive processes.

Moreover, this study was conducted based on the proposition that a cognitive gap between learners' problem-solving in real-world contexts and learners' problem-solving in a testing environment should be closer. Because learners' perceptions of problem-solving in real-world contexts and the perceptions of problem-solving in current testing environments are different, learners may only focus on hunting for the correct answers instead of solving problems. The cognitive affordances in a testing environment are the perceived relationships between what is afforded by perception and thinking in a given problem-solving situation. In other words, the perceived cognitive affordances from the testing environment in the context of problem-solving should be provided to students in order for them to determine how to solve the problems.

Research Questions

In order to investigate how different types of clinical assessments afforded medical students' different types of problem-solving, medical students' cognitive processes of clinical diagnostic problem-solving in three different types of clinical assessments were identified, and the purpose of the study was fulfilled by answering the following research questions:

Research Question 1. What are the cognitive processes of clinical diagnostic problem-solving in three different types of clinical assessments: clinical competence examination, multimedia case-based assessment, and clinical knowledge and reasoning examination?

> Research Question 1.1. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a clinical competence examination (Objective Structured Clinical Examination)?
>
> Research Question 1.2. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a multimedia case-based assessment?
>
> Research Question 1.3. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a clinical knowledge and reasoning examination (Modified Essay Question)?
>
> Research Question 1.4. How do medical students think similarly to solve clinical diagnostic problems in each assessment?

Research Question 2. How do medical students think differently to solve clinical diagnostic problems in the three different types of assessments?

Importance of the Study

Medical education is one of the places where real-world problem-solving ability development is needed. Jonassen (1997) writes that "ill-structured problem-solving becomes a process of iteratively restricting alternatives and refining arguments before selecting a solution" (p. 81). Clinical diagnostic reasoning is one type of ill-structured problem-solving and an iterative process "converting observed evidence into the names of diseases" (Feinstein, 1973, p. 212; Higgs & Jones, 2000). In order to assess medical students' clinical diagnostic reasoning in solving problems, it is important to first identify what the students' cognitive processes are while they encounter clinical problems in testing environments to understand how they solve problems.

Traditionally, there have been difficulties assessing learners' problem-solving abilities in real-world settings due to a lack of reliable test methods (Jonassen, 2011; Spector, 2006). Moreover, little attention has been given to exploring the kinds of thinking that occur during tests. Generally, multiple-choice types of test items afford thinking where students intentionally or unintentionally find the right answer by eliminating least likely possibilities or hunting for the most probably answer instead of problem-solving. In assuming that there is a limitation of current clinical examinations for creating an appropriate test context that affords students to think authentically and solve real-world problems.

Although "today's multimedia instructional systems are capable of creating highly realistic, situated problem-solving experiences for students, designing assessment components for these instructional systems remains a challenge" (Shaw, Effken, Fajen, Garrett, & Morris, 1997, p. 151). Another importance of the study is that a new affordable assessment method with multimedia platforms is needed for medical students to afford authentic thinking and reasoning to solve clinical problems in testing settings, where the assessment is helpful to assess real-world problem-solving to complement both clinical performance examination and written knowledge test in medical education. Furthermore, the new assessment method can provide alternative ways to deploy technology for various types of assessment methods using multimedia platforms to solve the current assessment problems mentioned in the problem statement section above. As a result, a cost-effective multimedia-based assessment method to assess the process of medical students' diagnostic reasoning was designed and tested.

CHAPTER 2

LITERATURE REVIEW

This chapter provides identified core theories and literature that are closely related to this study. A thorough literature review of the relevant theoretical and conceptual research was conducted to develop a foundation for the study. The first section of this chapter provides an overview of diagnostic reasoning, including the hypothetico-deductive reasoning model utilized for identifying medical students' cognitive processes in this study. The second section reviews various types of clinical examinations in South Korean medical schools. These examinations were used to compare medical students' cognitive processes for clinical problem-solving. The third section contains a brief overview of assessing real-world problem-solving and of the development process of a multimedia case-based assessment. The final section identifies the theory of affordances from ecological psychology as a core foundation for the study.

Diagnostic Reasoning

Diagnostic Reasoning Models

Diagnosing a medical problem is an ill-structured problem-solving situation that requires dynamic decision-making (Kassirer & Gorry, 1978; Spector, 2006). Diagnostic reasoning is a complex problem-solving process that physicians face every day as ill-structured clinical problem-solving in multidimensional contexts (Higgs & Jones, 2000). Efficient and accurate diagnosis is one of the most critical functions of physicians' roles (Croskerry, 2009; Higgs & Jones, 2000; Kassirer & Gorry, 1978; Kassirer, 1989). Although there are several types of diagnostic reasoning models or approaches in the medical field (Croskerry, 2009; Edwards, Jones, Carr, Braunack-Mayer, & Jensen, 2004; Hardin, 2002; Patel, Groen, & Norman, 1993; Payton, 1985; Kassirer, 1989), two models below can be represented as being broad concepts explaining diagnostic reasoning (Coderre, Mandin, Harasym, & Fick, 2003; Edwards et al., 2004; Patel & Groen, 1986; Patel, Groen, & Norman, 1993).

Novice model (Hypothetico-deductive Model; Backward Reasoning). Non-experts (medical students) typically use limited information to diagnose patients' diseases because of a lack of knowledge (Barrows, 1994; Kassirer, 1989). Hypothetico-deductive reasoning (HDR), a process of sequential steps from a hypothesis to data, is widely used for diagnosis by novice students (Edwards et al., 2004; Hardin, 2002; Patel et al., 1993). Due to their lack of knowledge, students struggle with generating specific hypotheses with the result that they lean toward using non-selective search-and-seek procedures (Kassirer, 1989). The goal of HDR is to validate information or data acquired from the patients within the limitations of available standards (Edwards et al., 2004). The reasoning process of HDR includes: (1) problem framing, (2) hypothesis generation, (3) inquiry strategy, (4) data analysis or synthesis, (5) diagnostic decisions, and (6) therapeutic decisions (Barrows, 1985; 1994; Barrows & Tamblyn, 1980; Edwards et al., 2004; Hardin, 2002; Ju & Choi, 2018).

Expert model (Illness-scripts Model; Forward Reasoning). Experts recognize key elements or features of a certain case almost directly and instantly, drawing from previous experiences and proper knowledge structures by recognizing patterns and formulating problems from meaningful schemes (Barrows, 1994; Edwards et al., 2004; Hardin, 2002; Kassirer, 1989; Patel et al., 1993). This expert type method, called illness-scripts model, is a process of sequential steps from data to formulate a hypothesis (Hardin, 2002; Patel et al., 1993). Extensive experience with knowledge structures emerging from continued exposure to patients help expert physicians acquire a repertoire of problems termed illness-scripts (Coderre et al., 2003). This process allows experts to solve problems by applying similar or identical solutions to those already solved and recalled (Coderre et al., 2003; Kassirer, 1989). Scheme-inductive reasoning or pattern recognition is used as a similar concept (Patel & Groen, 1986). The illness-scripts model has the following four stages: (1) development of elaborated causal networks, (2) compilation of relations between elaborated networks and abridged ones, (3) emergence of illness scripts, and (4) storing of patient encounters as instance scripts (Hardin, 2002).

Hypothetico-Deductive Reasoning Model

Much research has been done on physicians' problem-solving processes (Patel, Arocha, & Zhang, 2012). Clinical reasoning involves an inferential process based on higher level of cognitive thinking in medicine, such as problem-solving and decision making, for diagnostic or therapeutic decisions (Feinstein, 1973; Patel et al., 2012). Elstein, Shulman, and Sprafka (1978) investigated the problem-solving processes using experimental methods and theories of cognitive science to examine clinical competencies, and the research influenced the development of a model of hypothetico-deductive reasoning (Feltovich, Johnson, Moller, & Swanson, 1984; Patel et al., 2012).

Hypothetico-deductive reasoning (HDR) is defined as a process of means-end analysis used to narrow diagnostic hypotheses to reduce the distance between the point where the problem solver is and where the problem solver would like to be using clinical inquiry and data (Elstein et al., 1978). Physicians first generate hypotheses from clinical data and knowledge and test a set of hypotheses to account for clinical data (Higgs & Jones, 2000; Patel et al., 2012). Hypothetico-deductive reasoning is an effective way to deal with the challenges from patients' ill-structured clinical problems through a logical problem-solving process (Barrows, 1994).

The use of hypothetico-deductive reasoning may be different based on the level of physicians' experiences and case difficulty. According to Barrows (1985, 1994), physicians may use the hypothetico-deductive reasoning method to solve unfamiliar or difficult cases. Norman, Trott, Brooks, and Smith (1994) found that if clinical cases are simple or not difficult for physicians to solve, a pattern recognition type of method, which directly leads to diagnosis from data, can be used. On the other hand, difficult or problematic cases, which require an iterative process from hypothesis generation to data analysis or testing, can be resolved by the HDR process (Higgs & Jones, 2000).

Hypothetico-deductive reasoning process. The process of hypothetico-deductive reasoning is not fixed as a single-form process model; instead, the process may incorporate similar phases according to several studies on hypothetico-deductive reasoning in medicine (Barrows, 1985, 1994; Barrows & Tamblyn, 1980; Edwards et al., 2004; Elstein et al., 1978; Hardin, 2002; Patel et al., 2012).

Elstein et al. (1978) introduced hypothetico-deductive reasoning as a process of problem-solving whereby the problem-solver: generates a number of hypotheses, obtain limited data, and generates a set of possible solutions through seminal empirical studies. Novices and experts use a similar process to generate hypotheses and find solutions; however, the quality of content in the process may be deeply different (Higgs & Jones, 2000).

Barrows and Tamblyn (1980) described hypothetico-deductive reasoning as a clinical reasoning process. The reasoning process of HDR includes: (1) perceiving a variety of cues, such as patient responses, prior records, or non-verbal cues such as patient age or location of the visit, (2) generating hypotheses, including ideas, hunches, guesses, impressions, or diagnoses, usually from past experiences, (3) acquiring information that employs a problem-oriented search strategy in a well-disciplined, logical approach or rule-in/rule-out shortcuts, (4) adding all new data, and (5) making a diagnostic and therapeutic decision (Barrows & Tamblyn, 1980; Edwards et al., 2004; Hardin, 2002). Detailed information regarding the hypothetico-deductive reasoning process and its goal is provided in Table 2.1.

Table 2.1.
Hypothetico-deductive reasoning steps (Barrows & Tamblyn, 1980; Hardin, 2002)

Step	Goal	Procedure
Perceiving a variety of cues	To initiate the direction and scope of the reasoning process	Information can be taken from the patient and the setting either spontaneously or in response to the physician's own questions
Generating hypotheses	To generate possible explanations from the initial concept for the patient's problem	The collection of hypotheses serves as the guide for the physician's interview and examination of the patient
Acquiring information	To shape or refine the hypotheses by using a variety of data-collection techniques	Search and scan (when the physician is unable to further rank, verify, deny, or refine hypotheses) approach can be used at this stage. The physician's initial questions often tend to rule-in or rule-out a large number of possible hypotheses and to limit the patient's problem to a workable size. Then, the physical examination is used to confirm any hypotheses that still remain after the information is acquired
Formulating problems	To formulate the patient's problem by adding all new data to the initial concept	As this data is added, the physician's picture of the patient grows and evolves
Making a diagnostic and/or therapeutic decision	To make a decision with obtained data or decide that the patient's problem is urgent, and that immediate care or treatment is needed	- Scanning to be sure that data has not been overlooked and to gain more confidence in the hypotheses chosen - Making additional inquiries to get to know the patient as a person - Asking questions that are helpful in selecting treatment or management options - Asking questions about the convenience of hospitalization, willingness to have further investigations, other medications the patient may be taking or prior reactions to medications

In the later study on hypothetico-deductive reasoning, Barrows (1994) emphasized the importance of the HDR process for effective and efficient care for patients and presented the process in a more simplified way. As a result, which was based on previous studies, six phases were finally identified for hypothetico-deductive reasoning: (1) problem framing, (2) hypothesis generation, (3) inquiry strategy, (4) data analysis or synthesis, (5) diagnostic decision, and (6) therapeutic decision (Barrows, 1994; Ju & Choi, 2018).

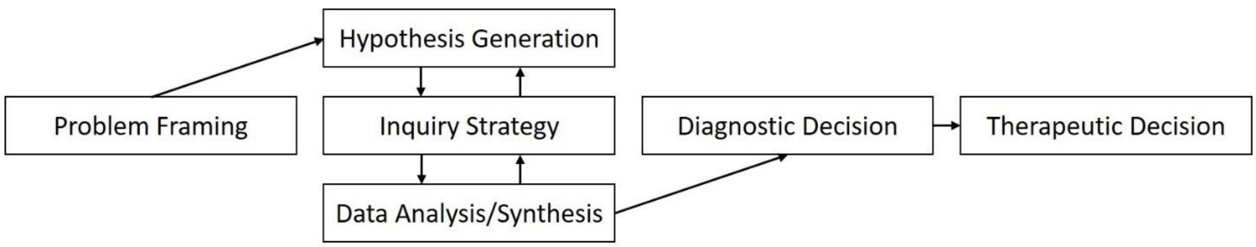


Figure 2.1. Hypothetico-deductive reasoning process. Adapted from Barrows (1994).

Patel et al. (2012) explained hypothetico-deductive reasoning as a combination of deductive and inductive processes, which involved following four stages: cue acquisition, hypothesis generation, cue interpretation, and hypothesis evaluation. Cues from clinical cases lead to generation of possible hypotheses in a deductive process, and then each hypothesis is interpreted and evaluated with cues based on clinical data in an inductive process.

Conflicting interpretations regarding hypothetico-deductive reasoning. Elstein et al. (1978) found that problem-solving in medicine is more dependent on physicians' mastery of a particular domain than a particular strategy or general procedures. Higgs and Jones (2000) explained Elstein's findings as case specificity that describes how successful diagnoses are represented by experiencing various clinical cases and where differences between physicians' diagnostic reasoning may be explained more as an understanding of numerous cases than from different reasoning models or strategies (Elstein et al., 1978). This finding challenges the real-world implications of the hypothetico-deductive model, and the model has further been criticized by researchers in cognitive psychology (Higgs & Jones, 2000). Based on the result of the studies, as mentioned before, not all physicians use hypothetico-deductive reasoning for diagnoses; in particular, experienced physicians in every day clinical situations utilize pattern recognition, based on rapid, automatic processes, more than HDR. Even if experts use hypothetico-deductive reasoning for problem-solving in medicine, HDR can be affected by prior experiences (Higgs & Jones, 2000; Schmidt, Norman, & Boshuizen, 1990). Furthermore, the experienced physicians' diagnostic reasoning could be nonanalytic (Norman, Young, & Brook, 2007). Patel et al. (2012) explained that based on the complexity of diagnostic reasoning nature, more than one type of reasoning is employed to solve clinical diagnostic problems. Recent research on clinical reasoning has tried to introduce various types of clinical reasoning models, such as dual-process theory (Evans, 2008).

Different Types of Clinical Examinations in Korean Medical Schools

According to ACGME (Accreditation Council for Graduate Medical Education), there are six core competencies medical doctors should have: patient care; medical knowledge; practice-based learning and improvement; interpersonal and communication skills; professionalism; and systems-based practice (Swing, 2007). The purpose of medical education is to help students develop essential competencies needed for medical treatments and foster students' competencies to solve problems (Small et al., 1993; Park, 2004; Park, 2008). In order to provide appropriate treatments for patients, medical doctors must have medical knowledge as well as clinical performance abilities to solve problems (Hwang & Jeong, 2011; Park, 2004; Han et al., 2004). Medical students have gained medical knowledge in order to perform it and have been trained to develop these competencies (Choi & Sunwoo, 2009).

Moreover, assessing medical students' core competencies is one of the major areas in medical education (Cho et al., 2011), and is important in assessing not only their medical knowledge but also medical performance (Hwang & Jeong, 2011). In order to assess students' competencies, a written examination was traditionally used for assessing medical knowledge, although clinical assessment includes all areas of medical treatment, knowledge (knows), competence (knows how), performance (shows how), and action (does) (Cho et al., 2011; Hwang & Jeong, 2011; Miller, 1990; Park, 2008). There have been many attempts to apply the framework that includes knowledge, competence, performance, and action for clinical assessment (Miller, 1990) to medical assessment. For example, a written examination or interview using graphic slides or pictures tried to assess the level of performance and action (Cho et al., 2011). However, it remains another type of knowledge test, and these attempts have a limit in assessing actual areas of students' performance and action (Cho et al., 2011; Park, 2008). Currently, the OSCE (objective structured clinical examination) and CPX (clinical performance examination) were developed to assess the performance and action areas, and these examinations have been widely used in medical assessment (Cho et al., 2011; Hwang & Jeong, 2011).

In most medical schools in Korea, two types of assessment have been used: a comprehensive clinical knowledge and reasoning examination and a comprehensive competence examination. The Korean National Health Personnel Licensing Examination has been changed several times since 1994 (Kim, 2009). In particular, the National Licensing Examination had been implemented for 15 clinical subjects but has been changed to 3 subjects, including general pathology, special pathology, and clinical laws. This changed occurred because it was deemed that the important knowledge for physicians is not fragmentary knowledge based on clinical subjects but comprehensive knowledge for problem-solving (Hwang & Jeong, 2011; Kim, 2009).

Moreover, clinical performance examinations were adopted since 2010 to assess not only students' medical knowledge but also their actual performance abilities (Cho et al., 2011; Hwang & Jeong, 2011; Park, 2008). In the next section, both the comprehensive knowledge examination and comprehensive performance examination are discussed based on their features, benefits, and limitations.

Comprehensive Clinical Knowledge and Reasoning Examination

For the comprehensive clinical knowledge and reasoning examination, a written examination is typically used (Cho et al., 2011; Park, Chung, Hong, Lee, & Shin, 2009). Some schools use a chart stimulated oral examination (CSR) (Accreditation Council for Graduate Medical Education [ACGME] & American Board of Medical Specialties [ABMS], 2000). These types of examinations include a comprehensive basic science examination and comprehensive clinical science examination.

Features. The MCQ (multiple choice questions) method is widely used for comprehensive clinical knowledge and reasoning examination in medical education (Park, Park, Kim, & Hwang, 2015; Park et al., 2009). A written or computer-based type of examination can be used for assessing students' understanding of medical knowledge and problem-solving with the MCQ. Several types of questions can be developed including selecting the best answer, one correct answer, negative answer, and combined response (Meng et al., 1994). The number of answer options in the multiple-choice questions varies according to the type and content of questions (Thorndike & Thorndike-Christ, 2010).

Each question is composed of an introductory statement, including a patient case, cues, symptoms, or medical information followed by answer options (ACGME & ABMS, 2000). The examinee should select an answer or answers according to the types of questions. In most medical schools in Korea, a written examination of comprehensive clinical science knowledge has been used for first and second-year medical students to assess their basic and clinical science knowledge (Park et al., 2015). Moreover, the current National Licensing Examination in Korea has 360 questions, and examinees need to solve all test items in 460 minutes.

Another way of assessing clinical knowledge and reasoning is the MEQ (modified essay question). The MEQ was introduced in the late 1960s as a paper-based test featuring an evolving situation to assess more general practices than traditional types of clinical assessments (Knox, 1989). Because several short-essay types of test items are included in the MEQ with a clinical scenario, the MEQs more often assess higher order cognitive skills than the MCQ (Palmer, 2010).

Benefits. The multiple-choice question is one of the most flexible types of objective tests (Thorndike & Thorndike-Christ, 2010) and assesses a variety of content (Osterlind, 1998). As already mentioned above, medical knowledge regarding basic and clinical science (knows) and competence levels (knows how) can be measured easily through the written examination with multiple-choice questions (ACGME & ABMS, 2000; Kim & Hur, 2005; Miller, 1990). Because the multiple-choice questions are simple to use, easy to grade, and highly reliable (Meng et al., 1994), it is typically used for high-stake tests with a number of examinees at the same time (Hogan, 2007). Collecting aggregated test scores is possible to compare an individual's scores and school's average scores and identify the strengths and limitations of individuals in order to help them improve (ACGME & ABMS, 2000).

The MEQ is a context-dependent test using clinical cases in written form, and examinees are required to recall their clinical knowledge and apply the knowledge to practice (Knox, 1989).

Due to the nature of the MEQ, it has been used for assessing reasoning and decision-making abilities, rather than simple, clinical knowledge recall (Feletti & Smith, 1986).

Limitations. Comprehensive clinical knowledge and reasoning examinations have limitations despite several benefits. First, a written test as a common method for the examination is limited in assessing medical students' performance and action areas (Park et al., 2015; Park, 2008). Here, performance and action include clinical performance and having empathy to establish a better relationship between physician and patient (Ogle, Bushnell, & Caputi, 2013; Park et al., 2015). According to Sloan et al. (1995), "other important aspects of clinical expertise, such as physical examination skills, interpersonal skills, technical skills, problem-solving abilities, decision-making abilities, patient treatment skills are not assessed objectively" (p. 736). There is a restriction that the test results cannot detect whether or not medical students perform well and communicate well with patients when the written examination is the only way to assess students' clinical abilities (Park, 2008). Furthermore, most of the questions from multiple-choice items are not used for assessing students' clinical problem-solving but their memorization skills (Hur et al., 2007; Kim, 2009; Meng et al., 1994). These questions assess only a single dimension of clinical competence based on content knowledge (Sloan et al., 1995). According to studies regarding the relationship between the result of the written examination and clinical performance abilities, conflicting results have shown that the clinical performance abilities cannot be predicted by written examinations (Hur et al., 2007; Hwang & Jeong, 2011; Kim et al., 2004; Kramer et al., 2002; Remmen et al., 2001). Moreover, the MEQ type of assessment has limitations for grading due to inadequate guidelines and inconsistencies in marking (Palmer et al., 2010).

Second, multiple-choice questions limit students' reasoning and thought processes regarding problem-solving by reducing the represented knowledge into simple statements (Osterlind, 1998). Although the multiple-choice format provides flexibilities and conveniences for developing and distributing test items, these advantages are construed as negative characteristics due to the simplification of knowledge. Moreover, when important knowledge necessary for problem-solving is reduced into discrete test items, the test creates the possibility of a guessing factor, which may distract students from problem-solving (Meng et al., 1994; Osterlind, 1998).

Comprehensive Clinical Competence Examination

The primary goal of medical education is not to train students who have enormous medical knowledge but to foster students who can perform well in clinical treatments (Han et al., 2004; Koh & Park, 2009; Ludmerer, 1985; Park, 2004). For this reason, a competency-based curriculum has been emphasized by developing new types of assessment methods for current medical education (Choi & Koh, 2008; Kim, 1995; Park, 2004; Small et al., 1993). Various assessment methods have been developed and implemented for clinical performance skills and abilities (ACGME & ABMS, 2000; Kim & Hur, 2005; Kramer et al., 2002; Martin, Lloyd, & Singh, 2002). Among the assessment methods, the Objective Structured Clinical Examination (OSCE) and the Clinical Performance Examination (CPX) are the most reliable tests, widely used around the world (Choi & Koh, 2008; Choi & Sunwoo, 2009; Hur et al., 2007; Park, 2004; Park, 2008; Park et al., 2005; Reznick et al., 1993; Williams et al., 1987).

Objective Structured Clinical Examination (OSCE). In order to avoid the disadvantages of the traditional clinical examination and close the gap between medical education and medical practice (Han et al., 2004; Hur et al., 2007), Harden et al (1975) introduced the OSCE in medical education, which can measure knowledge, behavior, and clinical skills at the same time. The OSCE has been implemented in Korean medical education as a method of teaching and assessment since 1994 (Choi & Sunwoo, 2009; Han et al., 2004; Park, 2004).

Features. The OSCE is a type of clinical competence assessment, a multidimensional practical examination of clinical skills, developed for assessing the extent of students' performance abilities and identifying the results of clinical practice education (Harden et al., 1975; Park, 2004; Sloan et al., 1995). In the testing environment, examinees rotate through a number of stations followed by a clinical procedure, and the examinees are asked to solve problems related to their findings and interpretations or perform appropriate treatments (Harden et al., 1975; Newble, 2004). Five to ten minutes are given to solve problems at each station, and examinees' answers and performance are assessed by examiners with a prepared checklist (Park, 2004). Standardized patients are considered to be the primary assessment tool; however, other assessment tools, including data interpretation exercises or clinical scenarios, can be used in the OSCE (ACGME & ABMS, 2000). There are guidelines for the OSCE as a clinical performance assessment (Park, 2004; Tervo et al., 1997): (a) Assessing comprehensive clinical skills and abilities is needed, (b) the level of difficulty needs to be controlled relative to the curriculum that students have learned, and (c) a test blueprint should be developed and provided based on learning objectives.

Benefits. As contrasted with clinical knowledge assessment, students' clinical performance based on clinical knowledge can be assessed using the OSCE in a standardized way for assessing clinical competence (Choi & Koh, 2008; Tervo et al., 1997; Hwang & Jeong, 2011). According to ACGME and ABMS (2000), "the OSCE format provides a standardized means to assess: physical examination and history taking skills; communication skills with patients and family members, breadth and depth of knowledge; ability to summarize and document findings; ability to make a differential diagnosis, or plan treatment; and clinical judgment based upon patient notes" (p. 7). Because a procedure of problem-solving in clinical practices can be assessed through the OSCE, it offers a significant advantage in improving the quality of physicians' clinical competence (Koh & Park, 2009).

In any procedure of the OSCE, the effect of cueing in solving problems can be diminished and therefore achieve similarity with real-world contexts (Harden et al., 1975). When examinees are presented with questions at the stations, they cannot revisit previously encountered stations to rectify their decisions or any missing points. Therefore, questions in the OSCE do not provide hints in solving problems or suggest tentative solutions to handle problems (Harden et al., 1975). Another benefit of the OSCE is that the reliability and validity of OSCE have been proved through research (Anderson et al., 1991; Choi & Sunwoo, 2009; Park, 2004; Park, 2008; Sloan et al., 1995).

Limitations. Although the OSCE is a useful assessment type for clinical skills and abilities, there are some drawbacks (ACGME & ABMS, 2000; Kramer et al., 2002; Park, 2004; Smee, 2003). First, creating a test set and administering the test are difficult (ACGME & ABMS, 2000). In order to create the test, the test objectives and difficulty of the test should be discussed in a preparation stage, and the number of stations and staff affects the costs and the quality of assessment (Carpenter, 1995; Kramer et al., 2002, Park, 2004). For example, if the number of stations is below 10, it is difficult to develop a test with the proper variety of cases and scenarios; on the other hand, if the number of stations is above 20, more staff and faculty members are needed, driving up costs considerably (Park, 2004). Second, using stations with limited time forces examinees to perform isolated aspects of the clinical encounter (Smee, 2003). This might be solved by increasing the number of stations to provide reliability; however, it causes the same problem mentioned above (Kramer et al., 2002; Park, 2004; Smee, 2003). Third, because OSCE uses a simulation-based approach with standardized patients and scenarios, there are limits to what can be simulated and what can be sampled for patient problems (Smee, 2003). Due to these limitations, administering the OSCE on a large-scale is restricted (Kramer et al., 2002).

Clinical Performance Examination (CPX). Another way to assess clinical performance is the Clinical Performance Examination (CPX). The CPX is a type of assessment to assess clinical performance in a variety of clinical settings that mimic real-world situations (Choi & Sunwoo, 2009). In comparison to the OSCE, which assesses examinees' specific clinical performance abilities in a certain area, the CPX assesses examinees' entire clinical performance abilities using a standardized patient who presents a particular disease (Hwang & Jeong, 2011; Kwon et al., 2005; Park et al., 2005). Barrows (1993) defined a standardized patient as a simulated patient or actual patient who has been trained to present his or her own illnesses. The CPX is a performance assessment that can predict a student's comprehensive clinical competence using standardized patients, and it has been widely used after a standardized patient concept was accepted in medical curricula and assessment (Cho et al., 2011; Choi & Koh, 2008; Choi & Sunwoo, 2009; Kim et al., 2004; Stillman & Swanson, 1987; Williams et al., 1987).

Features. A standardized patient method has provided a possibility to assess students' clinical performances in simulated situations. Examinees enter a station where a situated case is given, then they encounter a standardized patient to provide clinical treatment (Park et al., 2005).

Examinees are given 10 to 15 minutes to perform the clinical treatment followed by greeting, history taking, physical examination, clinical courtesy, patient-physician interaction, patient education, and bad news delivery to the standardized patient (Cho et al., 2011; Park et al., 2005). After finishing the clinical treatment, examinees are asked to complete a post-encounter note. Examiners are composed of faculty members and the standardized patient, and they assess examinees' clinical performance including patient and advisor satisfaction scores (Cho et al., 2011). Typically, grading is conducted with a checklist, and satisfaction scores are calculated with a Likert scale (Kim et al., 2004; Cho et al., 2011).

Benefits. The CPX has various benefits in assessing students' clinical performances. First, an entire clinical performance procedure, including history-taking skills, physical examination skills, communication skills, differential diagnosis, laboratory utilization, and treatment, can be assessed through standardized patients (ACGME & ABMS, 2000; Park et al., 2005). The reliability and validity of evaluations from standardized patients are supported by research (Elliot, 1994; Kwon et al., 2005; Shin, Lee, & Park, 2005; Vu et al., 1992). The most important benefit from the CPX is that students need to show their performance in patient encounters in a situated context. Applied knowledge, behaviors, and clinical skills can be assessed at the same time, and the examination provides authentic experiences for students to focus more on important behaviors, including communication with patients and clinical treatments rather than memorization of clinical knowledge (Kim et al., 2004; Park, 2008).

Limitations. Assessment through the CPX, however, has its limitations. Firstly, there is a major restriction regarding simulated situations (Park et al., 2005). The topic covered by the CPX lean toward internal medicine because standardized patients have difficulty expressing symptoms that represent the diseases of external medicine (Park et al., 2005). Some research indicates that a written examination can be an alternative method to clinical performance examination based on the correlation between the written examination scores and CPX scores (Kramer et al., 2002; Remmen et al., 2001). Secondly, another limitation of the CPX is that planning the examination, training the standardized patients, implementing the examination, and grading the scores require considerable costs, time, and effort (Carpenter, 1995). This limitation is similar to the restrictions of the OSCE, which similarly make the CPX difficult to administer on a large-scale (Kramer et al., 2002).

Assessing Real-world Problem-solving

Assessment is not only the most important component in learning and instruction (Spector, 2006) but also a significant area of interest. However, assessment is one of the weakest aspects of the learning process (Jonassen, 2011, Spector, 2006). Most educators know the best way to teach is exemplified by an end-product where a student can learn how to solve problems resulting in meaningful learning experiences. The problem is that most educators do not know how to construct, implement, and evaluate a meaningful assessment for problem-solving abilities (Jonassen, 2011). Moreover, constructing, implementing, and evaluating assessments for whole parts of problem-solving with the proper skill sets is not an easy task for educators; therefore, they simply use recall-test type assessment (Jonassen, 2011). However, if an appropriate form of assessment is not ready to assess learners' problem-solving abilities, and only recall tests based on learners' memorization skills are used for assessing problem-solving abilities, learners' mental efforts in learning to solve problems will not occur (Jonassen, 2011).

It is important to note that learners' learning outcomes can be assessed by a performance assessment concept that believes that an assessment should be performed to measure how learners construct meaningful solutions through the problem-solving process (Elliott, 1994;

Jonassen, 2011; Spector, 2006). Several research studies on developing assessment methods with problem-solving process have been conducted (Eseryel et al., 2013; Jonassen, 2011; Pirnay-Dummer, Ifenthaler, & Spector, 2010; Spector, 2006). However, several limitations of the methods, such as difficulties in developing and grading and the inefficiency of time and effort, were revealed (Eseryel et al., 2013; Jonassen, 2011; Pirnay-Dummer et al., 2010; Spector, 2006). Moreover, it remains difficult to develop an assessment method for problem-solving using traditional assessment types (e.g., standardized multiple-choice questions, short open-ended questions) for accurate and automatized assessments.

Multimedia Case-based Assessment as an Alternative Approach

Case-based learning refers to a set of learning and teaching methods using cases that have multifaceted events or narratives to provide authentic situations (Allen, Otto, & Hoffman, 2000; Barnes, Christensen, & Hoffman, 1994; Choi, Hong, Park, & Lee, 2013; Shulman, 1992). Through case-based learning, students' problem-solving abilities are promoted (Choi et al., 2013; Danielson et al., 2007), and it can be applied with flexibility in various learning environments (Choi et al., 2013). The characteristics of cases are as follows: (1) all cases are realistic, (2) cases provide contextualized accounts of teaching, and (3) cases are used for educational purposes (Grossman, 1992). Because cases have a narrative, a story, and a set of events (Shulman, 1992), students focus more on problems rather than learning the content of cases. For this reason, Shulman (1992) pointed out that the importance of cases is that students can reinterpret the cases by identifying context and background, and that they can have multiple representations of the cases. According to research, case-based learning is more motivating and promotes better learning transfer from theory to practice because cases tend to make implicit meaning more visible (Allen et al., 2000; Shulman, 1992). Thus, the fields of law, medicine, and business utilize case-based instruction more frequently in order to promote expertise in the exercise of the profession (Williams, 1992).

Based on the concepts of case-based learning, various case-based e-learning platforms were designed, developed, and implemented in anesthesiology (Choi et al., 2013; Choi, Lee, & Kang, 2009), veterinary clinical education (Creevy et al., 2017), and teacher education (Choi & Lee, 2008;; Choi & Lee, 2009;; Lee & Choi, 2008). Along with the case-based learning platforms, test items have also been developed. There are several development phases for multimedia case-based assessment, and video cases are considered an appropriate medium using current technology for providing authentic context (Spiro, Collins, & Ramachandran, 2006). The detailed development procedure of each phase is described below.

(1) Case selection. The development of multimedia case-based assessment starts from selecting cases. Based on discussions with subject-matter experts, a case is selected for the development of multimedia case-based assessment. The case is considered to be selected based on the availability of resources, time, and feasibility.

(2) Scenario development. After selecting cases for multimedia case-based assessment, a scenario development phase is next. In many cases, the scenario development is conducted through collaborative work with subject-matter experts for multimedia case-based assessment. Basic information of characters, contexts, background information, references, narrations, and scripts regarding conversations are included in a scenario. Figure 2.2 provides a sample page of a scenario developed for multimedia case-based assessment.

Patient Information					
Gender	Male	**Age**	59	**Name**	Jungho Kim
Economic Level	Middle Class	**Educational Level**	Bachelor	**Occupation**	Self-employed
Height	173cm	**Weight**	81kg		

Context	Mr. Kim, 59-year old, came to the hospital due to chest pain <Vital Signs> Blood pressure: 140/90 mmHg Pulse: 88/minute Breathing rate: 20/minute Body temperature: 36.7°C A doctor should • Establish a close rapport with the patient and find the chief complaint • Confirm the patient's medical history about the major symptom • Do a physical examination • Discuss initial diagnosis results and future treatment plans
Reference	1. Silverman, J., & Kurtz, S. (2013). Skills for communicating with patients, 3rd ed.

Clinical Scenario	
Physician	**Patient**
Hello. I am John Doe, your doctor today. Can you tell me your name and age?	**Yes, my name is Jungho Kim. I'm 59 years old.**
Okay, thank you for the information. If you have any problems or feel uncomfortable, please let me know.	**Okay.**
So, what's your occasion for the visit?	**[Symptom 1 – chest pain]** **(Grab your chest and pretend to breathe hard) I'm having chest pain… it's like pushing my chest…**
When did it start?	**Hmm… about 30-40 minutes before.**
About 30 minutes before… do you have any other pain?	**[Symptom 2 – dyspnea]** **Hmm… I feel pressure on my chest, it's hard to breathe.**
If you feel pressure on your chest, I believe you must be feeling very uncomfortable… anything else?	**[Symptom 3 – Anxiety]** **I don't know… feeling pressure on my chest, I'm like… I'm afraid I am going to fall.**

Figure 2.2. A sample of a developed scenario for multimedia case-based assessment

(3) Video case development. With the developed scenario, actors and video production teams are recruited to develop video cases. Filming video cases is an important phase for multimedia case-based assessment, and enhancing the level of authenticity is a important for students to focus more on the case. Due to the importance of voice quality, external mics are often used to enhance the audio quality. Figure 2.3 shows samples of the setting during filming.

Figure 2.3. A setting for the filming of a multimedia case-based assessment

(4) Test items development. Test items of each case for the multimedia case-based assessment are developed after video development finishes. After the video case development work is done, the initial edited video cases are analyzed for test item development through several discussions with subject-matter experts, who developed the original scenario and contexts. Based on the discussions regarding the test items, the videos are divided into several short videos according to major decision points. Furthermore, initial test items are developed for each video case. Table 2.2 provides a sample of the initial test items in a multimedia case-based assessment. Base on the initial test items, several meetings may be needed to finalize the test items through pilot tests.

Table 2.2.
A sample of test items developed for multimedia case-based assessment

DP 1 (20s)	
Question 1	Please describe all clinical cues from the patient.
Question 2	Please formulate clinical hypotheses with the patient's clinical characteristics.
Question 3	Please list the patient's history and explain your decision.
DP 2 (60s)	
Question 1	Please describe additional meaningful clinical cues from the patient.
Question 2	Please list three clinical hypotheses and explain your decision.
Question 3	Please describe additional history taking for diagnosis and explain your decision.
DP 3 (60s)	
Question 1	Please describe additional meaningful clinical cues from the patient.
Question 2	Please list three clinical hypotheses and explain your decision.
Question 3	Please describe a list of physical examinations needed to prove the hypotheses and explain your decision.
DP 4 (90s)	
Question 1	Please describe the important physical examination results.
Question 2	Please list tentative diagnoses (up to 3) and explain your decision.
Question 3	Please describe a list of lab tests or additional physical examinations to confirm the diagnosis and explain your decision.
Question 4	Please describe initial patient management plans at this point.

(5) Multimedia case-based assessment development. To distribute the multimedia case-based assessment, which includes several video clips that provide authentic situations in convenient and affordable way, an online version of the assessment is often used. Using an online survey development and distribution system, multimedia case-based assessment can be developed, modified, and distributed.

The Theory of Affordances

The term "affordances," first coined by Gibson (1977), has influenced a variety of fields, such as perceptual psychology, cognitive psychology, human-computer interaction (HCI), interaction design, and instructional design (Dabbagh & Dass, 2013; Greeno, 1994; Hartson, 2003; Hutchins, 2010; McGrenere & Ho, 2000). According to Gibson (1977), "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill" (p. 127). The original definition of affordances by Gibson (1977, 1979) refers to the relationship between an environment (a physical object) and an organism (a person). Norman (2013) explains that the affordance is "a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used" (p. 11). Warren and Whang (1987) provides a condition that "the organism perceive what actions are afforded by a given situation" (p. 371) to further explain the concept of affordances. Based on these explanations, affordances can be defined as an ability or the possibility of objects to be perceived as opportunities for action in the relationship between the environment and the objects (Gibson & Pick, 2000; Reed, 1996; Zukow-Goldring & Arbib, 2007).

The concept of affordances originated from the ecological approach to visual perception (Gibson, 1979), and Gibson's perspective regarding perception, that it is a system that picks up information supporting an organism's action in the environment, led to the development of the theory of affordances (Greeno, 1994; Reed, 1996). Gibson (1979) stresses the importance of the relationship between the environment and the organism; in particular, the direct perception of how the environment affords various actions to the organism. The purpose of perception is to obtain meaningful information that affords acting in the environment quickly and efficiently (Gaver, 1991; Gibson, 1979; Hutchins, 2010; Kaptelinin, n.d.), and the values and meanings can be directly perceived (Gibson, 1979; McGrenere & Ho, 2000). According to Gibson and Pick (2000), this perception-action relationship is not a unilateral protocol but a reciprocal arrangement. Obtained perception guides appropriate actions, and the actions in turn yield more information for further guidance (Gibson & Pick, 2000). Gibson's concept of perception being based on awareness and action rather than memory and retrieval was radical because the core idea in traditional cognitive psychology was information processing theory, which Gibson strongly opposed (Greeno, 1994; Reed, 1996).

However, Greeno (1994) and Norman (2013) point out that people have confused the concept of affordances in many ways. Greeno (1994) interprets affordances as “the characteristics of objects and arrangements in the environment that support their contributions to interactive activity and, therefore, the characteristics of the environment that agents need to perceive” (p. 341). Hartson (2003) also suggests that the terminologies surrounding the concept of affordances need to be calibrated and reorganized. The relationships among others’ use of the terminologies are summarized (Gaver, 1991; Gibson, 1979; McGrenere & Ho, 2000; Norman, 2013), and the concept of affordances distinguished in two ways: physical affordances and cognitive affordances (Hartson, 2003).

Physical Affordances

In Gibson’s view of affordances (1979), the affordances are explained in terms of what it offers the animal, what it provides or furnishes in a relationship between the environment and the organism. Gibson (1979) provides examples of ecological frames, such as mediums, surfaces, and substances, regarding affordances for visual perception. McGrenere and Ho (2000) and Hartson (2003) point out that Gibson considered the concept of affordances as a physical relationship to perceive directly from an object. In this context, physical affordances are independent of a person’s previous experiences and cultural differences (Gibson, 1979). In other words, the affordances of an object are not related to the person’s knowledge, culture, or ability to perceive (Hartson, 2003). Norman (1999) acknowledges that the term of affordances from Gibson’s perspective refers to real affordance, which are physical characteristics that allow its operation (Hartson, 2003). Hartson (2003) categorizes this as physical affordances and defined it as “a design feature that helps, aids, supports, facilitates, or enables physically doing something” (p. 319).

Cognitive Affordances

In most of the literature, affordances are explained as characteristics that provide clues or hints on how to operate an object (Hartson, 2003; Norman, 2013). Norman (2013) explains that the affordances are “a relationship between the properties of an object and the capabilities of the agent that determine just how the object could possibly be used” (p. 11). Hartson (2003) categorizes this as cognitive affordance, and defines it as “a design feature that helps, aids, supports, facilitates, or enables thinking and/or knowing about something” (p. 319). However, as Hartson (2003) points out, the concept of cognitive affordances has been used with different terminologies, such as perceived affordances, apparent affordances, or perceptual information about affordances. Norman (1999, 2013) stresses the differences between real affordances [physical affordances] and perceived affordances [cognitive affordances] and the importance of perceived affordances (McGrenere & Ho, 2000). Because perceived affordances are closely dependent on previous experiences, knowledge, or the cultural perspectives of a person (McGrenere & Ho, 2000), nothing is directly perceived; instead, a cognitive process is needed to acquire information from an object (Norman, 2013). Based on this assumption, cognitive affordances refer to the relationship between what is afforded by perception and thinking in a given situation.

The concept of cognitive affordances is considered to be an essential element in case-based assessment. The most important challenge of current assessments is an alignment between instruction and assessment, and therefore an alternative assessment to close the gap between the between the ways of teaching and learning and the ways of assessing knowledge is required.

Moreover, as much as it is possible, closing the cognitive gap between students' problem-solving in real-world contexts and students' problem-solving in a testing environment will be ideal. Because students' perceptions of problem-solving in real-world contexts and the perception of problem-solving in current testing environments are different, students may only focus on hunting for the correct answers instead of solving problems.

The cognitive affordances in a testing environment are the perceived relationships between what is afforded by perception and thinking in a given problem-solving situation. In other words, the perceived cognitive affordances from the testing environment in the context of problem-solving should be provided to students for them to determine how to solve the problems.

CHAPTER 3

METHODOLOGY

This chapter is an overview of the methodology proposed for the qualitative case study. This study is designed to identify medical students' cognitive processes when solving clinical problems in three different types of clinical assessments, and, by comparing the differences of cognitive reasoning processes, to understand how the different types of clinical assessments afford medical students in their problem-solving approaches. In this chapter on the research methodology, the first section describes research design regarding a cross-case qualitative study, overview of research design procedure, and intervention design. In the second section, the research site and participant information for the study are provided. The third section explains the details of the data collection method and protocol, including stimulated recall protocol, general data collection procedure, and a detailed data collection protocol for each condition. The fourth section presents the data analysis framework and protocol. The fifth section provides information on the validity and reliability of the research. The last section describes methodological limitations to this study.

Research Design

A qualitative case study research method was employed for this study. A case study refers to "a qualitative design in which the researcher explores in depth a program, event, activity, process, or one or more individuals" (Creswell, 2014, p. 241). More importantly, a case study can be used for analyzing case(s), a bounded system by time and activity, "through detailed, in-depth data collection" in context (Creswell, 1998, p. 61; Creswell, 2014). This research was conducted as a case study to collect complex and in-depth explanations of phenomena (de Vaus, 2001) because the purpose of this study is to identify medical students' cognitive processes for clinical problem-solving in three different types of clinical assessments. Patton (2015) states that "cases can be empirical units" (p. 259), and in this study, a case was defined as an individual participant's problem-solving process derived from each testing environment.

However, different views on the case study also exist (Patton, 2015). One common misunderstanding is to define a case study as a preliminary stage to explore a participant's activity or conduct fieldwork for a certain research method (Yin, 2014). Moreover, there is some confusion regarding what exactly a case study is (Merriam, 1998; Yin, 2014). One perspective on case study defines it as an end product similar to Creswell's definition mentioned above (Merriam, 1988; Patton, 2015). A case study from this perspective focuses on a detailed and rich story about the case, which is a person, organization, event, or program as a unit of analysis, to understand the case itself (Patton, 2015; Stake, 2006). According to Merriam (1988), a qualitative case study can be defined as "an intensive, holistic description and analysis of a single instance, phenomenon, or social unit" (p. 21). On the other hand, a case study defines a case study is a process of research or method of inquiry (Patton, 2015). Additionally, Yin (1994) described case study as "an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are clearly evident" (p. 13).

In this study, a case study is considered to be a holistic description and analysis of a single unit as Creswell (1998, 2014) and Merriam (1988) highlight, as well as a method of inquiry as Yin (1994) described. The case study here can be defined as a method of inquiry where the researcher studies a case through detailed, in-depth data collection from multiple viewpoints of information in context (Creswell, 1998; Merriam, 1988; Patton, 2015).

Cross-case Study

Multiple cases were analyzed for the research to answer the research questions listed in Chapter 1. In this study, a cross-case analysis method was framed within the research design of a case study. Cross-case analysis can be performed as a case study to compare or contrast research findings by aggregating the findings across a series of individual studies (Yin, 2014). Because a cross-case analysis allows a researcher to understand how relationships exist among individual cases and develop accumulated knowledge from the original case (Khan & VanWynsberghe, 2008), the researcher can compare or contrast the similarities and differences of cases by analyzing the aggregated data. A data analysis method may vary according to how many cases are available, and quantitative methods or meta-analysis can be conducted to draw cross-case conclusions, if possible (Yin, 2014). However, cross-case analysis results still strongly rely on "argumentative interpretation, not numeric tallies" (Yin, 2014, p. 167).

To answer the research questions and align with the purpose of the study, cross-case analysis was employed to identify similarities and differences of medical students' cognitive processes when solving clinical problems in multiple cases. First, each case data was collected with the same interview protocol and analyzed to identify cognitive processes in solving problems. Second, a cross-case analysis for this study was conducted to understand the relations among different types of clinical assessments. Detailed case definitions and research design procedures are described in the next section along with a graphical representation.

An Overview of the Research Design and Procedure

This study was designed to identify medical students' cognitive processes while they solve a clinical problem in three different types of assessments and to understand the similarities and differences of the processes. Based on a cross-case study design, two medical students participated for the study, and three different types of conditions for each participant were selected to compare medical students' cognitive processes: clinical competence examination (OSCE: Objective Structured Clinical Examination), multimedia case-based assessment (CBA), and clinical knowledge and reasoning examination (MEQ: Modified Essay Question). The participants completed each type of assessment in order of the condition numbers: 1 for the OSCE, 2 for the CBA, and 3 for the MEQ. In this study, therefore, a case was defined as a condition in which a medical student participated, and a total of six cases were collected and analyzed.

Data were collected from these three conditions for each of the two participants. First, this study started with condition 1: the OSCE. Two 4th year medical students as research participants first participated in the OSCE. All events were video captured to collect participants' performances from the OSCE. Collected video data from those two participants were utilized for follow-up interviews, which were video and audio captured, to identify the similarities of clinical reasoning processes in the OSCE. Second, condition 2 was a multimedia case-based assessment one week after the condition 1 data collection was finished. The two participants joined the pre-developed multimedia case-based assessment for condition 2. All performances were captured with video and audio recorders to collect data. The collected data were used for follow-up interviews, which were also video and audio captured, to identify the similarities of clinical reasoning processes and patterns in the CBA. Third, the participants took the MEQ test for condition 3 one week after condition 2 data collection was finished. All processes were captured with video and audio recorders for data collection as well. The accumulated data from condition 3 were employed to identify the similarities of clinical reasoning processes in the MEQ. Lastly, comparison among identified data from all three conditions was made to understand how medical students reason to solve clinical problems in different types of clinical assessments. Detailed data collection methods and procedures are provided in the next section. The graphical representation of the research design and procedure is provided in Figure 3.1.

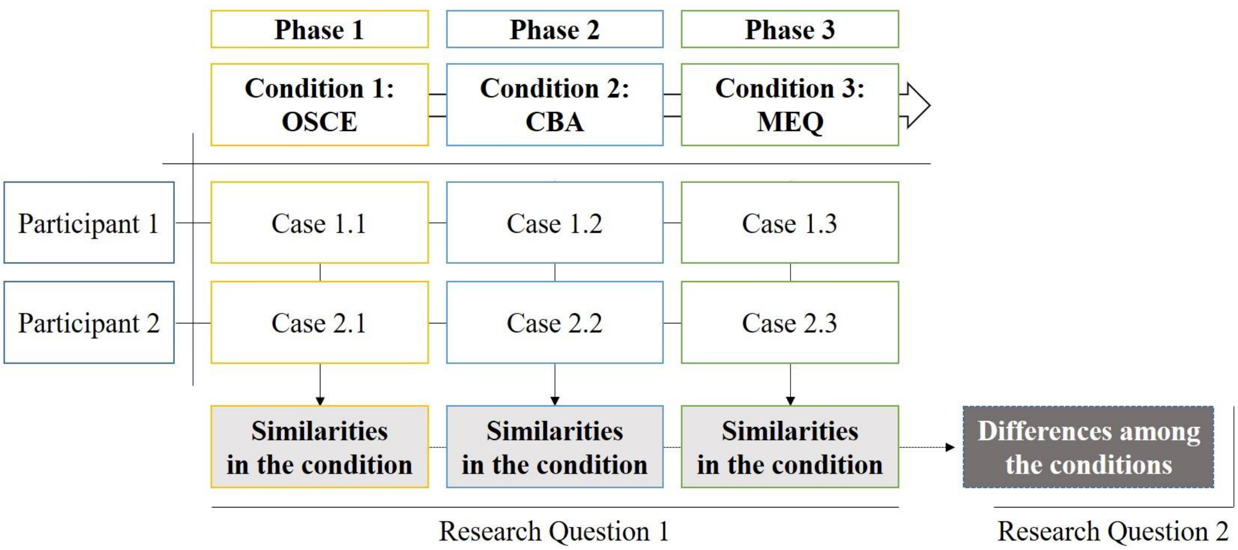


Figure 3.1. Research design and procedure: Three different types of conditions

Intervention Design: Three Different Types of Conditions

In order to identify medical students' cognitive processes of problem-solving in the three types of, a clinical presentation of chest pain was selected as the subject matter for all three assessments in the study. Based on preliminary discussions with subject-matter experts at Inje University College of Medicine (IUCM) and my dissertation committee members, clinical presentations of chest pain and abdominal pain were initially selected for the study. There are several reasons for carefully choosing from those two clinical presentations. First, these two clinical presentations are typical diseases allowing a doctor to use conventional reasoning for diagnosis. Second, chest pain and abdominal pain are familiar clinical presentations regularly practiced in medical school, which would allow this study to better control for participants' lack of knowledge as a factor in solving problems. This in turn favors participants' use of reasoning behaviors instead of hunting or guessing the right answer. Finally, the clinical presentation of chest pain was selected for the reason that data had to be collected from a licensure preparatory OSCE, one that had already been prepared by a regional consortium of OSCE examiners in South Korea using a presentation of chest pain. Thus, the three conditions of the OSCE, CBA, and MEQ with chest pain as the clinical presentation were prepared and provided for the research participants in order to compare the similarities and differences of their cognitive processes in solving examination questions.

To prevent learning effects from the previous test, the final diagnosis of each assessment was different, although each assessment maintained the same clinical presentation: The clinical presentation of chest pain was from stable angina, unstable angina, and aortic dissection for the OSCE, CBA, and MEQ, respectively. Each condition was conducted in that order: the OSCE, the multimedia case-based assessment (CBA), and the MEQ. The OSCE was provided for the first assessment because a preparatory OSCE organized by a regional consortium of OSCE examiners in South Korea, which is responsible for administering the national licensure examination process, was taken by all 4th year medical students at IUCM. Due to protecting the content of the preparatory OSCE and preventing the test effects from other assessments to the preparatory OSCE, the OSCE was selected as the first of the three conditions. The graphical representation of the intervention design and procedures are provided in Figure 3.2.

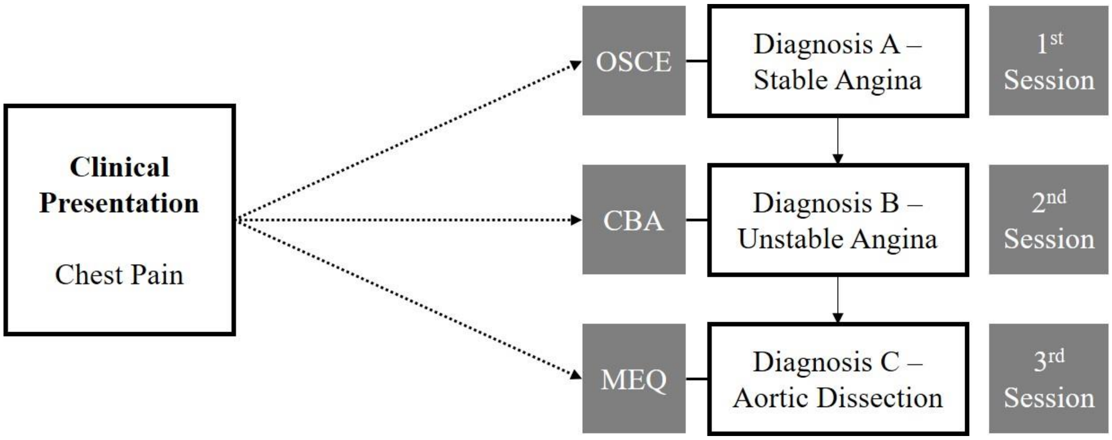


Figure 3.2. Intervention design and procedure

Research Site and Participants

Research Site

This study was conducted within the context of medical examinations at Inje University College of Medicine (IUCM) in South Korea. Since an international cooperative agreement between IUCM and the University of Georgia in 2014, full access has been granted, including working with faculty members at IUCM.

IUCM has a 6-year curriculum model. Medical students at IUCM learn liberal arts, English, medical humanities, and basic science/basic medical science as premedical courses for the first 2 years. After premedical courses, students learn various medical knowledge as pre-clerkship courses, such as clinical competency development, pathophysiology of clinical presentation, growth and aging, medical genetics, infection, hematopoietic system, gastrointestinal system, cardiovascular system, nervous system, and so on. During pre-clerkship courses, clinical knowledge and reasoning examinations (MEQ) with problem-based learning (PBL) modules are provided to enhance problem-solving abilities in real-world situations. In their final 2 years, students can have real-world experiences through clinical clerkships at one of five hospitals affiliated with IUCM. The objective structured clinical examination (OSCE) and clinical performance examination (CPX) as clinical competence examinations are administered for students to develop their basic clinical skills and clinical reasoning abilities.

Research Participants

Two 4th year medical students out of 119 4th year students at IUCM were selected as research participants for this study. As mentioned above, IUCM students have clinical clerkship experiences from affiliated hospitals from their 3rd year. After clinical clerkship courses are finished, students have an optional clinical clerkship experience and/or an intensive clinical clerkship course within a certain department.

An IRB application was submitted in May 2017 for participants recruiting; participant recruitment began immediately after IRB approval. The two participants recruited through purposeful sampling were 4th year medical students at IUCM, and the initially targeted participants were students who already had had clinical experiences in most of the clinical presentations from all departments, which were related to the area that would be covered in the clinical assessments. Thus, purposeful sampling eliminated 3rd year medical students in order to prevent participants from guessing in their problem-solving for the areas they had not yet covered.

The initial plan for selecting research participants was to recruit a total of four students for the study. Therefore, eight possible candidates were pre-selected based on their GPA and OSCE/CPX results. In order to select the research participants in a well-balanced way, two students were selected in the 75th percentile of GPA, four students were selected in the 50th percentile of GPA, and other two students were selected in the 25th percentile of GPA. Moreover, to eliminate outliers from OSCE/CPX scores, students who were below the average OSCE/CPX score were not selected for the study. The OSCE/CPX average score was 67.4 (out of 100), and standard deviation was 5.62. Hence, a total of four participants were initially recruited for the study. In addition to these four participants, two additional students participated for a pre-interview for condition 1 of the study. Based on the quality of the interviews, these two participants were also added for the research study; however, one participant could not participate in the following two interviews after the first interview due to a schedule conflict. Therefore, a total five participants, one in the 90^{th} percentile of GPA, one in the 75^{th} percentile of GPA, two in the 50^{th} percentile of GPA, and one in the 25^{th} percentile of GPA, were recruited in this research study, which screened for GPA, gender, and OSCE/CPX scores. Among the five candidates, two participants in particular richly elaborated their thinking during the interviews; therefore, based on a consideration of robustness and the details of interview data, the two high-GPA students, in the 90^{th} and 75^{th} percentile of GPAs, were finally selected for this study.

Data Collection Method and Protocol

After research participants were recruited, this study utilized video- and audio-recorded, in-depth stimulated recall interviews for data collection. To collect data, a two-step data collection protocol, collecting test performances and interview data, was designed where the stimulated recall interviews for each condition could identify the similarities and differences of medical students' cognitive processes in the three different types of clinical assessments. Detailed data collection settings and protocol are provided in the following sections.

Data Collection Method: Stimulated Recall Protocol

A series of in-depth interviews for cross-case study was conducted to fulfill the purpose of this research. This study used a stimulated recall protocol as a cognitive task analysis method for collecting data, enabling the illustration of research participants' cognitive processes while they encounter clinical problems in the different types of clinical assessments. The reason for using stimulated recall interviews was to capture how medical students think to solve clinical problems in different clinical assessments, and the stimulated recall protocol is an effective way to understand the cognitive processes when the students encounter clinical problems in assessments (Mackey & Gass, 2005). Using this retrospective way of interviewing helps a researcher assess cognitive processes by "inviting subjects to recall, … their concurrent thinking during that event" (Lyle, 2003, p. 861).

Cognitive task analysis (CTA) is a method to understand and describe how people think, what people pay attention to, what strategies people use to make decisions or solve problems, and what people try to accomplish by capturing the way the mind works, and cognition (Crandall et al., 2006). The purpose of CTA is to comprehend "how cognition makes it possible for humans to get things done and then turning that understanding into aids for helping people get things done better" (Crandall et al., 2006, p. 2). Thus, a researcher who tries to know "how participants view the work they are doing and how they make sense of events" (Crandall et al., 2006, p. 9) can understand the processes through CTA. CTA can capture "what people are thinking about, what they are paying attention to, the strategies they are using to make decisions or detect problems, what they are trying to accomplish, and what they know about the way a process works" (Crandall et al., 2006, p. 9). In other words, knowledge, cognitive processes, and goal structures from observable task performances can be captured through various ways of CTA methods (Chipman, Schraagen, & Shalin, 2000). There are numerous methods of CTA, such as critical decision method, cognitive function model, goal-directed task analysis, precursor, action, result, interpretation method, and so on, and three primary aspects of CTA include knowledge elicitation, data analysis, and knowledge representation (Crandall et al., 2006).

The stimulated recall protocol is a CTA retrospective method, and was used for this study. Stimulated recall protocol refers to a method that "represent a means of eliciting data about thought processes involved in carrying out a task or activity" (Gass & Mackey, 2000, p. 1). Through the stimulated recall method, a researcher can observe participant's internal thought processes while the researcher observes their external activities or events. The assumption of this stimulated recall method is that a visual or verbal prompt of a certain situation or event helps people recall their mental processes of the situation, thus stimulating recall of the situation. In other words, the stimulated recall protocol assesses research participants' thinking processes by analyzing previous events or activities that the participants encountered through recalling their reflections of critical moments (Gass & Mackey, 2000). Video or audio recordings can be often used as a tangible reminder of the encountered events or activities for participants to recall their thought (Beers, Boshuizen, Kirschner, Gijselaers, & Westendorp, 2008). This study used audio-visual recordings in the stimulated recall interview to identify qualitative insights of participants' cognitive processes in each of the three assessments. Thus, this study could use stimulated recall protocol to avoid distracting the participants while in their testing environments and reduce as much as possible any memory loss and distortion of the actual conditions the participants were asked to recall.

General procedure and guideline. The stimulated recall protocol is an effective way to retrieve cognitive processes and extract learner strategies from an unconscious level (Gass & Mackey, 2000). As discussed above, the stimulated recall method is typically used to identify participant's cognitive processes by asking several questions to stimulate recall of activities or events. Generally, the stimulated recall method can occur in two steps. First, to accomplish the goal of identifying participants' cognitive processes from a certain activity or event, the researcher records all activities using video or audio recorders. Second, the video or audio data is played back to the participants, and the researcher asks questions to the participants at each critical moment. These stimulated recall sessions are also videotaped or audiotaped for data analysis.

For better results, stimulated recall sessions may be conducted immediately after the activity or event finishes, and the strong stimulus, such as video sources, can help participants recall more vividly (Gass & Mackey, 2000). Moreover, a well-trained researcher can ask additional questions related to the participants' answers as well as a series of questions based on an interview protocol. Because it is important to be aware that "not all participants respond in the same way to the stimulus" (Gass & Mackey, 2000, p, 53), sometimes simple instructions may enough for the participants to retrieve their thought processes and patterns, but sometimes the researcher needs to ask appropriate additional questions to the participants to stimulate stronger recall after the initial stimulus. Regarding the structure of recall procedure, there is no general way to conduct the interviews; however, interview questions and protocols were developed based on the research questions in the study, and data collection through the stimulated recall protocol was conducted in the ways discussed above.

General Data Collection Procedure

As mentioned in the intervention design section, the three conditions (the OSCE, CBA, and MEQ) were set for data collection. Each condition was prepared through communications with faculty members from the research site and was designed through collaborative work. The data collection was conducted in two phases: (1) participants' performance data collection in each condition and (2) participants' cognitive activity data collection by stimulated recall interviews. A general procedure regarding the performance data collection (phase 1) and cognitive activity data collection (phase 2) is described below along with detailed information regarding data collection for each condition.

Phase 1 – research participants' performance data collection. The first phase of the data collection was capturing participants' performance data in each of the three conditions. To identify medical students' cognitive process during solving clinical problems in three different types of clinical assessments, each participant's performance on the assessments was videotaped. The video-captured data were used for the stimulated recall interviews in phase 2 of the data collection.

Phase 2 – research participants' cognitive activity data collection. After each performance data were collected, the data were used for the participants' cognitive activity data collection in phase 2. In order to identify the participants' cognitive processes in each condition, the video data captured from the three different types of clinical assessments were each divided into 20 segments. Each segment of the video was played back to the participants, and, while watching their own performance on video, they were asked to recall their cognitive occurrences by answering questions posed by the interviewer. This data collection was conducted as video-recorded, retrospective interviews following by the stimulated recall protocol. An audio recorder was also used as a backup, which captured the participants' voices. The participants were asked to participate in one-hour individual interviews for each condition. For better results, the stimulated recall interviews were conducted right after the participants finished each assessment; however, due to the limitation of using a preparatory OSCE organized by a regional consortium of OSCE examiners for condition 1, the OSCE interviews were conducted one week after the participant had finished the OSCE.

Interview protocol. The interview sessions were started by signing the approved IRB consent form. The researcher explained the background and purpose of the research, and a data management plan was also provided. The participants' performance videos in each condition had been divided into 20 segments, and general questions were posed to the participants to capture their cognitive occurrences. Additional questions were asked, if needed. Due to the differences of the participants' time spent on each assessment, each segment of the OSCE was 30 seconds, each segment of the CBA was 3 minutes, and each segment of the MEQ was 2.5 minutes. A list of retrospective interview questions used for the data collection is provided in Table 3.1.

Table 3.1.
Interview Questions for Each Condition

Session	**Interview Protocol and Interview Question**	**Timing**
Introduction	I will show your performance video taken in [the name of the assessment], and the video is divided into 20 segments. I will ask some questions at every [time].	At the beginning

General Questions	• What did you *see* here if you can recall the moment in this [time] video clip?
	• What did you *hear* here if you can recall the moment in this [time] video clip?
	• What did you *read* here if you can recall the moment in this [time] video clip?
	• What did you *feel* here if you can recall the moment in this [time] video clip?
	• What did you *do* or *think* at that moment?
	    ○ Why did you *think* this way?
	    ○ Why did you *make that decision* at that moment?
	    ○ Why did you *do* it this way?
	• Did you plan ahead for the next step? What did you expect would happen after this?
		At each segment

Session	Interview Protocol and Interview Question	Timing
Additional Questions	• Did you think about the meaning of what you saw, heard, read, and felt? What was that meaning? • Did you make a decision at that moment? What was it? • Can you explain why you did this? • What was the most challenging experience while you solved the problem?	If needed
Closing	Thank you for your cooperation. Your participation will be important data for me to conduct this research. If you are interested, I will contact you after the research is finished to provide you with the results.	

Detailed Data Collection Protocol for Each Condition

Clinical competence examination (OSCE – Condition 1). As described in the intervention design above, the OSCE as a clinical competence examination was selected for the first condition. A detailed data collection protocol for condition 1 is provided below.

Phase 1 – research participants' performance data collection. For the data collection of a clinical competence examination, a preparatory OSCE, organized by a regional consortium of OSCE examiners, was used. The preparatory OSCE was well designed and prepared by the regional consortium to help 4th year medical students prepare for the National Licensing Examination at the end of the year; therefore, the OSCE was opportunistically selected for data collection for condition 1.

The preparatory OSCE was held on June 4th, 2017 at one of the IUCM-affiliated hospitals in Seoul, South Korea. Half of the 4th year medical students at IUCM visited the hospital to take the preparatory OSCE on that day. A total of 60 students were divided into five groups for the preparatory OSCE, and six clinical presentations were given to the students during the OSCE. Each student was given 10 minutes to complete each clinical presentation. Before students started the preparatory OSCE, they were given 2 minutes to read instructions for a certain clinical presentation. The instruction provided basic information, including the patient's age, symptoms, body temperature, pulse, breathing rate, and blood pressure. The examinees then entered a room to meet a standardized patient who had been trained to present the clinical presentation. The examinees could use the 10 minutes to consult with the standardized patient regarding the presented illness through a history taking, physical examination, and finally to diagnose the illness. After the examinees finished meeting with the standardized patient, test items related to the situation were given to the students, and the students were required to provide answers for these questions within 5 minutes. Because the OSCE is a type of assessment to assess clinical performances in a variety of clinical settings that mimics real-world situations (ACGME & ABMS, 2000; Harden et al, 1975), the preparatory OSCE was also taken in a testing environment similar to a real-world clinical setting.

For the study, a single video recorder was set up in the chest pain station to capture the clinical performances of the research participants without distracting the examinees. Figure 3.3 shows the video-capture setting in the station; a sample screenshot of a video-captured preparatory OSCE is provided in Figure 3.4.

Figure 3.3. Research setting: Videotaping for clinical competence examination (OSCE)

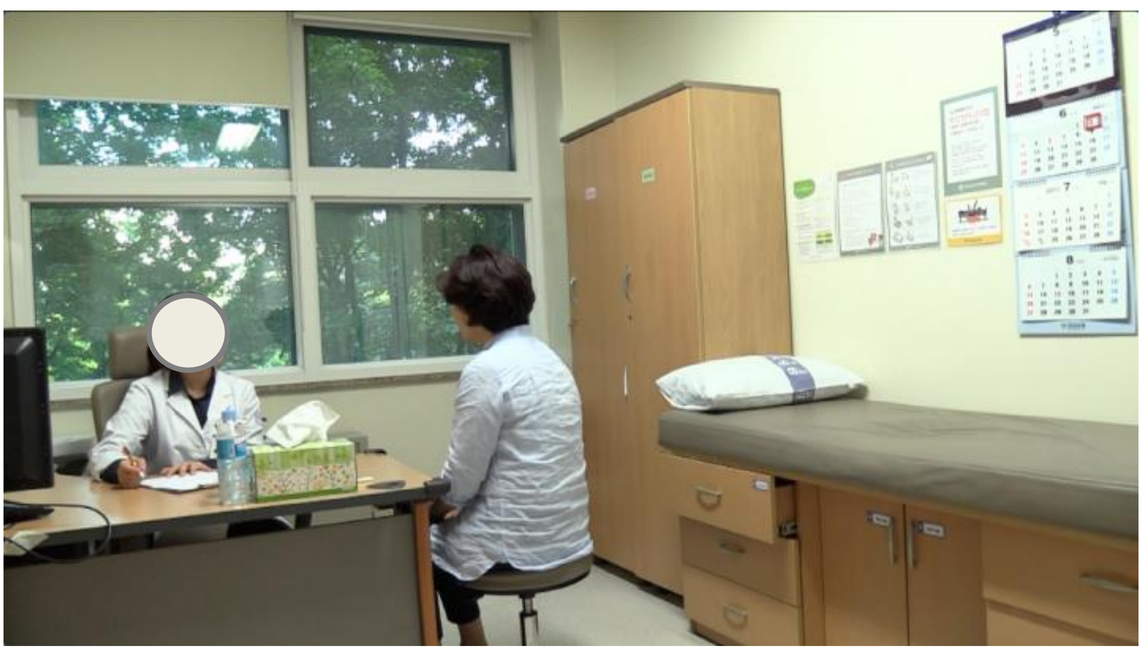


Figure 3.4. A sample screenshot of the preparatory OSCE

Phase 2 – research participants' cognitive activity data collection. The data for phase 2 was collected by stimulated recall interviews. In order to identify the participants' cognitive processes during the OSCE, the captured video data were divided into 20 segments, and each segment was equally fixed for 30 seconds in duration to collect data in a more detailed way.

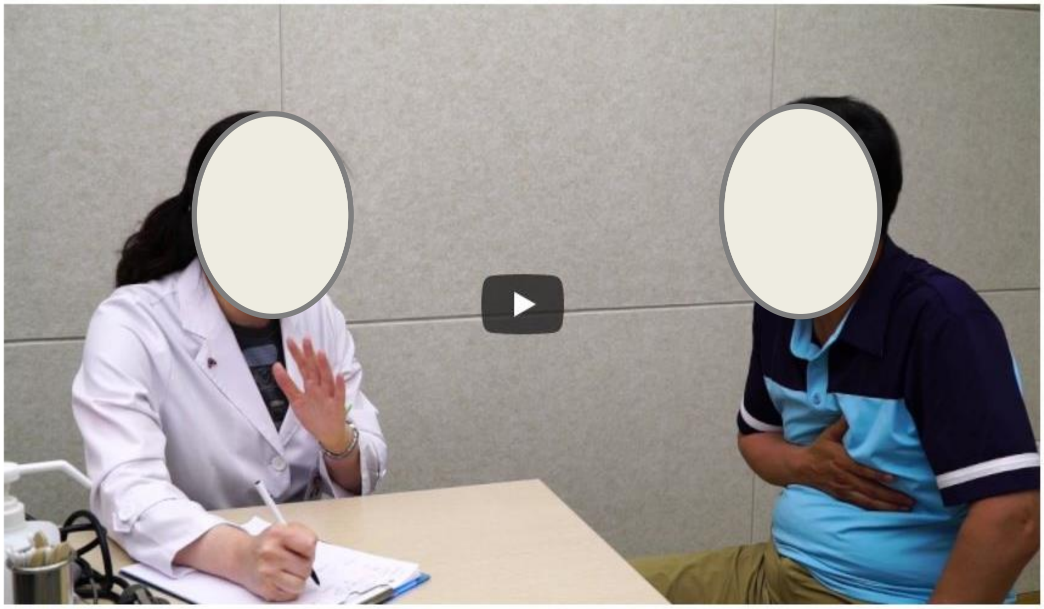
Multimedia case-based assessment (CBA – Condition 2). As described in the intervention design above, the multimedia case-based assessment was selected for the second condition. For condition 2, the CBA was implemented for data collection. According to the research procedures to avoid test-retest biases, this multimedia case-based assessment was given to the research participants one week after they had finished the OSCE (condition 1). The detailed data collection protocol for condition 2 is provided below.

Phase 1 – research participants' performance data collection. The multimedia case-based assessment in this study used chest pain the clinical presentation, and two discussion meetings with subject-matter experts, who are faculty members at IUCM, had been held to decide the video cases for the multimedia case-based assessment in March of 2017.

The multimedia case-based assessment was distributed for data collection using an online survey development system (Qualtrics). Based on the discussion with the subject matter experts, an appropriate time for examinees to finish each test set was decided and technical guidelines were embedded in the test to prevent collecting any missing or misguided answers from the examinees. A sample page of the developed multimedia case-based assessment is presented in Figure 3.5.

Watch the video below by clicking the play button, then answer the questions (12 minutes).

Scene #2



Please describe important clinical cues from the video clip to diagnose the disease.

Please list three possible clinical hypotheses to diagnose the disease and explain your decision.

Hypothesis 1

Hypothesis 2

Hypothesis 3

Please describe additional history taking lists to diagnose the disease and explain your decision(s).

Figure 3.5. A sample page of the developed multimedia case-based assessment

According to their schedules, the students participated in taking the multimedia case-based assessment at IUCM or at one of the affiliated hospitals. The research participants were given paper and a pencil for note taking during the test. All sessions occurred only in a lab setting and were video-taped for data collection and analysis. A desktop (Mac) computer was used by the examinees to take the assessment, and at the same time the desktop computer captured the screen via Camtasia (Mac) 3.0.6 while the desktop computer recorded the examinees themselves through the desktop computer's built-in camera. Before the participants started the test, technical instructions were provided by the researcher to eliminate any distractions in computer usage and taking the test. The captured videos were edited and divided into 20 segments immediately after the participants finished their tests. Figure 3.6 shows the video-capturing setting in the studio with a participant.

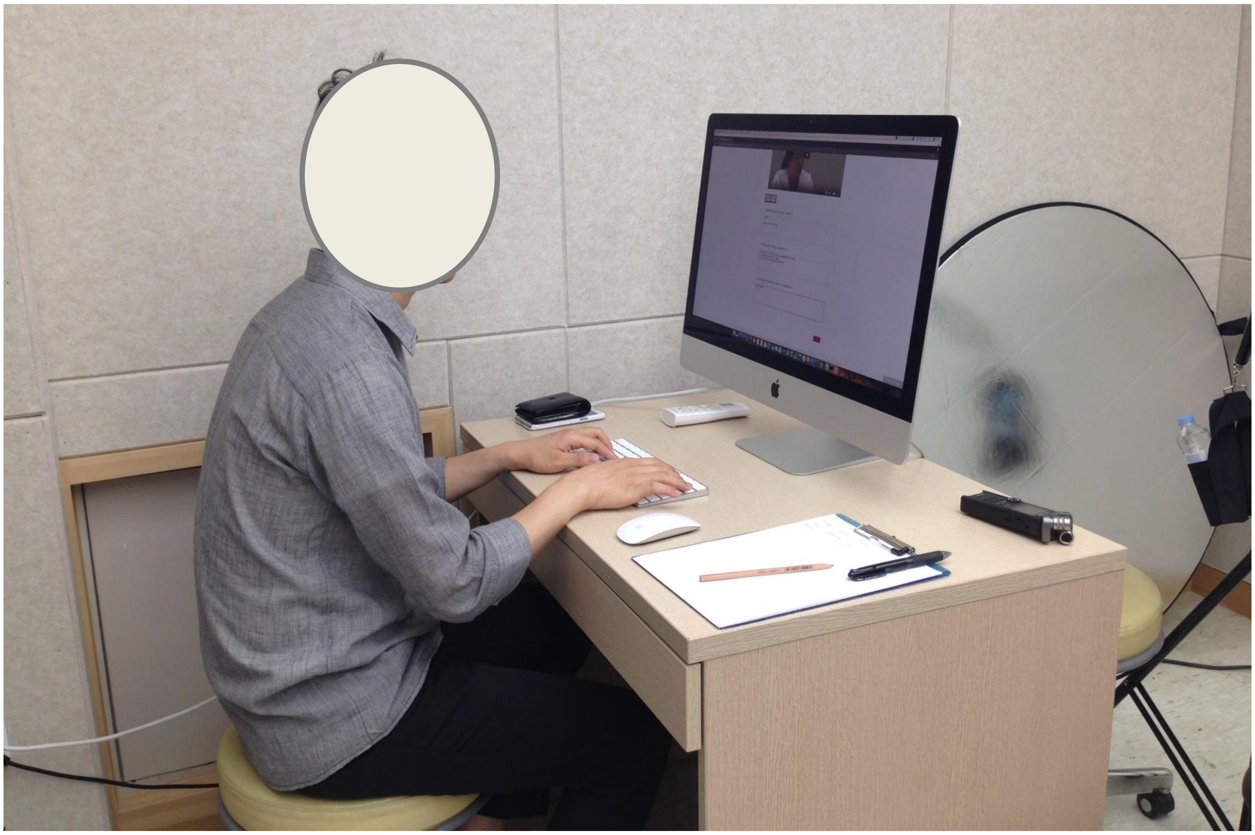


Figure 3.6. A test taking setting for the multimedia case-based assessment

Phase 2 – research participants' cognitive activity data collection. Like condition 1, the data for the condition 2 was also collected by stimulated recall interviews. Video-recorded, retrospective interviews following a stimulated recall protocol were conducted while watching each individual participant's original videos of their multimedia case-based assessment performance.

The total test time for the multimedia case-based assessment was around 60 minutes, depending on each participant's performance, and each segment was fixed for 3 minutes to collect data in a more detailed way. Due to time restrictions, a fast-forward function was used to reduce the time to spent watching the test performance video for the interview. For better results where the participants could recall their memories more vividly (Gass & Mackey, 2000), the stimulated recall interview sessions were conducted immediately after the multimedia case-based assessment's conclusion.

Clinical knowledge and reasoning examination (MEQ – Condition 3). Condition 3 in the study was the clinical knowledge and reasoning examination. As described in the intervention design above, a pre-developed MEQ was implemented as the clinical knowledge and reasoning examination for data collection. According to research procedure to avoid test-retest biases, the MEQ was provided to the research participants one week after they had finished the multimedia case-based assessment (condition 2).

Phase 1 – research participants' performance data collection. For the clinical knowledge and reasoning examination, an IUCM-developed MEQ was modified for this research. Through two discussion meetings with subject matter experts at IUCM, the final version of the MEQ was prepared in the last week of June 2017. The paper-based MEQ test consists of a number of open-ended questions within the context of a patient's information.

Participants read a patient's clinical scenario regarding chest pain, the patient's physical history, and symptoms, then encounter the questions to solve.

Two video recorders were prepared to capture the participants' faces and the MEQ papers they were writing on. Figure 3.7 shows a sample screenshot of the MEQ paper and the video-capture setting in the meeting room. A sample test set of the MEQ used in this study is provided in Figure 3.8.

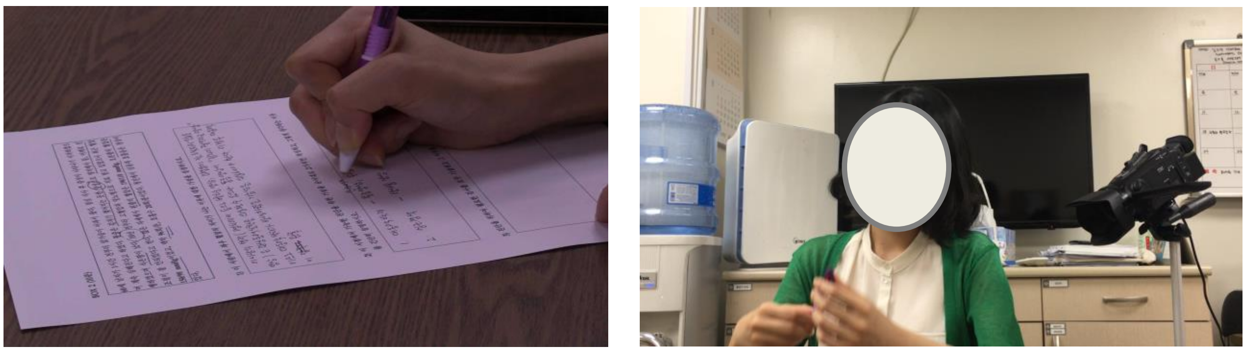


Figure 3.7. Test taking setting of the modified essay question (MEQ)

Phase 2 – research participants' cognitive activity data collection. In the data collection for the clinical knowledge and reasoning examination (MEQ – condition 3), video-recorded, retrospective interviews following the stimulated recall protocol were conducted while watching each individual participant's performance videos during the clinical knowledge and reasoning examination. The total test duration for the MEQ was around 50 minutes depending on each participant's performance, and each segment was fixed for 2.5 minutes to collect data in a more detailed way. Due to time restrictions, a fast-forward function was used to reduce the time spent watching the test performance for the interview. For better results where the participants could recall their memories more vividly (Gass & Mackey, 2000), the stimulated recall interview sessions were conducted immediately after the MEQ's conclusion.

Box 2 (10 minutes)

60-year old, …

CLINICAL SCENARIO

1) **Generate clinical hypotheses to explain the difference of the blood pressures.**

2) **Generate two clinical hypotheses to diagnoses the possible disease in this case, and explain your decision.**

3) **List all necessary physical examinations to diagnose the possible disease in this case, and explain your decision(s).**

Figure 3.8. A sample of the MEQ scenario and questions

Data Analysis

The results were analyzed based on the hypothetico-deductive reasoning (HDR) processes and other cognitive occurrences themes. The OSCE cases were analyzed by the OSCE processes in addition to the HDR process and other cognitive occurrences themes. An initial goal of the analysis was to identify medical students' cognitive processes; therefore, clinical reasoning processes were identified with the hypothetico-deductive reasoning and other cognitive occurrences were discovered from each case. For the analysis, a framework was used for data coding using the HDR process and the OSCE process.

Analysis Framework

Clinical reasoning process. Several clinical reasoning models as cognitive process have been introduced in medical education field: hypothetico-deductive reasoning (Barrows & Tamblyn, 1980; Barrows, 1994; Elstein et al., 1978), pattern recognition (Barrows & Feltovich, 1987), forward and backward reasoning (Patel & Groen, 1986; Arocha et al., 1993), knowledge reasoning integration (Schmidt et al., 1990; Boshuizen & Schmidt, 1992), and intuitive reasoning (Agan, 1987; Rew, 1990; Rew & Barrow, 1987). Among these models, the hypothetico-deductive reasoning model is one of the more suitable models for undergraduate medical students to apply and practice their clinical reasoning for diagnosis (Barrows, 1994; Barrows & Tamblyn, 1980). In order to present medical students' clinical reasoning processes in the three different types of clinical assessments, the hypothetico-deductive reasoning (HDR) model was selected to represent the clinical reasoning processes in this study.

Hypothetico-Deductive Reasoning process. Hypothetico-deductive reasoning (HDR) as clinical reasoning refers to a form of means-end analysis used to narrow diagnostic hypotheses to reduce the distance between the point where the problem solver is and where the problem solver would like to be using clinical inquiry and data (Elstein et al., 1978). This approach generates hypotheses from clinical data and knowledge and tests the hypotheses through clinical inquiries (Higgs & Jones, 2000). In this study, the following HDR process was incorporated and used to identify the participants' clinical reasoning processes (Barrows & Tamblyn, 1980; Barrows, 1994; Hardin et al., 2002; Ju & Choi, 2017).

(1) Problem Framing

(2) Hypothesis Generation

(3) Inquiry Strategy

(4) Data Analysis or Synthesis

(5) Diagnostic Decision and Explanation

(6) Therapeutic Decision and Treatment Options

The Objective Structured Clinical Examination (OSCE) process. There are various clinical presentations that can be assessed by the OSCE, and each clinical presentation may have different steps or processes to be solved. However, the OSCE was developed based on general clinical procedures for diagnosis and used with questions related to clinical findings and interpretations (Harden et al., 1975). Therefore, the OSCEs are organized with a general procedure for diagnosis, and medical students are taught to follow the procedure in solving clinical problems in the OSCE.

There are two main parts in the procedure of the OSCE: Diagnosis and patient education and counseling. In the diagnosis part, history-taking and physical examination are processes to obtain more information and analyze data. The history-taking includes some or all of the following elements: chief complaint, history of present illness, review of systems, and medical history, family history, and social history. After finishing the physical examination, a doctor evaluates all the possibilities from the collected data and decides on an initial diagnosis. Included in the patient education and counseling part is an explanation of the diagnosis and treatment options.

The entire process of the OSCE is listed below, which was also used in this study.

(1) Diagnosis: History-taking – Chief Complaint

(2) Diagnosis: History-taking – History of Patient Illness

(3) Diagnosis: History-taking – Review of Systems

(4) Diagnosis: History-taking – Medical, Family, and Social History

(5) Diagnosis: Physical Examination

(6) Patient Education and Counseling: Explanation of Diagnosis

(7) Patient Education and Counseling: Treatment Options

Other Cognitive Occurrences. In order to identify medical students' cognitive processes in solving clinical problems, other cognitive occurrences were also detected if the participants' cognitive processes did not fall under clinical reasoning when solving clinical problems. Other cognitive occurrences refer to reasoning that actually may not occur when a doctor sees patients and/or not authentic clinical problem-solving, the kind that may not necessarily occur when encountering patients and occur only in certain testing contexts. Due to each clinical assessment affording varied other cognitive occurrences, the other cognitive occurrences were processed separately. The detailed protocol of recording the emerging other cognitive occurrences themes is described in the next section.

Data Analysis Protocol

Data analysis in this study was conducted with the following steps. After the unit of analysis was selected, the interview data were transcribed and classified based on a naturalistic decision-making model. The classified data were coded with the hypothetico-deductive reasoning process and the OSCE process, and any un-coded data were classified under other cognitive occurrences. Then, the coded clinical reasoning and other cognitive occurrences results were chronologically reorganized to reconstruct the participants' cognitive occurrences. Lastly, a graphical representation was provided to observe the reasoning processes.

The unit of analysis. In this study, the unit of analysis was the smallest phrases or sentences representing a meaningful cognitive occurrence. All the phrases or sentences of the narratives were derived from the stimulated recall interviews after each clinical assessment. With this unit of analysis, data were analyzed in the following phases.

Phase 1 – Data classification. After transcribing the interviews from the video and audio files, the transcribed data were transferred to an Excel sheet. Then, the data were organized according to the unit of analysis with reference to a naturalistic decision making (NDM) model. The NDM has been used to describe "how people actually make decisions in real-world settings" (Klein, 2008, p. 456), and it provides how people make decisions and take actions based on cognitive processes (Zsambok & Klein, 1997). The content for the NDM was adapted from Crandall, Klein, & Hoffman (2006), Hoffman, Crandall, & Shadbolt (1998), and Klein, Calderwood, & MacGregor (1989). The following NDM content was used to divide the cognitive incidents: cue identification, cue interpretation, information, analogs, knowledge, action, appraisal, goal setting, anticipation, metacognition, situation assessment, situation procedure, previous experience, mental models, decision-making, rationale, plans, expected results, results, self-reflection, and lesson learned. Each data was divided into several chunks based on the NDM content, and decision content was added to explain and support the details of the NDM model. Each chunk was named as an occurrence and numbered to avoid being mixed for later analysis. A sample of data table with the occurrence, NDM content, and decision content columns is provided in Figure 3.9.

Transcript (English)	Original Number	Occurrence (English)	Naturalistic Decision Making	Decision Content
First, the reason why I asked the patient about more details of her chest pain was to get more information from open-ended questions. But because she didn't give me much information, I asked closed questions one-by-one. Initially [I asked about] location, duration [how long is each painful episode], timeframe [when the pain began]... Because these can be big hints to a diagnosis, I asked these questions first. So, I asked: "Can you point to where it hurts?" instead of asking "Where does it hurt?" in order to get a more accurate location of the pain. At that time the patient pointed to the middle of her chest and said: "It hurts here." After I found the location of the chest pain, I was thinking of what I needed to ask next for the differential diagnosis. There are many possibilities for having pain in the center of the chest; for example, heart disease... So I thought I needed to keep my [diagnosis] options open. For example, although she pointed at the middle of her chest, I was thinking was it really the front part of her chest or could it be the back part? Also, it may not even be heart disease. So, I thought I had to keep my options open.	33	First, the reason why I asked the patient about more details of her chest pain	Rationale	Reason of the Pain
	34	was to get more information from open-ended questions.	Rationale	Open Question
	35	But because she didn't give me much information,	Information	Information
	36	I asked closed questions one-by-one.	Rationale	Close Question
	37	Initially [I asked about] location, duration [how long is each painful episode], timeframe [when the pain began]...	Cue Identification	Location, duration, & timeframe
	38	Because these can be big hints to a diagnosis,	Rationale	Diagnosis
	39	I asked these questions first.	Cue Identification	Location, duration, & timeframe
	40	instead of asking "Where does it hurt?" in order to get a more accurate location of the pain.	Rationale	Location
	41	So, I asked: "Can you point to where it hurts?"	Cue Identification	Location
	42	At that time the patient pointed to the middle of her chest and said: "It hurts here."	Cue Interpretation	Location
	43	After I found the location of the chest pain,	Information	Location
	44	I was thinking of what I needed to ask next for the differential diagnosis	Goal Setting	Diagnosis
	45	There are many possibilities for having pain in the center of the chest; for example, heart disease...	Appraisal	Hypothesis
	46	So I thought I needed to keep my [diagnosis] options open.	Decision-making	Hypothesis
	47	For example, although she pointed at the middle of her chest, I was thinking was it really the front part of her chest or could it be the back part?	Rationale	Location
	48	Also, it may not even be heart disease.	Rationale	Hypothesis
	49	So, I thought I had to keep my options open.	Decision-making	Hypothesis

Figure 3.9. Data Analysis Table (Phase 1)

Phase 2 – Data coding. Each chunk of the cognitive occurrences derived from the original data was coded with reference to the hypothetico-deductive reasoning process as clinical reasoning in phase 2. In particular, case 1 and 2 were also coded with reference to the OSCE process, and the OSCE process was used for distinguishing sections of the OSCE. A sample of data table with HDR process columns is presented in Figure 3.10.

Original Number	Occurrence (English)	Naturalistic Decision Making	Decision Content	HDR Coding (2nd)	OSCE Process (2nd)
33	First, the reason why I asked the patient about more details of her chest pain	Rationale	Reason of the Pain	Inquiry Strategy	History-taking: HPI
34	was to get more information from open-ended questions.	Rationale	Open Question	Inquiry Strategy	History-taking: HPI
35	But because she didn't give me much information,	Information	Information	Data Analysis or Synthesis	History-taking: HPI
36	I asked closed questions one-by-one.	Rationale	Close Question	Inquiry Strategy	History-taking: HPI
37	Initially [I asked about] location, duration [how long is each painful episode], timeframe [when the pain began]...	Cue Identification	Location, duration, & timeframe	Inquiry Strategy	History-taking: HPI
38	Because these can be big hints to a diagnosis,	Rationale	Diagnosis	Inquiry Strategy	History-taking: HPI
39	I asked these questions first.	Cue Identification	Location, duration, & timeframe	Inquiry Strategy	History-taking: HPI
40	instead of asking "Where does it hurt?" in order to get a more accurate location of the pain.	Rationale	Location	Inquiry Strategy	History-taking: HPI
41	So, I asked: "Can you point to where it hurts?"	Cue Identification	Location	Inquiry Strategy	History-taking: HPI
42	At that time the patient pointed to the middle of her chest and said: "It hurts here."	Cue Interpretation	Location	Data Analysis or Synthesis	History-taking: HPI
43	After I found the location of the chest pain,	Information	Location	Data Analysis or Synthesis	History-taking: HPI
44	I was thinking of what I needed to ask next for the differential diagnosis.	Goal Setting	Diagnosis	Data Analysis or Synthesis	History-taking: HPI
45	There are many possibilities for having pain in the center of the chest; for example, heart disease...	Appraisal	Hypothesis	Hypothesis Generation	History-taking: HPI
46	So I thought I needed to keep my [diagnosis] options open.	Decision-making	Hypothesis	Hypothesis Generation	History-taking: HPI
47	For example, although she pointed at the middle of her chest, I was thinking was it really the front part of her chest or could it be the back part?	Rationale	Location	Hypothesis Generation	History-taking: HPI
48	Also, it may not even be heart disease.	Rationale	Hypothesis	Hypothesis Generation	History-taking: HPI
49	So, I thought I had to keep my options open.	Decision-making	Hypothesis	Hypothesis Generation	History-taking: HPI

Figure 3.10. Data Analysis Table (Phase 2)

Phase 3 – Un-coded data classification. After all cognitive occurrences were coded as clinical reasoning with the HDR process, un-coded data from the original data were classified under other cognitive occurrences. In phase 3, any obvious cognitive occurrences that was not related to the testing situation, any duplicated answers, any post-assessment, such as what they should have done or could have done differently, instead of revealing what they were actually thinking during the moment they were doing the examination, and any simple confirming answers, such as "yes" or "that's right" were eliminated for the analysis. For the other cognitive occurrences coding, the un-coded chunks were initially coded according to the contents of cognitive occurrences; then the data were re-categorized according to emerging themes through axial coding. A sample of a data table with other cognitive occurrences is presented in Figure 3.11.

Original Number	Occurrence (English)	Naturalistic Decision Making	Decision Content	HDR Coding (2nd)	OSCE Process (2nd	Other Cognitive Occurrences	Note
185	but I was thinking [about time] in my mind. If I finish the history taking, already around 5 minutes is gone.	Self-reflection	Time			Off-protocol Behavior	
186	If a patient is talkative while I am doing history taking so that I am delayed,	Rationale	Time Delay			Off-protocol Behavior	
187	I do physical examination more quickly. I need to omit unnecessary things…	Decision-making	Time			Off-protocol Behavior	
188	Because there is a grading criteria made by the Health Personnel Licensing Examination Institute,	Information	Grading Criteria			Point-seeking/hunting	
189	and the Patient Education and Counselling section is a higher portion [of the grade] than I thought....	Information	Grading Criteria			Point-seeking/hunting	
190	Because when I read the patient-physician interaction section [of the guidebook], asking the patient if he/she has any questions or asking if he/she does not undertand something are part of the patient education section. Also, telling the patient about a summary of symptoms and making contact with the patient's eyes are mandatory.	Information	Grading Criteria			Point-seeking/hunting	
191	because if I don't do it, I can't get the score. So I have to do it.	Rationale	Grading Criteria			Point-seeking/hunting	
192	So at this point [after finishing history taking], I tried to move to the physical examination quickly.	Plans	Protocol			Off-protocol Behavior	

Figure 3.11. Data Analysis Table (Phase 3)

Phase 4 – Data re-organization with chronological orders. All cognitive occurrences were coded based on the clinical reasoning process and the themes of other cognitive occurrences. Then, the coded data were re-organized according to the chronology of each participant's performance in order to reconstruct the interpreted narrative of cognitive occurrences. A sample of a data table from the reorganized cognitive occurrences in chronological order is provided in Figure 3.12.

Narrative Number	Original Number	Reconstructed Narrative of Occurrences (EN)	Naturalistic Decision Making (of Narrative Occurrences)	Decision Content	HDR Coding	OSCE Coding	Other Cognitive Occurrences	Note
252	289	But this time I omitted tapping and touching,	Decision-making	Physical Examination	Inquiry Strategy	Physical Examination		
253	290	because I thought the possibility of it not being respiratory disease was higher than other [conditions]...	Rationale	Hypothesis	Data Analysis or Synthesis	History-taking: ROS		
254	293	For the respiratory symptoms.... I had already asked quite a lot of questions to do differential diagnosis for the respiratory system.	Cue Identification	History-taking	Inquiry Strategy	History-taking: ROS		
255	294	So I thought I was finished with the respiratory system,	Anticipation	Diagnosis	Data Analysis or Synthesis	History-taking: ROS		
256	291	Much time had already passed, and I didn't think I needed to do it.	Appraisal	Physical Examination	Inquiry Strategy	Physical Examination		
257	292	so I skipped those two: tapping and touching.	Goal Setting	Physical Examination			Off-protocol Behavior	
258	295	and I didn't do tapping and touching because of the time limit.	Self-reflection	Time Pressure			Off-protocol Behavior	
259	296	But I was not sure about criteria of the checklist [grade book] so I thought if I did not do it... the mission was doing physical examination for chest pain but if I did not do tapping and touching... what if I can't get a score for this?	Self-reflection	Grading Criteria			Point-seeking/hunting	
260	297	Then I had just wasted my time...	Self-reflection	Grading Criteria			Off-protocol Behavior	
261	298	First, I thought I could do differential diagnosis with only the stethoscope for the heart and the lung of the disease I was thinking about.	Self-reflection	Physical Examination	Inquiry Strategy	Physical Examination		
262	299	Then the next step was asking the patient, when I pressed here, if the patient felt pain,	Plans	Physical Examination	Inquiry Strategy	Physical Examination		
263	300	[to see] if there were any pressure points.	Plans	Physical Examination	Inquiry Strategy	Physical Examination		

Figure 3.12. Data Analysis Table (Phase 4)

Phase 5 – Data visualization. After reorganizing the chronology of the cognitive occurrences, a color-coded graphical representation was developed for each clinical assessment to observe the patterns of clinical reasoning along with other cognitive occurrences in solving clinical problems. Each of the hypothetico-deductive reasoning processes have a unique color theme, and the other cognitive occurrences have another unique color theme to be distinguished in the representative process map. In this study, the HDR processes were presented by gradations of blue colors, and the other cognitive process were displayed as the color red. A sample of a data table with the graphical representation for the cognitive patterns is shown in Figure 3.13.

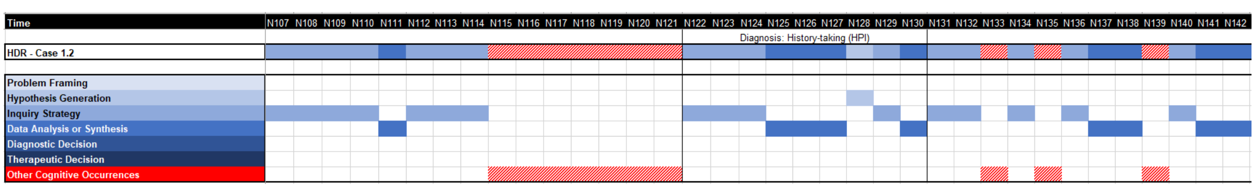


Figure 3.13. Data Analysis Table (Phase 5)

Validity and Reliability

Creswell (2007) defined both validity and reliability in qualitative research. Qualitative validity refers to when "the researcher checks for the accuracy of the findings by employing certain procedures, while qualitative reliability indicates that the researcher's approach is consistent across different researchers and different projects" (Creswell, 2007, p. 201). To fulfill the accuracy of the findings and the consistency of the analysis procedures, the data analysis was conducted with a validation procedure through inter-rater reliability. By acquiring high inter-rater reliability, the results based on reliable and valid coding schemes were not affected to a large extent by the researcher (Gwet, 2014).

The inter-rater reliability process was conducted with a medical doctor from the research site (IUCM) in two phases. First, all cognitive occurrences were coded by the researcher first, then the expert validated the coded data set. Second, a negotiation process was conducted for any unmatched chunks of cognitive occurrences. The inter-rater reliability results are provided in Table 3.2.

Table 3.2.
Inter-rater reliability result

Condition	Case		Initial Agreement	Second Agreement
Condition 1 - OSCE	Case 1.1	HDR Process	74.76%	100.00%
		OSCE process	69.90%	100.00%
		Other cognitive occurrences	67.86%	100.00%
	Case 1.2	HDR Process	78.57%	100.00%
		OSCE process	85.71%	100.00%
		Other cognitive occurrences	74.75%	100.00%
Condition 2 - CBA	Case 2.1	HDR Process	80.95%	100.00%
		Other cognitive occurrences	93.68%	100.00%
	Case 2.2	HDR Process	86.11%	100.00%
		Other cognitive occurrences	87.10%	100.00%
Condition 3 - MEQ	Case 3.1	HDR Process	82.88%	100.00%
		Other cognitive occurrences	86.30%	100.00%
	Case 3.2	HDR Process	97.54%	100.00%
		Other cognitive occurrences	97.56%	100.00%

Through two-step inter-rater reliability, validation procedures were processed to enhance the accuracy of research findings with the subject-matter expert by minimizing researcher bias.

Moreover, spending prolonged time on the research site to increase trustworthiness, authenticity, and credibility of the study during the data collection helped the validation (Creswell, 2007; Maxwell, 2013).

In addition, all research procedures and protocols in terms of the case study were documented to determine that the research approaches are reliable (Creswell, 2007; Yin, 2009). In this study, all interview transcripts were double-checked to avoid obvious mistakes (Creswell, 2007), and analysis sheets were individually documented to prevent data from being mixed up.

Methodological Limitations

A case study method has limitations regarding reliability, validity, and generalizability (Merriam, 1998). However, as Stake (1995) highlighted, one of the important purposes of a case study is not generalization but understanding the case itself. In this study, I tried to identify medical students' cognitive processes while they solve clinical problems in different types of clinical assessments to gain an initial understanding of how medical students actually think in solving clinical problems. As a result, each individual case had unique characteristics. Therefore, the research findings were made through the uniqueness of each case result. Furthermore, to minimize the issues of reliability and validity in the case study, inter-rater reliability was used as mentioned in the previous section.

With the research design, I decided that each participant would experience each type of clinical assessments in one-week intervals to remove the factor that the results could be influenced or different by the participant's constantly developing knowledge and skill. Then I used three different diagnoses from the same clinical presentation to identify the cognitive processes in solving clinical problems. One limitation here is that the level of difficulty between each test set may be different if the final diagnoses are different. Moreover, even though the final diagnoses are different for each assessment, there may be the presence of a learning effect from the previous test within such a short time interval, as well as possible learning effects from the follow up interviews after each test. An additional possibility is that the students could have reasoned that each assessment would involve a different final diagnosis; thus, by process of elimination, they would know which diagnoses would not be on the third assessment. Another limitation is that I used hypothetico-deductive reasoning as clinical reasoning to identify the cognitive processes in this study; because the hypothetico-deductive reasoning is used when physicians encounter unfamiliar clinical cases (Barrows, 1985, 1994; Barrows & Tamblyn, 1980), if the participants was too familiar with clinical cases, even if the final diagnosis of each assessment was different, the participants' clinical reasoning may not be identified as hypothetico-deductive reasoning.

Regarding the stimulated recall interviews, Yin (2003) suggested six major sources of data for case studies: documents, archival records, interviews, direct observation, participant observation, and physical artifacts. For this multiple case study, interviews, participant observations, and documents were selected as data collection methods that have been most commonly used in case studies (Bassey, 1999; Creswell, 1998; Merriam, 1998; Stake, 2010). In this study, however, the interviews were used as a major data collection method, which is a limitation of this research.

The stimulated recall method has limitations as well. The method requires in-depth interviews with the participants in this study. However, the interviews may not extract all information based on the cognitive events, and the information relies solely on the participants' narratives. Although their time and effort are one of the key aspects to collect good data, the participants may be fallible in many ways, such as missing the details of some events, confusing one event for another, or have limited information about what happened or why (Crandall et al., 2006).

CHAPTER 4

RESULTS

Data regarding medical students' reconstructed narratives of cognitive occurrences in three different types of assessments were collected and analyzed. This chapter provides the analyzed data results according to the research questions. In the first section, the results from each case are provided, including frequencies and percentages of clinical reasoning, graphical representations of cognitive occurrences, representative quotes of clinical reasoning taken from participants' narratives, frequencies and percentages of other cognitive occurrences, and representative quotes of other cognitive occurrences taken from participants' narratives. Then, comparisons of clinical reasoning and other cognitive occurrences between two participants in each condition are provided to identify the similarities of the cognitive processes are provided. In the second section, the cross-case analysis results that combined the two participants' clinical reasoning and other cognitive occurrences to reveal the differences of the cognitive processes in three different conditions are provided.

Clinical Reasoning Processes (HDR) and Other Cognitive Occurrences from Each Case

Data analysis was conducted to report the results of the first research question, leading to an identification of the clinical reasoning processes and other cognitive occurrences, which afford cognitive aids or distractions to problem-solving, respectively, from each condition. The results are presented here according to the sub-questions under research question 1. In order to answer the research questions, all meaningful cognitive occurrences of participants' narratives were analyzed using hypothetico-deductive reasoning as clinical reasoning processes, and uncoded data (i.e. data not categorizable under hypothetico-deductive reasoning) were classified as other cognitive occurrences. The results of each case are presented respectively, then the results of the clinical reasoning processes and other cognitive occurrences from each condition are provided. Research question 1 is provided below.

Research Question 1. What are the cognitive processes of clinical diagnostic problem-solving in three different types of clinical assessments: Clinical competence examination, multimedia case-based assessment, and clinical knowledge and reasoning examination?

Under the first research question, four sub-research questions were probed to identify medical students' cognitive processes in each condition. This section presents (1) identified frequencies and percentages of clinical reasoning and other cognitive occurrences, (2) graphical representations of cognitive occurrences, (3) representative quotes from participants' narratives for each process of hypothetico-deductive reasoning, (4) themes of other cognitive occurrences, definitions of each theme, and representative quotes from participants' narratives for each case, and (5) results based on the similarities of clinical reasoning and other cognitive occurrences from each condition.

Research Question 1.1. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a clinical competence examination (Objective Structured Clinical Examination)?

Case 1.1 – participant 1 in condition 1 (female; 75th percentile in GPA). The analyzed result of case 1.1 data from a female 4th year medical student, 75th percentile GPA of all 4th year students at IUCM, is provided below.

Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 1.1 were coded as clinical reasoning and other cognitive occurrences. Identified frequencies and percentages of each process of clinical reasoning, analyzed by hypothetico-deductive reasoning model, are presented in Table 4.1.

Table 4.1.
Frequencies and percentages of cognitive occurrences as HDR Processes from case 1.1

Hypothetico-deductive Reasoning	Case 1.1 (Female; 75th percentile)	
	Frequency	Percentage
Clinical Reasoning	**98**	**79.67%**
Problem Framing	5	4.07%
Hypothesis Generation	5	4.07%
Inquiry Strategy	49	39.84%
Data Analysis or Synthesis	25	20.33%
Diagnostic Decision and Explanation	7	5.69%
Therapeutic Decision and Treatment Options	7	5.69%
Other Cognitive Occurrences	**25**	**20.33%**
Total	**123**	**100.00%**

As shown in Table 4.1., 79.67% of 123 cognitive occurrences from case 1.1 were coded for clinical reasoning in the OSCE (condition 1). This result revealed that the most frequent cognitive process in case 1.1 was inquiry strategy (49, 39.84%), followed by data analysis or synthesis (25, 20.33%). 20.49% of other cognitive occurrences from case 1.1 were also detected.

Identified frequencies and percentages of OSCE processes. For case 1.1, which took place in the context of the OSCE, all meaningful cognitive occurrences were coded as clinical reasoning using the OSCE processes as well as hypothetico-deductive reasoning. Identified frequencies and percentages of each process of the OSCE are presented in Table 4.2.

Table 4.2.

Frequencies and percentages of OSCE process cognitive occurrences from case 1.1

	Case 1.1 (Female; 75th percentile)	
	Frequency	Percentage
Clinical Reasoning	**98**	**79.67%**
Diagnosis: History-taking (CC[a])	8	6.50%
Diagnosis: History-taking (HPI[b])	22	17.89%
Diagnosis: History-taking (ROS[c])	21	17.07%
Diagnosis: History-taking (MH, FH, SH[d])	23	18.70%
Diagnosis: Physical Examination	10	8.13%
Education & Counseling: Explanation of Diagnosis	9	7.32%
Education & Counseling: Treatment Options	5	4.07%
Other Cognitive Occurrences	**25**	**20.33%**
Total	**123**	**100.00%**

Note. [a] refers to chief complaint. [b] refers to history of present illness. [c] refers to review of systems. [d] refers to medical history, family history, and social history.

As presented in Table 4.2., 79.67% of clinical reasoning occurred in the OSCE based on the processes of diagnosis and education and counseling. 85.72% of the OSCE process were the diagnosis part, which included history-taking for the history of present illness, review of systems, and medical, family, and social history. This result revealed that this participant spent the most time on taking medical history from the patient during the OSCE.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes regarding hypothetico-deductive reasoning and the OSCE, clinical reasoning and other cognitive occurrences of case 1.1 are presented as a graphical representation map in Figure 4.1.

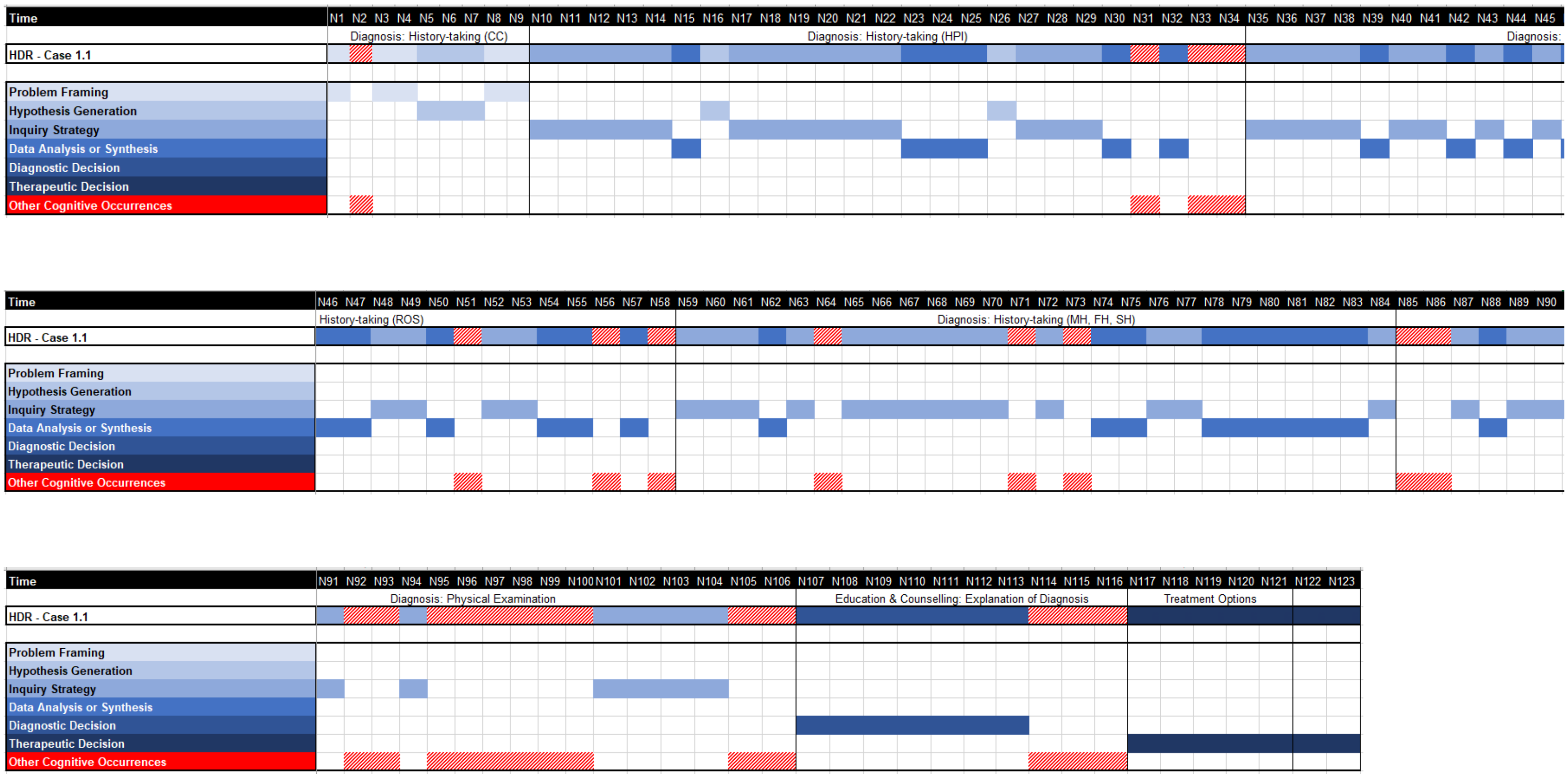


Figure 4.1. A graphical representation of cognitive occurrences from case 1.1

Representative quotes from the participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 1.1 was identified. Representative quotes for each reasoning process are presented in Table 4.3.

Table 4.3.
Representative quotes for clinical reasoning from narratives in case 1.1

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	1	"What I saw was that the patient was calm and didn't look in pain"
	Chief complaint	3	"Then after that, I asked where her pain was in order to find the chief complaint"
	Perceiving a variety of cues: Patient's remarks and responses	4	"The patient said she has a pain in the chest"
Hypothesis Generation	Hypothesis Generation	5	"I was thinking: I'm familiar with chest pain, so digestive, circulatory, respiratory systems, and infections…"
		16	"So, I thought the disease might be close to angina from acute coronary syndrome"
Inquiry Strategy	History-taking	61	"If high blood pressure can be controlled, it will not be a big deal; if it cannot be controlled, there is a disease that can cause complications in the form of high blood pressure, which in turn can cause chest pain. That's why I asked [about the high blood pressure]"
	Physical examination	90	"So, I felt her neck to see if there might be a swollen lymphatic gland"
Data Analysis or Synthesis	Data analysis or synthesis	14	"The patient said: The pain was on the right side of her chest; the duration was not that long, and felt some tightness and stiffness"
	Ongoing summary	82	"From the family history-taking, I found connections to angina"

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	108	"I was thinking I had to tell the patient what the diagnosis was,"
		110	"I intentionally told the patient she was susceptible to the disease [angina] due to high blood pressure and hypercholesterolemia"
Therapeutic Decision & Treatment Options	Therapeutic Decision & Treatment Options	119	"I was explaining about the medication treatments… there are two types of treatment plans: One is surgery, and the other is medication"
		121	"In this case, I thought I needed to explain that if she did not get treatment quickly, she could develop a myocardial infarction"

Themes of other cognitive occurrences, definitions, and representative quotes from the participant's narratives. Uncoded data were classified as other cognitive occurrences. The uncoded data from the original data set was initially coded according to the contents of cognitive occurrences, then the data were re-categorized according to emerging themes. The emerging themes from condition 1, the OSCE, were the following: Test-taking strategy and point-seeking/hunting. Under test-taking strategy, three sub-themes were identified: inattentive action, off-protocol behavior, and educated guessing, and under point-seeking/hunting, two sub-themes were identified: unnecessary behavior and checklist awareness.

Identified frequencies and percentages of each other cognitive occurrences theme from case 1.1 are presented in Table 4.4.

Table 4.4.

Frequencies and percentages of other cognitive occurrences from case 1.1

Other Cognitive Occurrences	Case 1.1 (Female; 75th percentile)	
	Frequency	Percentage
Test-taking Strategy	**19**	**76.00%**
Inattentive Action	9	36.00%
Off-protocol Behavior	7	28.00%
Educated Guessing	3	12.00%
Point-seeking/hunting	**6**	**24.00%**
Unnecessary Behavior	5	20.00%
Checklist Awareness	1	4.00%
Total	**25**	**100.00%**

A total 25 out of 122 cognitive occurrences (20.49%) were detected as other cognitive occurrences from case 1.1, and the most frequent other cognitive occurrences was inattentive action (9, 36.00%) under test-taking strategy. Due to the context of the testing environment in the OSCE, such as physical examination, the examinee pretended to be doing the procedure of physical examination in front of the standard patient, such as thinking about the next steps of action while using a stethoscope on the standardized patient and not actually listening to the sound from the stethoscope.

Definitions and representative quotes of each other cognitive occurrences theme from case 1.1 are explained in Table 4.5.

Table 4.5.

Definition and representative quotes for other cognitive occurrences from narratives in case 1.1

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Inattentive Action	An examinee thinks non-clinically while simultaneously displaying clinical communication or procedure	94	“Honestly, I did not even try to listen to the sound of the patient’s heart because I know the patent was [in real-life] normal”
		98	“… [I didn’t try to listen the sound] because I think of the next step [while I was listening]”

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Off-protocol (Time-limit Influence)	An examinee thinks of the exam's time limits to make decisions, influencing behavior that skips or cuts short clinical communication or procedure	90	"But I don't think doing every procedure is necessary all the time"
Educated Guessing	An examinee thinks of the test's scenario limitations to make decisions and/or to display clinical communication or behavior	54	"Because I have never seen a psychiatric condition [diagnosis] be the correct answer for a test"
	An examinee thinks of a standardized patient's limitations to make decisions and/or to display clinical communication or behavior	94	"Honestly, [doing stethoscope] examination for the heart… standardized patient should be a normal… [heartbeat]"
Point-seeking/hunting	Definition	Narrative Number	Sample Quote
Unnecessary Behavior	An examinee displays clinical communication or procedure believed to be unnecessary or for longer/more than believed necessary	62	"… but I decided to just ask about diabetes. If the patient did have diabetes, then I would have thought about what to do next"
Checklist Awareness	An examinee thinks of a grader's checklist to make decisions and/or to display clinical communication or behavior	83	"I didn't have to do it [checking eyes], but I did anyways"

Case 1.2 – participant 2 in condition 1 (male; 90th percentile in GPA). The analyzed result of case 1.2 data from a male 4th year medical student, 90th percentile GPA of all 4th year students at IUCM, is provided below.

Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 1.2 were coded as clinical reasoning and other cognitive occurrences. Identified frequencies and percentages of each process of clinical reasoning, analyzed by the hypothetico-deductive reasoning model, are presented in Table 4.6.

Table 4.6.
Frequencies and percentages of cognitive occurrences as HDR processes from case 1.2

	Case 1.2 (Male; 90th percentile)	
Hypothetico-deductive Reasoning	Frequency	Percentage
Clinical Reasoning	**209**	**69.44%**
Problem Framing	21	6.98%
Hypothesis Generation	20	6.64%
Inquiry Strategy	87	28.90%
Data Analysis or Synthesis	54	17.94%
Diagnostic Decision and Explanation	16	5.32%
Therapeutic Decision and Treatment Options	11	3.65%
Other Cognitive Occurrences	**92**	**30.56%**
Total	**301**	**100.00%**

As shown in Table 4.6., 69.44% of 301 cognitive occurrences from case 1.2 were coded for clinical reasoning in the OSCE (condition 1). Similar to case 1.1's result, the most frequent cognitive process in case 1.2 was inquiry strategy (87, 28.90%), followed by data analysis or synthesis (54, 17.94%). 30.56% of other cognitive occurrences from case 1.2 were also detected.

Identified frequencies and percentages of OSCE processes. For case 1.2, which occurred in the context of the OSCE, all meaningful cognitive occurrences were also coded as clinical reasoning using OSCE processes. The identified frequencies and percentages of each process of the OSCE are presented in Table 4.7.

Table 4.7.

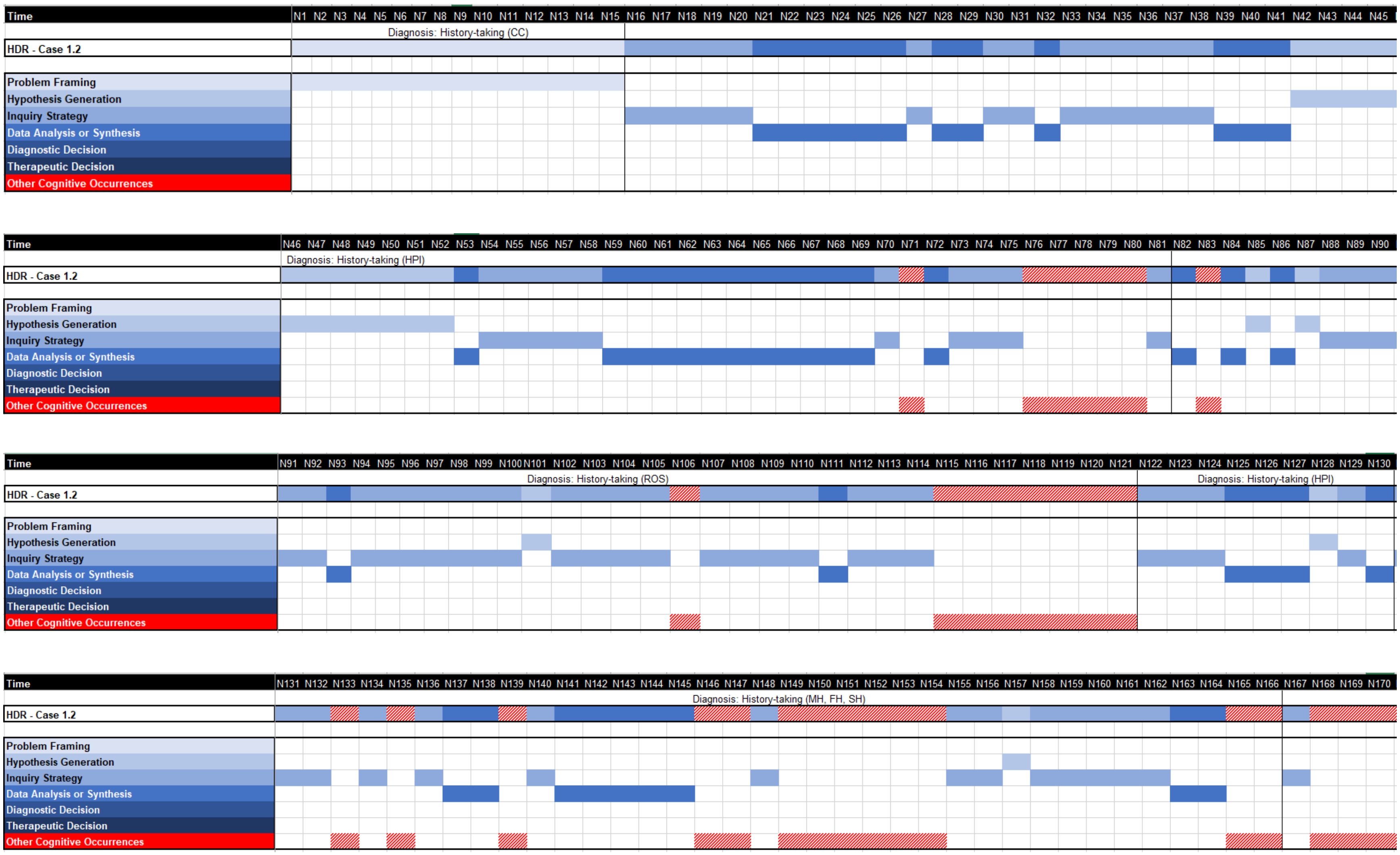
Frequencies and percentages of OSCE process cognitive occurrences from case 1.2

	Case 1.2 (Male; 90th percentile)	
	Frequency	Percentage
Clinical Reasoning	**209**	**69.44%**
Diagnosis: History-taking (CC[a])	15	4.98%
Diagnosis: History-taking (HPI[b])	72	23.92%
Diagnosis: History-taking (ROS[c])	33	10.96%
Diagnosis: History-taking (MH, FH, SH[d])	23	7.64%
Diagnosis: Physical Examination	39	12.96%
Education & Counseling: Explanation of Diagnosis	16	5.32%
Education & Counseling: Treatment Options	11	3.65%
Other Cognitive Occurrences	**92**	**30.56%**
Total	**301**	**100.00%**

Note. [a] refers to chief complaint. [b] refers to history of present illness. [c] refers to review of systems. [d] refers to medical history, family history, and social history.

As presented in Table 4.7., 69.44% of 301 clinical reasoning occurred in the OSCE belonged to the processes of diagnosis and education and counseling. 87.08% of the OSCE process were the diagnosis part, with history-taking for the history of the present illness comprising the largest portion. Review of systems, medical, family, and social history, and physical examination parts followed in frequency. This result also revealed that this participant, like the participant in case 1.1, spent the most time on taking medical history from the patient during the OSCE.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes regarding hypothetico-deductive reasoning and the OSCE, clinical reasoning and other cognitive occurrences of case 1.2 are presented as a graphical representation in Figure 4.2.


Time
N1 N2 N3 N4 N5 N6 N7 N8 N9 N10 N11 N12 N13 N14 N15 N16 N17 N18 N19 N20 N21 N22 N23 N24 N25 N26 N27 N28 N29 N30 N31 N32 N33 N34 N35 N36 N37 N38 N39 N40 N41 N42 N43 N44 N45
Diagnosis: History-taking (CC)
HDR - Case 1.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N46 N47 N48 N49 N50 N51 N52 N53 N54 N55 N56 N57 N58 N59 N60 N61 N62 N63 N64 N65 N66 N67 N68 N69 N70 N71 N72 N73 N74 N75 N76 N77 N78 N79 N80 N81 N82 N83 N84 N85 N86 N87 N88 N89 N90
Diagnosis: History-taking (HPI)
HDR - Case 1.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N91 N92 N93 N94 N95 N96 N97 N98 N99 N100 N101 N102 N103 N104 N105 N106 N107 N108 N109 N110 N111 N112 N113 N114 N115 N116 N117 N118 N119 N120 N121 N122 N123 N124 N125 N126 N127 N128 N129 N130
Diagnosis: History-taking (ROS)
Diagnosis: History-taking (HPI)
HDR - Case 1.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N131 N132 N133 N134 N135 N136 N137 N138 N139 N140 N141 N142 N143 N144 N145 N146 N147 N148 N149 N150 N151 N152 N153 N154 N155 N156 N157 N158 N159 N160 N161 N162 N163 N164 N165 N166 N167 N168 N169 N170
Diagnosis: History-taking (MH, FH, SH)
HDR - Case 1.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences

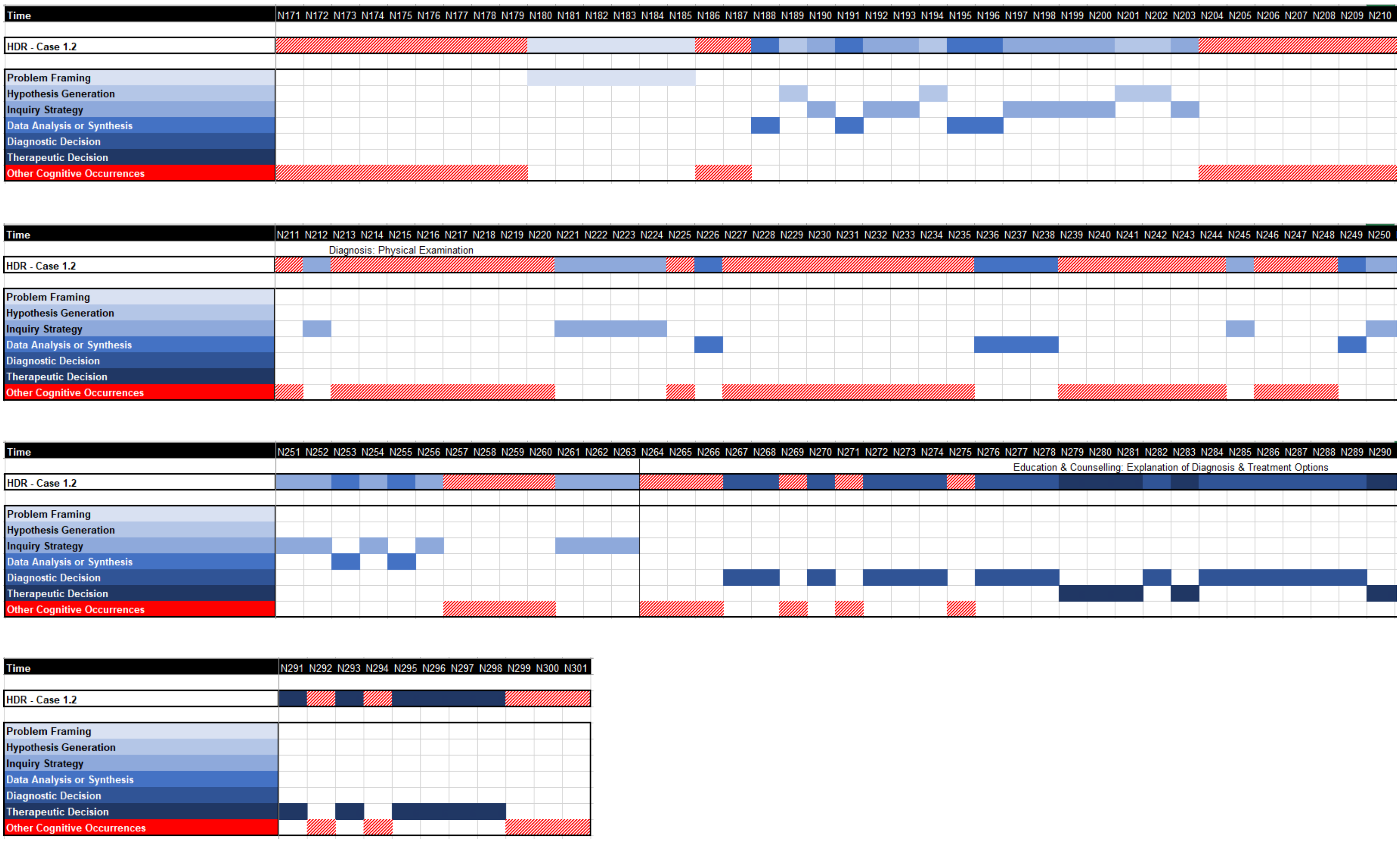


Figure 4.2. A graphical representation of cognitive occurrences from case 1.2

Representative quotes from the participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 1.2 was identified. Representative quotes of each reasoning process are presented in Table 4.8.

Table 4.8.
Representative quotes for clinical reasoning from narratives in case 1.2

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	4	"First, I looked once at the patient's face; I tried to guess the patient's age, patient is female…"
	Chief complaint	10	"and the patient's symptoms of chest pain may not be serious, but there are many urgent conditions I may need to diagnose"
	Perceiving a variety of cues: Patient's remarks and responses	13	"First, what I asked the patient was the reason for the visit. The reason I asked this was because I wanted to hear the patient tell me it was chest pain and to tell me where the pain was bothering her the most"
Hypothesis Generation	Hypothesis Generation	42	"There are many possibilities for having pain in the center of the chest, for example, heart disease…"
		49	"because there are many diseases related to chest pain, if you ask specifically about the pain's frequency and duration, you can rule out many related illnesses"
Inquiry Strategy	History-taking	56	"After that I asked about characteristics of the pain; for example, if it was a stabbing pain, or a squeezing pain…"
	Physical examination	261	"First, I thought I could do differential diagnosis with only the stethoscope for the heart and the lung of the disease I was thinking about"

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Data Analysis or Synthesis	Data analysis or synthesis	196	"the patient might actually be mistaken; there is an arrhythmia where people don't feel any symptoms"
	Ongoing summary	145	"At this point, I was thinking that this diagnosis would pretty much be stable angina"
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	277	"Yes, for the patient education, I always mention the possible diagnosis first, simply, this is your diagnosis"
		288	"so, I explained simply that if the heart is abnormal, the enzyme level will be high, so the test will have to be done"
Therapeutic Decision & Treatment Options	Therapeutic Decision & Treatment Options	280	"So, I needed to explain that this was not a simple situation for the patient"
		297	"because the patient should not postpone the test, especially for this angina case"

Themes of other cognitive occurrences, definitions, and representative quotes from the participant's narratives. Uncoded data were classified as other cognitive occurrences. The uncoded data from the original data set was initially coded according to the contents of cognitive occurrences and then the data were re-categorized according to emerging themes. The emerging themes from condition 1, OSCE, were the following: Test-taking strategy and point-seeking/hunting. Under test-taking strategy were found: inattentive action, off-protocol behavior, and educated guessing, and under point-seeking/hunting were found: unnecessary behavior and checklist awareness.

The identified frequencies and percentages of each other cognitive occurrences theme from case 1.2 are presented in Table 4.9.

Table 4.9.
Frequencies and percentages of other cognitive occurrences from case 1.2

	Case 1.2 (Male; 90th percentile)	
Other Cognitive Occurrences	Frequency	Percentage
Test-taking Strategy	**37**	**40.22%**
Inattentive Action	1	1.09%
Off-protocol Behavior	25	27.17%
Educated Guessing	11	11.96%
Point-seeking/hunting	**55**	**59.78%**
Unnecessary Behavior	2	2.17%
Checklist Awareness	53	57.61%
Total	**92**	**100.00%**

A total 92 out of 301 cognitive occurrences (30.56%) were detected as other cognitive occurrences from case 1.2, and the most frequent other cognitive occurrences was checklist awareness (53, 57.61%). The examinee was acutely aware of solving test questions in the examination situation instead of natural problem-solving. As a result, more than half of the other cognitive occurrences in case 1.2 were detected as checklist awareness where the participant focused on seeking hints from the test and/or hunting for the right answer. Also, 27.17% of the other cognitive occurrences were coded as off-protocol behavior, which are behaviors where the examinee would skip or shorten an important procedure due to the time constraint.

Definitions and representative quotes of each other cognitive occurrences theme from case 1.2 are explained in Table 4.10.

Table 4.10.
Definition and representative quotes for other cognitive occurrences from narratives in case 1.2

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Inattentive Action	An examinee thinks non-clinically while simultaneously displaying clinical communication or procedure	243	“I think this is like ‘pretend to do it’”
Off-protocol (Time-limit Influence)	An examinee thinks of the exam’s time limits to make decisions, influencing behavior that skips or cuts short clinical communication or procedure	186	“For the history-taking… I didn’t do many things… somewhat intentionally…”
Educated Guessing	An examinee thinks of the test’s scenario limitations to make decisions and/or to display clinical communication or behavior	248	“because [I assumed that] doing physical examinations in this context means there [should be] no problem with the patient”
	An examinee thinks of a standardized patient’s limitations to make decisions and/or to display clinical communication or behavior	247	“Always, I didn’t think there’s a problem with the patient in the physical exam”
Point-seeking/hunting	**Definition**	**Narrative Number**	**Sample Quote**
Unnecessary Behavior	An examinee displays clinical communication or procedure believed to be unnecessary or for longer/more than believed necessary	147	“so after that I started to ask some more questions [for the exam only]”
Checklist Awareness	An examinee thinks of a grader’s checklist to make decisions and/or to display clinical communication or behavior	80	“so, I somewhat thought I needed to meet those test requirements quickly and move on”

Research Question 1.2. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a multimedia case-based assessment?

Regarding research question 1.2, the same students from the previous cases participated in the multimedia case-based assessment, which is the second condition of this study.

Case 2.1 – participant 1 in condition 2 (female; 75th percentile in GPA). The analyzed result of case 2.1 data from the same female 4th year medical student, 75th percentile GPA of all 4th year students at IUCM, is provided below.

Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 2.1 were coded as clinical reasoning and other cognitive occurrences. The identified frequencies and percentages of each process of clinical reasoning, based on hypothetico-deductive reasoning model, are presented in Table 4.11.

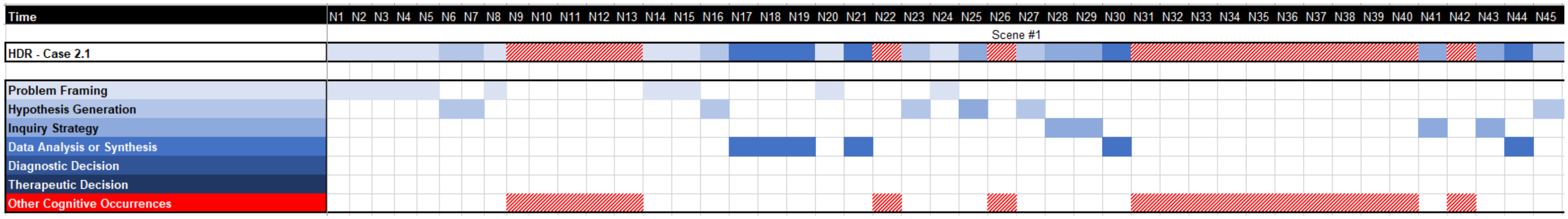
Table 4.11.
Frequencies and percentages of cognitive occurrences as HDR processes from case 2.1

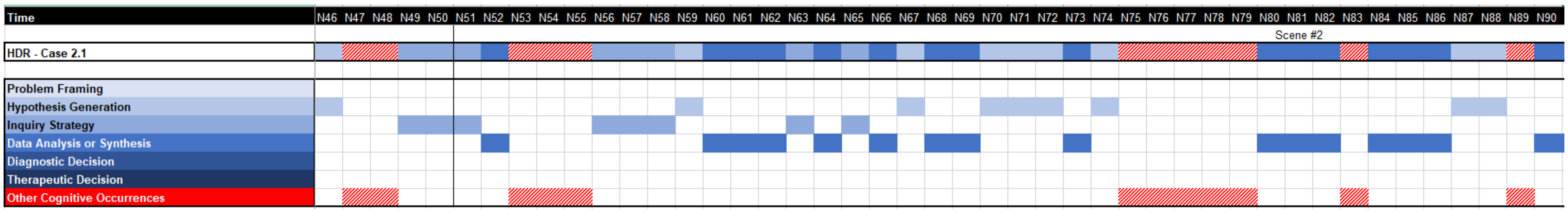
	Case 2.1 (Female; 75th percentile)	
Hypothetico-deductive Reasoning	Frequency	Percentage
Clinical Reasoning	**143**	**61.37%**
Problem Framing	10	4.29%
Hypothesis Generation	27	11.59%
Inquiry Strategy	49	21.03%
Data Analysis or Synthesis	50	21.46%
Diagnostic Decision and Explanation	0	0.00%
Therapeutic Decision and Treatment Options	7	3.00%
Other Cognitive Occurrences	**90**	**38.63%**
Total	**233**	**100.00%**

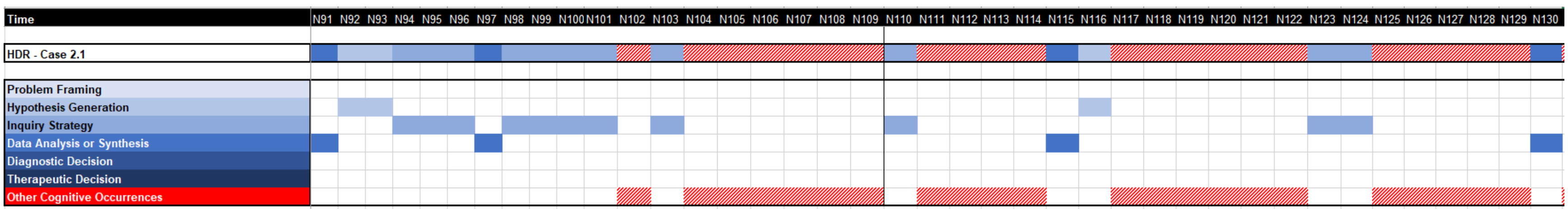
In case 2.1, 61.37% of 233 cognitive occurrences were coded as clinical reasoning from the multimedia case-based assessment (condition 2). Inquiry strategy (49, 21.03%) and data analysis or synthesis (50, 21.46%) were the most frequent reasoning processes in case 2.1.

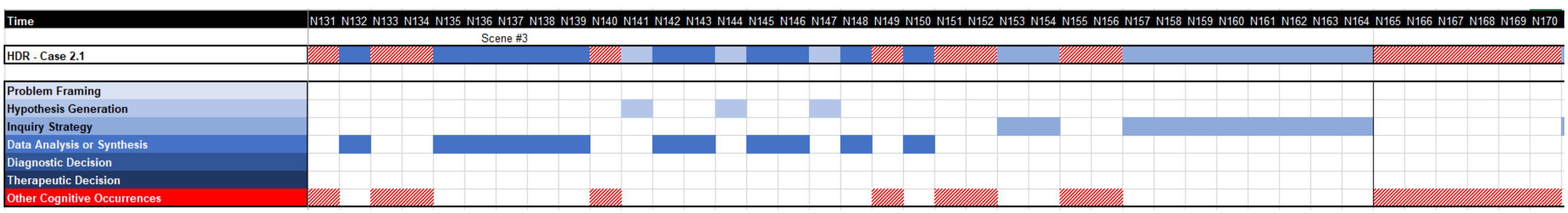
Diagnostic decision and explanation process was not detected in case 2.1. A total 38.63% of 233 cognitive occurrences from case 2.1 were detected and coded as other cognitive occurrences.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes regarding hypothetico-deductive reasoning, clinical reasoning and other cognitive occurrences of case 2.1 are presented as a graphical representation in Figure 4.3.


Time
N1 N2 N3 N4 N5 N6 N7 N8 N9 N10 N11 N12 N13 N14 N15 N16 N17 N18 N19 N20 N21 N22 N23 N24 N25 N26 N27 N28 N29 N30 N31 N32 N33 N34 N35 N36 N37 N38 N39 N40 N41 N42 N43 N44 N45
Scene #1
HDR - Case 2.1
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences


Time
N46 N47 N48 N49 N50 N51 N52 N53 N54 N55 N56 N57 N58 N59 N60 N61 N62 N63 N64 N65 N66 N67 N68 N69 N70 N71 N72 N73 N74 N75 N76 N77 N78 N79 N80 N81 N82 N83 N84 N85 N86 N87 N88 N89 N90
Scene #2
HDR - Case 2.1
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences


Time
N91 N92 N93 N94 N95 N96 N97 N98 N99 N100 N101 N102 N103 N104 N105 N106 N107 N108 N109 N110 N111 N112 N113 N114 N115 N116 N117 N118 N119 N120 N121 N122 N123 N124 N125 N126 N127 N128 N129 N130
HDR - Case 2.1
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences


Time
N131 N132 N133 N134 N135 N136 N137 N138 N139 N140 N141 N142 N143 N144 N145 N146 N147 N148 N149 N150 N151 N152 N153 N154 N155 N156 N157 N158 N159 N160 N161 N162 N163 N164 N165 N166 N167 N168 N169 N170
Scene #3
HDR - Case 2.1
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences

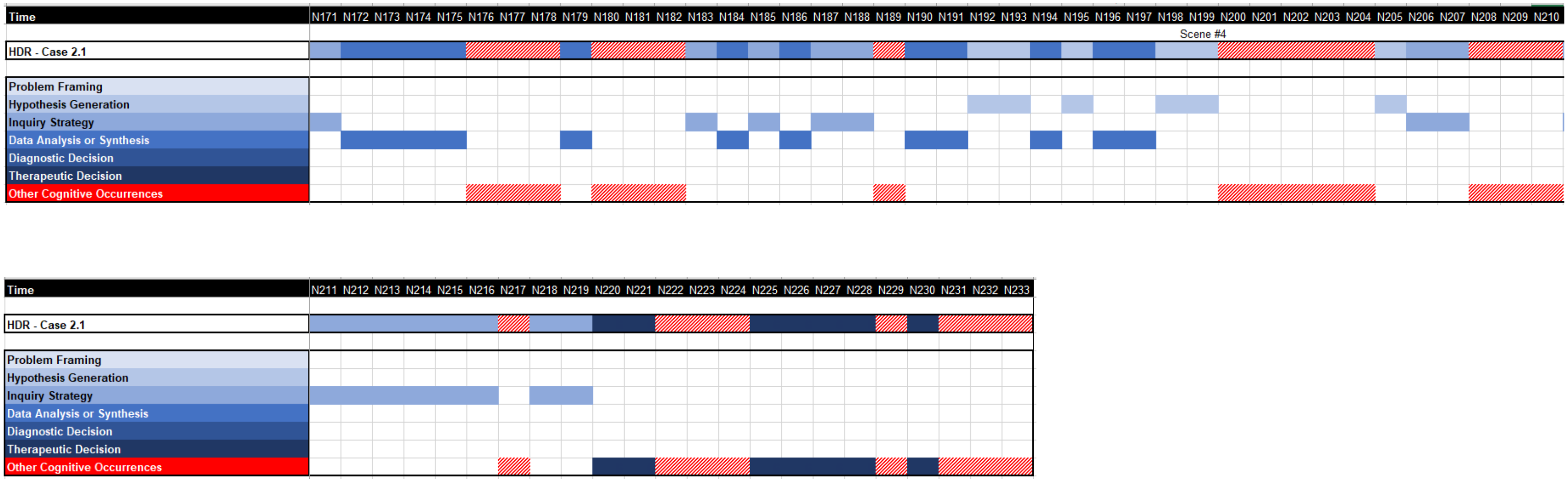


Figure 4.3. A graphical representation of cognitive occurrences from case 2.1

Representative quotes from participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 2.1 was identified. Representative quotes of each reasoning process are presented in Table 4.12.

Table 4.12.
Representative quotes for clinical reasoning from narratives in case 2.1

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	1	"So far… a male patient, 58-year-old… and he was grabbing his left chest"
	Chief complaint	3	"The chief complaint was chest pain, so I was thinking naturally because I am familiar with the case"
	Perceiving a variety of cues: Patient's remarks and responses	8	"Currently, I checked the patient's chief complaint, site of the pain, and character of the pain – left chest and pain that was pressing down"
Hypothesis Generation	Hypothesis Generation	6	"I thought that I need to check digestive, respiratory, circulatory, musculoskeletal system, and psychiatric diseases"
		192	"I was thinking, oh I could eliminate cardiac failure from my hypotheses, so the second hypothesis was changed"
Inquiry Strategy	History-taking	56	"[in the video] the doctor asked about exacerbating/relieving factors and associated symptoms, so I focused more on asking detailed questions such as duration of the pain, and radiating pains"
	Physical examination	154	"I wrote that I needed to do palpation to find out any changes, so I could include the possibility of cardiac failure or eliminate it"

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Data Analysis or Synthesis	Data analysis or synthesis	135	"Among these past histories, the patient has had high blood pressure for five years, but he doesn't take any pills; his father died at an early age because of a stroke, and he is drinking and smoking a lot, so I think these can be a positive signal"
	Ongoing summary	179	"I thought myocardial infarction could be the first possibility from the beginning"
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	-	-
Therapeutic Decision & Treatment Options	Therapeutic Decision & Treatment Options	227	"I thought I need to tell the patient that the current status is very serious, so I wanted to mention risk factors"

Themes of other cognitive occurrences, definitions, and representative quotes from the participant's narratives. As in the previous cases, uncoded data were classified as other cognitive occurrences. The uncoded data from the original data set was initially coded according to the contents of cognitive behaviors and then the data were re-categorized according to emerging themes. The emerging themes from condition 2, the CBA, were the following: test-taking strategy (including test-taking thinking, off-protocol behavior, and educated guessing), point-seeking/hunting, and unnecessary constraints.

The identified frequencies and percentages of each other cognitive occurrences theme from case 2.1 are presented in Table 4.13.

Table 4.13.
Frequencies and percentages of other cognitive occurrences from case 2.1

	Case 2.1 (Female; 75th percentile)	
Other Cognitive Occurrences	Frequency	Percentage
Test-taking Strategy	**54**	**60.00%**
Test-taking Thinking	29	32.22%
Off-protocol Behavior	20	22.22%
Educated Guessing	5	5.56%
Point-seeking/hunting	**25**	**27.78%**
Unnecessary Constraints	**11**	**12.22%**
Total	**90**	**100.00%**

A total 90 out of 233 cognitive occurrences (38.63%) were detected as other cognitive occurrences from case 2.1. Test-taking thinking (29, 32.22%) was the most frequent other cognitive occurrences when the participant encountered clinical problems-to-solve in the CBA. Point-seeking/hunting behavior (25, 27.78%) followed as the second most frequent other cognitive occurrences that interrupted natural problem-solving.

Detailed definitions and representative quotes of each other cognitive occurrences theme from case 2.1 are explained in Table 4.14.

Table 4.14.
Definition and representative quotes for other cognitive occurrences from narratives in case 2.1

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Test-taking Thinking	An examinee displays reasoning or thinking related to testing situations or the test itself rather than solving problems	121	"The question was very similar to MEQ questions I had taken a long time ago…"
Off-protocol Behavior	An examinee is limited in decision-making thinking due to a lack of knowledge	75	"I didn't remember exactly how to classify angina as stable, unstable, and atypical"

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Off-protocol Behavior	An examinee thinks in ways that influences behavior that skips or cuts short clinical reasoning	181-182	"Based on the previous information, I didn't think I could do any more for a differential diagnosis, so I didn't even try to check the blood pressure"
	An examinee thinks of the exam's time limits to make decisions, influencing behavior that interrupts clinical reasoning	126	"It took too much time to answer this question because I was asked to write the medical history in detail."
Educated Guessing	An examinee thinks of the limitations of the test instead of making decisions for problem-solving	47-48	"I was thinking of some urgent diseases, but they are very rare cases, so I thought it is not that important."
Point-seeking/Hunting	An examinee anticipates a grader's checklist and/or fishes for test points instead of making decisions for problem-solving	76-77	"I didn't think unstable angina is the answer [but I wasn't sure], so I put the other two, stable and atypical angina, in the blank."
Unnecessary Constraints	An examinee is misguided or confused by the presentation of the test items or case videos	9	"Yes… I could not see accurately where the patient pointed at his chest"
		12-13	"I didn't know if I could watch the video again. If I knew that I could, I would watch the video again to confirm what I [thought I] saw."

Case 2.2 – participant 2 in condition 2 (male; 90th percentile in GPA). The analyzed results of case 2.2 data from the same male 4th year medical student, 90th percentile GPA of all 4th year students at IUCM, is provided below.

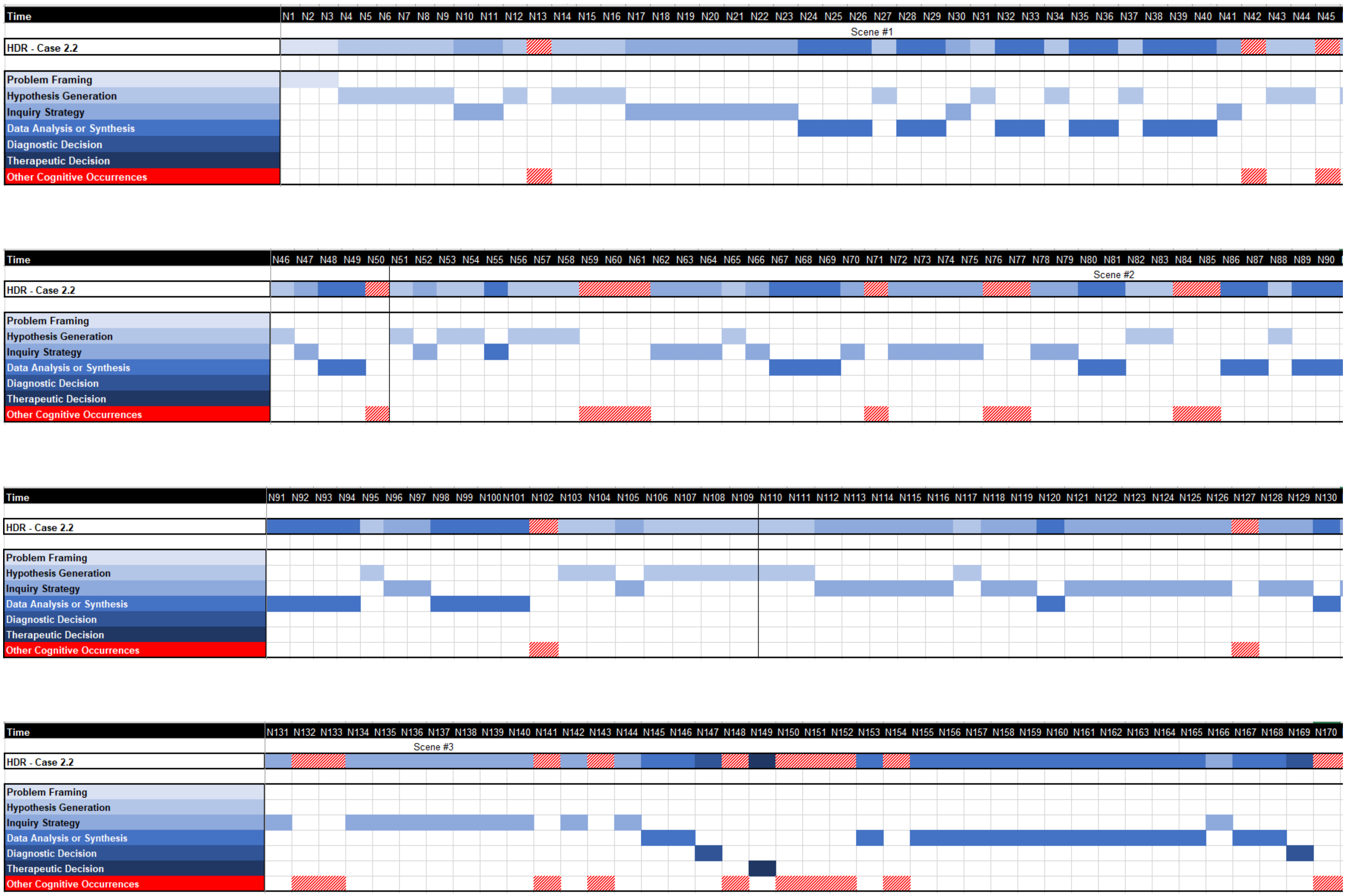
Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 2.2 were coded as clinical reasoning and other cognitive occurrences. The identified frequencies and percentages of each process of clinical reasoning, based on the hypothetico-deductive reasoning model, are presented in Table 4.15.

Table 4.15.
Frequencies and percentages of cognitive occurrences as HDR processes from case 2.2

	Case 2.2 (Male; 90th percentile)	
Hypothetico-deductive Reasoning	Frequency	Percentage
Clinical Reasoning	**174**	**86.57%**
Problem Framing	3	1.49%
Hypothesis Generation	37	18.41%
Inquiry Strategy	65	32.34%
Data Analysis or Synthesis	51	25.37%
Diagnostic Decision and Explanation	4	1.99%
Therapeutic Decision and Treatment Options	14	6.97%
Other Cognitive Occurrences	**27**	**13.43%**
Total	**201**	**100.00%**

As provided in Table 4.15., 86.57% of 201 cognitive occurrences from the case 2.2 were coded as clinical reasoning in the multimedia case-based assessment (condition 2). Similar to the previous case 2.1, inquiry strategy (65, 32.34%) and data analysis or synthesis (51, 25.37%) were the most frequent reasoning processes. A total of 13.43% from case 2.2 were detected and coded as other cognitive occurrence.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes of hypothetico-deductive reasoning, clinical reasoning, and other cognitive occurrences, a graphical representation of case 2.2 is presented in Figure 4.4.


Time
N1 N2 N3 N4 N5 N6 N7 N8 N9 N10 N11 N12 N13 N14 N15 N16 N17 N18 N19 N20 N21 N22 N23 N24 N25 N26 N27 N28 N29 N30 N31 N32 N33 N34 N35 N36 N37 N38 N39 N40 N41 N42 N43 N44 N45
Scene #1
HDR - Case 2.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N46 N47 N48 N49 N50 N51 N52 N53 N54 N55 N56 N57 N58 N59 N60 N61 N62 N63 N64 N65 N66 N67 N68 N69 N70 N71 N72 N73 N74 N75 N76 N77 N78 N79 N80 N81 N82 N83 N84 N85 N86 N87 N88 N89 N90
Scene #2
HDR - Case 2.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N91 N92 N93 N94 N95 N96 N97 N98 N99 N100 N101 N102 N103 N104 N105 N106 N107 N108 N109 N110 N111 N112 N113 N114 N115 N116 N117 N118 N119 N120 N121 N122 N123 N124 N125 N126 N127 N128 N129 N130
HDR - Case 2.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N131 N132 N133 N134 N135 N136 N137 N138 N139 N140 N141 N142 N143 N144 N145 N146 N147 N148 N149 N150 N151 N152 N153 N154 N155 N156 N157 N158 N159 N160 N161 N162 N163 N164 N165 N166 N167 N168 N169 N170
Scene #3
HDR - Case 2.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences

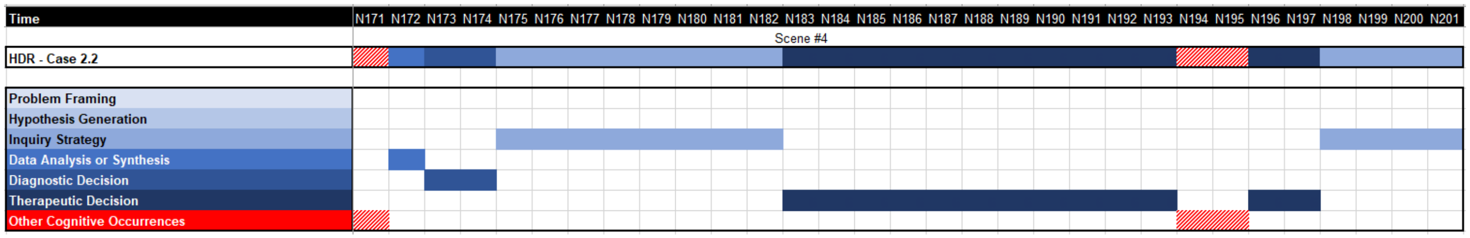


Figure 4.4. A graphical representation of cognitive occurrences from case 2.2

Representative quotes from participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 2.2 was identified. Representative quotes of each reasoning process are presented in Table 4.16.

Table 4.16.
Representative quotes for clinical reasoning from narratives in case 2.2

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	2	"There were many cues in this 20-30 seconds video… the patient's age, location of the pain through watching the patient touched on his chest, character of the pain from the patient's answer that the pain was pretty severe"
	Chief complaint	1	"First, it was a very short video, like 30 seconds, and it was about a clinical presentation: chest pain; the clinical presentation is very important for a patient…"
	Perceiving a variety of cues: Patient's remarks and responses	3	"And also, the patient said that the pain just happened suddenly"
Hypothesis Generation	Hypothesis Generation	5	"The location the patient pointed to was close to the heart, so I was thinking it could be a heart disease, and then I was trying to formulate hypotheses"
		9	"I wasn't sure about the character of the pain at this point, so it could be angina, aortic dissection or GID… or pneumonia is also another possibility, so I was open to more hypotheses"

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Inquiry Strategy	History-taking	17	"There were many things I needed to take for the patient's medical history that was in addition to what I had already taken, such as character of the pain, radiation pain, exacerbating/relieving factors, onset of the pain and so on… and I typed that I needed to do history taking to find associated symptoms. I focused on these things… if I do history taking intensively about the clinical presentation…"
	Physical examination	105	"Typically for the physical examination of chest pain, I hear the heart sounds and heart murmurs, then when I thought about it, I just remembered…"
Data Analysis or Synthesis	Data analysis or synthesis	89-90	"I received the information from the patient that he has underlying diseases, high blood pressure, diabetes, family history where his father died at an early age because of a stroke, smoking, and so on…. These histories are risk factors for acute coronary syndrome"
	Ongoing summary	93-94	"I typed my hypothesis here, and then I thought I could explain that these medical histories support my hypothesis. I thought I have more information for my hypothesis now, and I believed my hypothesis is almost right"
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	174	"I didn't put anything in diagnosis 3 [because the possibility is very low for hypothesis 3], and I was trying to focus on diagnosis 1 and 2."
Therapeutic Decision & Treatment Options	Therapeutic Decision & Treatment Options	193	"and because it [lifestyle] is directly related to the patient's illness I thought giving guidance on the patient's dietary life and exercise plans to improve lifestyle would be enough for now"

Themes of other cognitive occurrences, definitions, and represented quotes from the participant's narratives. As in the previous cases, any uncoded data were classified under other cognitive occurrences. The uncoded data from the original data set was initially coded according to the contents of the other cognitive occurrences, then the data were re-categorized according to emerging themes. The emerging themes from condition 2, the CBA, were the following: test-taking strategy (including test-taking thinking, off-protocol behavior, and educated guessing), point-seeking/hunting, and unnecessary constraints.

The identified frequencies and percentages of each other cognitive occurrences theme from case 2.2 are presented in Table 4.17.

Table 4.17.
Frequencies and percentages of other cognitive occurrences from case 2.2

	Case 2.2 (Male; 90th percentile)	
Other Cognitive Occurrences	Frequency	Percentage
Test-taking Strategy	**18**	**66.66%**
Test-taking Thinking	8	29.63%
Off-protocol Behavior	4	14.81%
Educated Guessing	6	22.22%
Point-seeking/hunting	**2**	**7.41%**
Unnecessary Constraints	**7**	**25.93%**
Total	**27**	**100.00%**

A total 27 out of 201 cognitive occurrences (13.43%) were detected as other cognitive occurrences from case 2.2. Among the five identified as other cognitive occurrences in condition 2 (CBA), test-taking thinking (8, 29.63%), unnecessary constraints (7, 25.93%), and educated guessing (6, 22.22%) were similarly detected as other cognitive occurrences when the participant encountered clinical problems-to-solve in the CBA.

Detailed definitions and representative quotes of each other cognitive occurrences theme from case 2.2 are explained in Table 4.18.

Table 4.18.

Definition and representative quotes for other cognitive occurrences from narratives in case 2.2

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Test-taking Thinking	An examinee displays reasoning or thinking related to testing situations or the test itself rather than solving problems	141, 143	"Taking the patient's blood pressure… for example, if I take the OSCE, I can't take the patient's blood pressure because it takes so much time in the exam…" "so I didn't put the answer because I wouldn't do it in other exams"
Off-protocol Behavior	An examinee is limited in decision-making thinking due to a lack of knowledge	61	"so, I skipped hypothesis 3 in order to think about it more later on, so I did the next question instead."
	An examinee thinks in ways that influences behavior that skips or cuts short clinical reasoning	195	"I wasn't sure if I should put chest PA or not… so I answered it last."
	An examinee thinks of the exam's time limits to make decisions, influencing behavior that interrupts clinical reasoning	12	"[I would like to think more about the hypotheses] but I didn't have enough time…"
Educated Guessing	An examinee thinks of the limitations of the test instead of making decisions for problem-solving	76	"I thought this clinical presentation was similar to the clinical presentation I already did in the OSCE, but a little bit different as well. Because I took the OSCE [one week ago], I was comparing the OSCE and this [CBA]."
Point-seeking/Hunting	An examinee anticipates a grader's checklist and/or fishes for test points instead of making decisions for problem-solving	132	"because the professor in this video is an expert, I was thinking that oh… an expert approaches [the situation] like this… [then I changed the way I thought]"

	Definition	Narrative Number	Sample Quote
Unnecessary Constraints	An examinee is misguided or confused by the presentation of the test items or case videos	102	"yes, I was kind of confused… in the previous question, it asked for a third hypothesis… and here, I was again asked to write down hypotheses, so I was thinking I needed to put a revised hypothesis with an explanation of why it was wrong from my previous answer because [the test screen] showed my previous answer of GID. Or, should I put a new hypothesis with an explanation of why this [new] hypothesis has more possibility [than the previous one] …"

Research Question 1.3. What are the cognitive processes of 4th year medical students in solving clinical diagnostic problems during a clinical knowledge and reasoning examination (Modified Essay Question)?

For research question 1.3, the same students from the previous cases participated in the modified essay question, which is the third condition of this study.

Case 3.1 – participant 1 in condition 3 (female; 75th percentile in GPA). The analyzed result of case 3.1 data from the female 4th year medical student, 75th percentile GPA of all 4th year students at IUCM, is provided below.

Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 3.1 were coded as clinical reasoning or other cognitive occurrences. The identified frequencies and percentages of each process of clinical reasoning, based on the hypothetico-deductive reasoning model, are presented in Table 4.19.

Table 4.19.
Frequencies and percentages of cognitive occurrences as HDR processes from case 3.1

	Case 3.1 (Female; 75th percentile)	
	Frequency	Percentage
Clinical Reasoning	**109**	**63.74%**
Problem Framing	3	1.75%
Hypothesis Generation	63	36.84%
Inquiry Strategy	10	5.85%
Data Analysis or Synthesis	31	18.13%
Diagnostic Decision and Explanation	2	1.17%
Therapeutic Decision and Treatment Options	0	0.00%
Other Cognitive Occurrences	**62**	**36.26%**
Total	**171**	**100.00%**

In case 3.1, the modified essay question (condition 3), 63.74% of 171 cognitive occurrences were coded as clinical reasoning from the modified essay question (condition 3). Hypothesis generation (63, 36.84%) was the most frequent reasoning process, with data analysis or synthesis (31, 18.13%) as the second most frequent process. Therapeutic decision and treatment options was not detected in case 3.1. A total 36.26% of 171 cognitive occurrences from case 3.1 were identified and coded as other cognitive occurrences.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes of hypothetico-deductive reasoning, clinical reasoning, and other cognitive occurrences, a graphical representation of case 3.1 is presented in Figure 4.5.

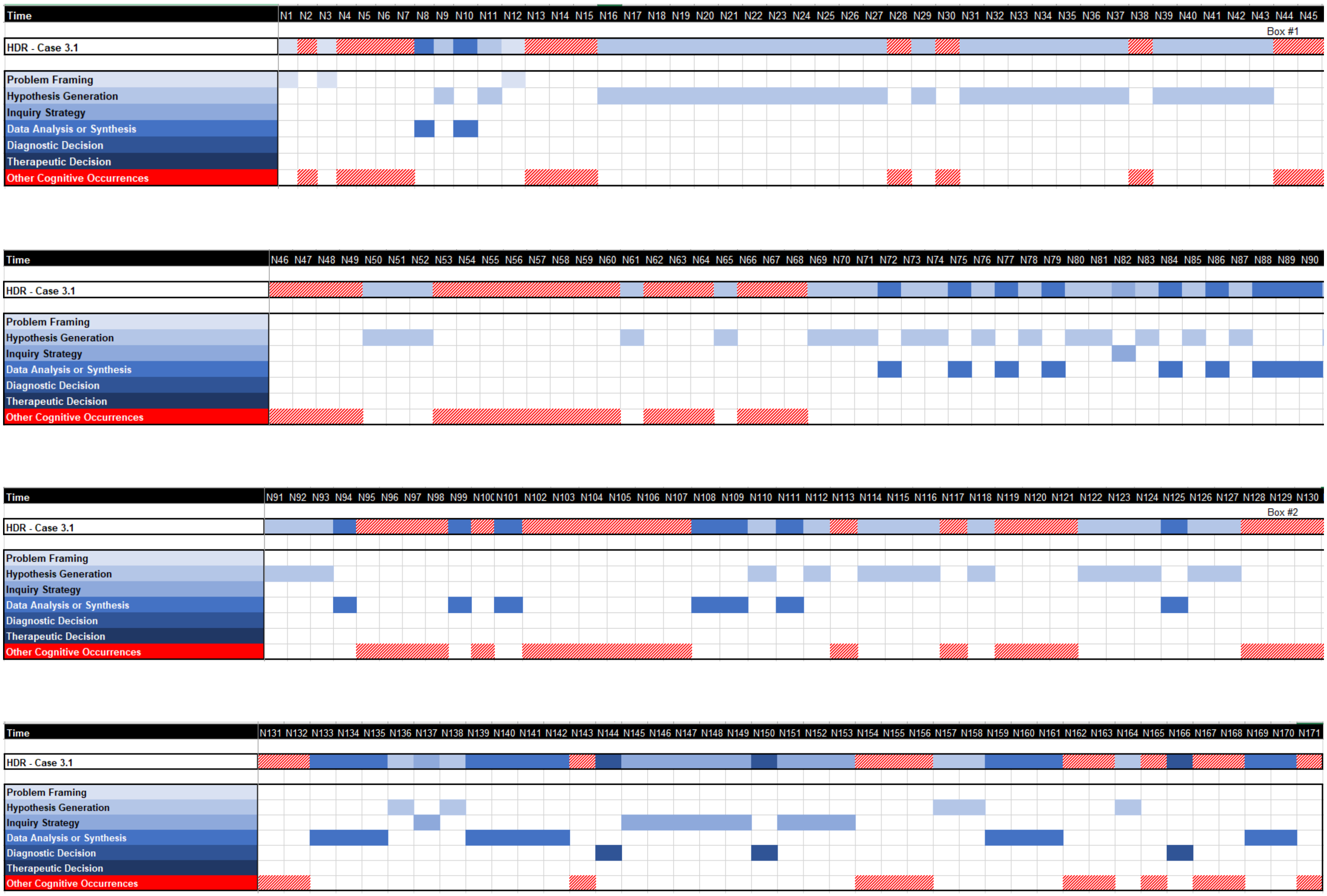


Figure 4.5. A graphical representation of cognitive occurrences from case 3.1

Representative quotes from participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 3.1 was identified. Representative quotes of each reasoning process are presented in Table 4.20.

Table 4.20.

Representative quotes for clinical reasoning of narratives in case 3.1

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	1	"I received information here about the patient's gender, age, clinical presentation, a short description of the clinical presentation, exacerbating/relieving factors, and clinical history"
	Chief complaint	12	"After I went on to the questions, I didn't think about it [hypothesis generation], then just focused on the clinical presentation"
Hypothesis Generation	Hypothesis Generation	11	"I thought I might also need to think about the digestive system"
		127	"Yes, I was thinking about the circulatory system before this [and eliminated digestive system]"
Inquiry Strategy	History-taking	82	"I thought I needed to do differential diagnosis here [for pulmonary illness]"
	Physical examination	137	"I tried to write down types of examinations that could identify the best hypothesis among the major hypotheses given at the beginning"
Data Analysis or Synthesis	Data analysis or synthesis	86	"The character of the pain wasn't provided before but suddenly the patient mentioned the pain was like lancinating [i.e., piercing] … at this [particular] point"
	Ongoing summary	144	"I kept aortic dissection as a diagnosis in mind, then thought about what I should do for the aortic dissection…"
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	166	"I was almost sure that aortic dissection is the disease"

Themes of other cognitive occurrences, definitions, and representative quotes from the participant's narratives. As in the previous cases, uncoded data were classified as other cognitive occurrences. The uncoded data from the original data set was initially coded according to the contents of cognitive occurrences, then the data were re-categorized according to emerging themes. The emerging themes from condition 3, the MEQ, were the following: test-taking strategy (including test-taking thinking and off-protocol behavior), point-seeking/hunting, and unnecessary constraints.

The identified frequencies and percentages of each cognitive occurrences theme from case 3.1 are presented in Table 4.21.

Table 4.21.
Frequencies and percentages of other cognitive occurrences from case 3.1

	Case 3.1 (Female; 75th percentile)	
Other Cognitive Occurrences	Frequency	Percentage
Test-taking Strategy	**27**	**43.55%**
Test-taking Thinking	13	20.97%
Off-protocol Behavior	14	22.58%
Educated Guessing	11	17.74%
Point-seeking/hunting	**16**	**25.81%**
Unnecessary Constraints	**8**	**12.90%**
Total	**62**	**100.00%**

A total 62 out of 171 cognitive occurrences (36.26%) were identified as other cognitive occurrences from case 3.1, and in this case, point-seeking/hunting (16, 25.81%) was the most frequent other cognitive occurrences when the participant encountered clinical problems in the MEQ. The rest of the other cognitive occurrences were as follows: off-protocol behavior (14, 22.58%), test-taking thinking (13, 20.97%), and educated guessing (11, 17.74%).

Detailed definitions and representative quotes of each other cognitive occurrences theme from case 3.1 are explained in Table 4.22.

Table 4.22.

Definition and representative quotes for other cognitive occurrences from narratives in case 3.1

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Test-taking Thinking	An examinee displays reasoning or thinking related to testing situations or the test itself rather than solving problems	55	"I was rushed for time, so I kept thinking about the other parts while I was writing this answer. I wrote something in the next question because I just remembered, then came back to the previous question…"
Off-protocol Behavior	An examinee is limited in decision-making thinking due to a lack of knowledge	163	"I wasn't familiar with the question about the difference between left- and right-hand pulses, and I thought I didn't know about it well enough"
	An examinee thinks in ways that influences behavior that skips or cuts short clinical reasoning	56	"[if I had more time] I would be able to arrange a pneumothorax or aortic dissection according to the appropriate branch of the concept map…"
Educated Guessing	An examinee thinks of the exam's time limits to make decisions, influencing behavior that interrupts clinical reasoning	102-103	"I assumed that the fact that the left-hand pulse was lower than the right-hand was the key point, so it was emphasized… it turned out to me it was a meaningful [hint] from the question"
Point-seeking/Hunting	An examinee anticipates a grader's checklist and/or fishes for test points instead of making decisions for problem-solving	62	"for example, herpes zoster or costal bone fracture… I believed there was no high possibility that these were the problem, but I had to write them"
Unnecessary Constraints	An examinee is misguided or confused by the presentation of the test items	95-96	"I couldn't focus on the questions…", "because I had to distinguish old information [that I had already read in the first textbox] with the new information in the second textbox"

Case 3.2 - participant 2 in condition 3 (male; 90th percentile in GPA). The analyzed result of case 3.2 data from the male 4th year medical student, 90th percentile GPA of all 4th year students at IUCM, is provided below.

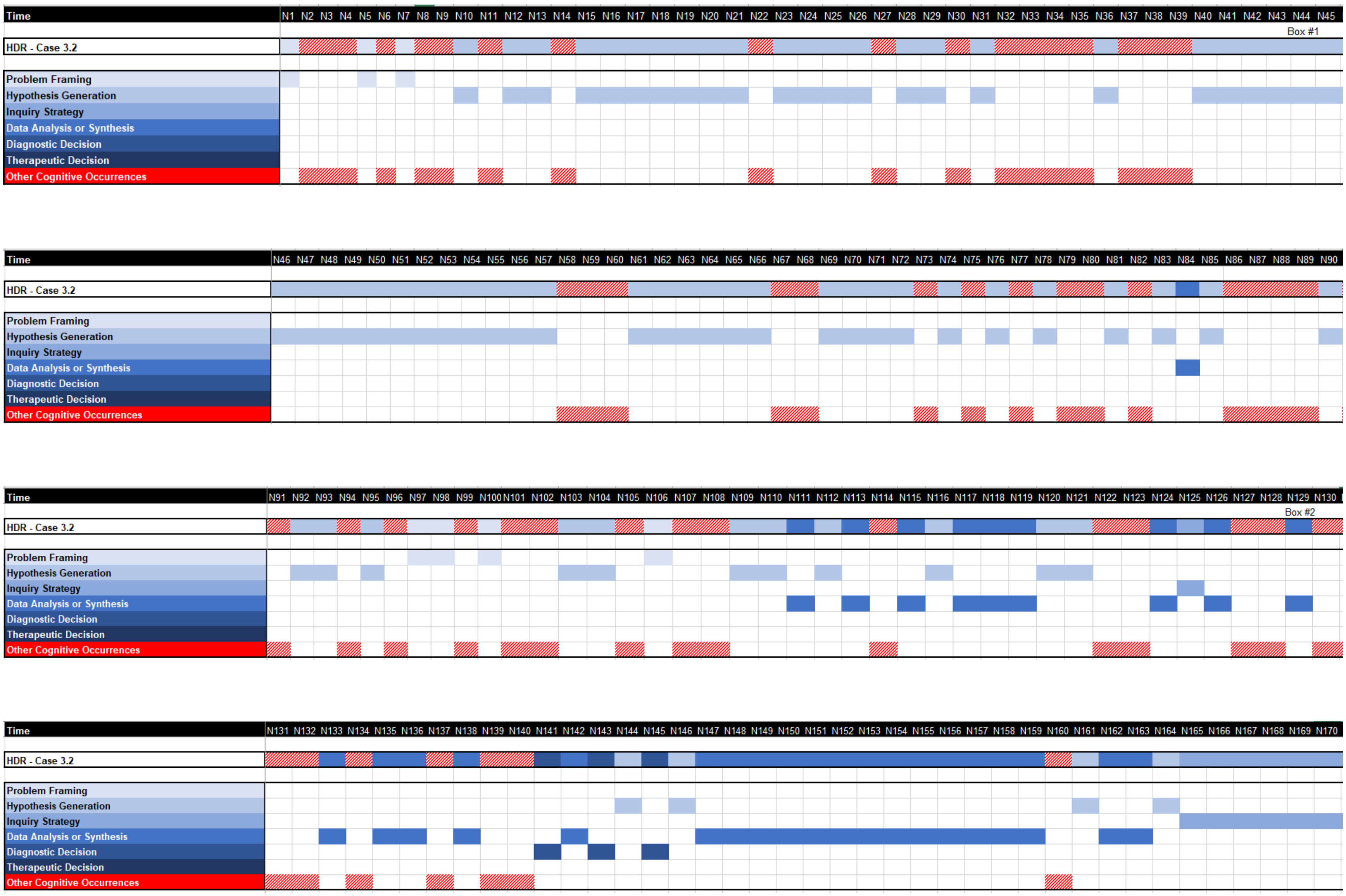
Identified frequencies and percentages of clinical reasoning (HDR). All meaningful cognitive occurrences from case 3.2 were coded as clinical reasoning and other cognitive occurrences. The identified frequencies and percentages of each process of clinical reasoning, based on the hypothetico-deductive reasoning model, are presented in Table 4.23.

Table 4.23.
Frequencies and percentages of cognitive occurrences as HDR processes from case 3.2

	Case 3.2 (Male; 90th percentile)	
Hypothetico-deductive Reasoning	Frequency	Percentage
Clinical Reasoning	**120**	**67.04%**
Problem Framing	7	3.91%
Hypothesis Generation	68	37.99%
Inquiry Strategy	12	6.70%
Data Analysis or Synthesis	30	16.76%
Diagnostic Decision and Explanation	3	1.68%
Therapeutic Decision and Treatment Options	0	0.00%
Other Cognitive Occurrences	**59**	**32.96%**
Total	**179**	**100.00%**

In case 3.2, 67.04% of 179 cognitive occurrences were coded as clinical reasoning from the modified essay question (condition 3). Hypothesis generation (68, 37.99%) was the most frequent reasoning process, and data analysis or synthesis (30, 16.76%) followed as the second most frequent process. As in the previous case 3.1., therapeutic decision and treatment options was not identified. A total 32.96% of 179 cognitive occurrences from case 3.2 were identified and coded as other cognitive occurrences.

Representation of cognitive occurrences. Based on the result of clinical reasoning processes of hypothetico-deductive reasoning, clinical reasoning, and other cognitive occurrences, a graphical representation of case 3.2 is presented in Figure 4.6.


Time
N1 N2 N3 N4 N5 N6 N7 N8 N9 N10 N11 N12 N13 N14 N15 N16 N17 N18 N19 N20 N21 N22 N23 N24 N25 N26 N27 N28 N29 N30 N31 N32 N33 N34 N35 N36 N37 N38 N39 N40 N41 N42 N43 N44 N45
Box #1
HDR - Case 3.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N46 N47 N48 N49 N50 N51 N52 N53 N54 N55 N56 N57 N58 N59 N60 N61 N62 N63 N64 N65 N66 N67 N68 N69 N70 N71 N72 N73 N74 N75 N76 N77 N78 N79 N80 N81 N82 N83 N84 N85 N86 N87 N88 N89 N90
HDR - Case 3.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N91 N92 N93 N94 N95 N96 N97 N98 N99 N100 N101 N102 N103 N104 N105 N106 N107 N108 N109 N110 N111 N112 N113 N114 N115 N116 N117 N118 N119 N120 N121 N122 N123 N124 N125 N126 N127 N128 N129 N130
Box #2
HDR - Case 3.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences
Time
N131 N132 N133 N134 N135 N136 N137 N138 N139 N140 N141 N142 N143 N144 N145 N146 N147 N148 N149 N150 N151 N152 N153 N154 N155 N156 N157 N158 N159 N160 N161 N162 N163 N164 N165 N166 N167 N168 N169 N170
HDR - Case 3.2
Problem Framing
Hypothesis Generation
Inquiry Strategy
Data Analysis or Synthesis
Diagnostic Decision
Therapeutic Decision
Other Cognitive Occurrences

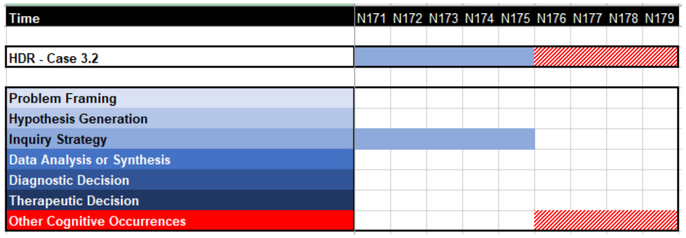


Figure 4.6. A graphical representation of cognitive occurrences from case 3.2

Representative quotes from participant's narratives. Each hypothetico-deductive reasoning process as clinical reasoning from case 3.2 was identified. Representative quotes of each reasoning process are presented in Table 4.24.

Table 4.24.
Representative quotes for clinical reasoning from narratives in case 3.2

HDR Process	Sub-indicator	Narrative Number	Sample Quote
Problem Framing	Perceiving a variety of cues: Observation	100	"the patient grabbed his chest, and he was breathing hard… at lunch time… I read [the entire] textbox"
	Chief complaint	1	"There is a clinical presentation: chest pain at first"
Hypothesis Generation	Hypothesis Generation	28	"I wrote down valve dysfunction, vessel obstruction, cardiomyopathy, and arrhythmia here [under cardiovascular system]"
		49	"the lower-level branches kept coming out from my head; then, I expanded the upper-level branches, deleted or revised them"
Inquiry Strategy	History-taking	-	-
	Physical examination	165	"yes, first of all, all needed examinations, such as an examination to differentiate myocardial infarction and aortic dissection"
Data Analysis or Synthesis	Data analysis or synthesis	118	"this information was not provided before, and the pain occurred newly, and the pain was lancinating [i.e., piercing] …"
	Ongoing summary	156	"honestly, they [the cues] were close to aortic dissection too"
Diagnostic Decision & Explanation	Diagnostic Decision & Explanation	145	"I tried to write down the reason why the diagnosis is aortic dissection, and I was almost sure about it"
Therapeutic Decision & Treatment Options	Therapeutic Decision & Treatment Options	-	-

Themes of other cognitive occurrences, definitions, and representative quotes from the participant's narratives. As in the previous cases, uncoded data were classified under other cognitive occurrences. The uncoded data from the original one was initially coded according to the contents of cognitive occurrences, then the data were re-categorized according to emerging themes. The emerging themes from condition 3, the MEQ, were the following: test-taking strategy (including test-taking thinking, off-protocol behavior, and educated guessing), point-seeking/hunting, and unnecessary constraints.

The identified frequencies and percentages of each other cognitive occurrences theme from case 3.2 are presented in Table 4.25.

Table 4.25.
Frequencies and percentages of other cognitive occurrences from case 3.2

	Case 3.2 (Male; 90th percentile)	
	Frequency	Percentage
Test-taking Strategy	**34**	**57.61%**
Test-taking Thinking	7	11.86%
Off-protocol Behavior	22	37.29%
Educated Guessing	5	8.48%
Point-seeking/hunting	**9**	**15.25%**
Unnecessary Constraints	**16**	**27.12%**
Total	**59**	**100.00%**

A total 59 out of 179 cognitive occurrences (32.96%) were identified as other cognitive occurrences from case 3.2, and off-protocol behavior (22, 37.29%) was the most frequent other cognitive occurrences when the participant encountered clinical problems in the MEQ. Unnecessary constraints (16, 27.12%) was the next most frequent other cognitive occurrences that interrupted natural problem-solving.

Detailed definitions and representative quotes of each other cognitive occurrences theme from case 3.2 are explained in Table 4.26.

Table 4.26.

Definition and representative quotes for other cognitive occurrences from narratives in case 3.2

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Test-taking Thinking	An examinee displays reasoning or thinking related to testing situations or the test itself rather than solving problems	160	"I was guided to answer only two hypotheses in question 2, so I thought these two hypotheses are enough"
Off-protocol Behavior	An examinee is limited in decision-making thinking due to a lack of knowledge	134	"I didn't think about why this [the difference between left- and right-hand pulses] happened before… I just knew about it and moved on… [so I wasn't sure about the question]"
	An examinee thinks in ways that influences behavior that skips or cuts short clinical reasoning	91	"It took so much time to do it [concept map branches] so when I saw the next question, I had only 1-minute"
Educated Guessing	An examinee thinks of the limitations of the test instead of making decisions for problem-solving	101	"Oh, I thought this could be a similar test like before [OSCE and CBA]. So, I was thinking this research is comparing the same case among different tests and is about comparing how I think among the tests… It was so similar."
Point-seeking/Hunting	An examinee anticipates a grader's checklist and/or fishes for test points instead of making decisions for problem-solving	60	"Honestly, I don't remember all the systems… I didn't write down the systems by priority, I just wrote down what I remembered first. I just wrote it like that."

Test-taking Strategy	Definition	Narrative Number	Sample Quote
Unnecessary Constraints	An examinee is misguided or confused by the presentation of the test items	4	"It was just simple chest pain… the patient came because of chest pain, and I wasn't sure how I could define this chest pain…"
		68	"Yes, a paper-based test, MEQ was always… I needed to pay attention to my hand writing because an examiner should be able to read what I wrote. I must pay attention to that. So, if an answer box is too small, I paid too much attention to it"

Research Question 1.4. How do medical students think similarly to solve clinical diagnostic problems in each assessment?

The goal of research question 1 was to explore medical students' cognitive processes of clinical diagnostic problem-solving in three different types of clinical assessments and to identify the similarities of the clinical reasoning processes that occurred between the two participants in each condition. Based on the results from research questions 1.1 through 1.3, each case result was presented with hypothetico-deductive reasoning as the clinical reasoning processes in the previous section. The aggregated result regarding clinical reasoning processes is provided in Table 4.27.

Table 4.27.
Percentages of clinical reasoning in three different types of conditions

Clinical Reasoning				
Hypothetico-deductive Reasoning		Condition 1 OSCE	Condition 2 CBA	Condition 3 MEQ
Clinical Reasoning	**Participant 1**	**79.67%**	**61.37%**	**63.74%**
	Participant 2	**69.44%**	**86.57%**	**67.04%**
Problem Framing	Participant 1	4.07%	4.29%	1.75%
	Participant 2	6.98%	1.49%	3.91%
Hypothesis Generation	Participant 1	4.07%	11.59%	36.84%
	Participant 2	6.64%	18.41%	37.99%
Inquiry Strategy	Participant 1	39.84%	21.03%	5.85%
	Participant 2	28.90%	32.34%	6.70%
Data Analysis or Synthesis	Participant 1	20.33%	21.46%	18.13%
	Participant 2	17.94%	25.37%	16.76%
Diagnostic Decision and Explanation	Participant 1	5.69%	0.00%	1.17%
	Participant 2	5.32%	1.99%	1.68%
Therapeutic Decision and Treatment Options	Participant 1	5.69%	3.00%	0.00%
	Participant 2	3.65%	6.97%	0.00%
Other Cognitive Occurrences	**Participant 1**	**20.33%**	**38.63%**	**36.26%**
	Participant 2	**30.56%**	**13.43%**	**32.96%**

In order to identify the similarities of the clinical reasoning processes that occurred between the two participants in each condition, a comparative analysis was conducted to compare clinical reasoning patterns using line graphs. Graphical representations of each condition are presented in Figure 4.7 for condition 1 (OSCE), in Figure 4.8 for condition 2 (CBA), and in figure 4.9 for condition 3 (MEQ).

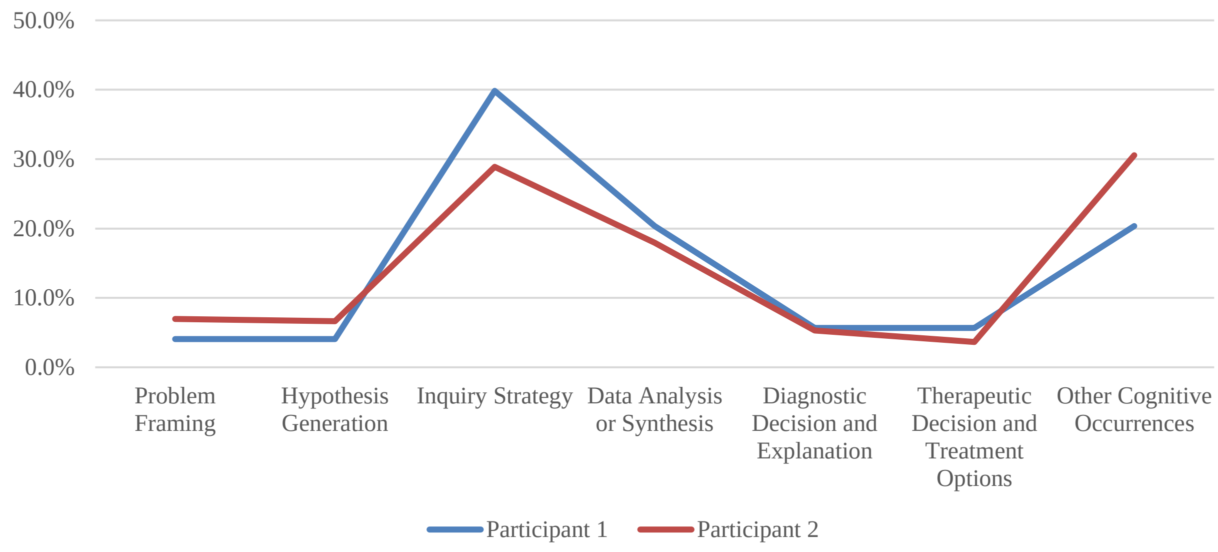


Figure 4.7. A graphical representation of clinical reasoning processes in condition 1 (OSCE)

As shown in Figure 4.7, a similar clinical reasoning process pattern between both participants is evident. In general, more clinical reasoning processes from participant 1 were observed than for participant 2. The inquiry strategy phase was the most frequent clinical reasoning process for both participants, and the data analysis or synthesis phase was the next most frequent. Based on the graphical representation, the result revealed that the two participants followed similar cognitive processes in condition 1 (OSCE).

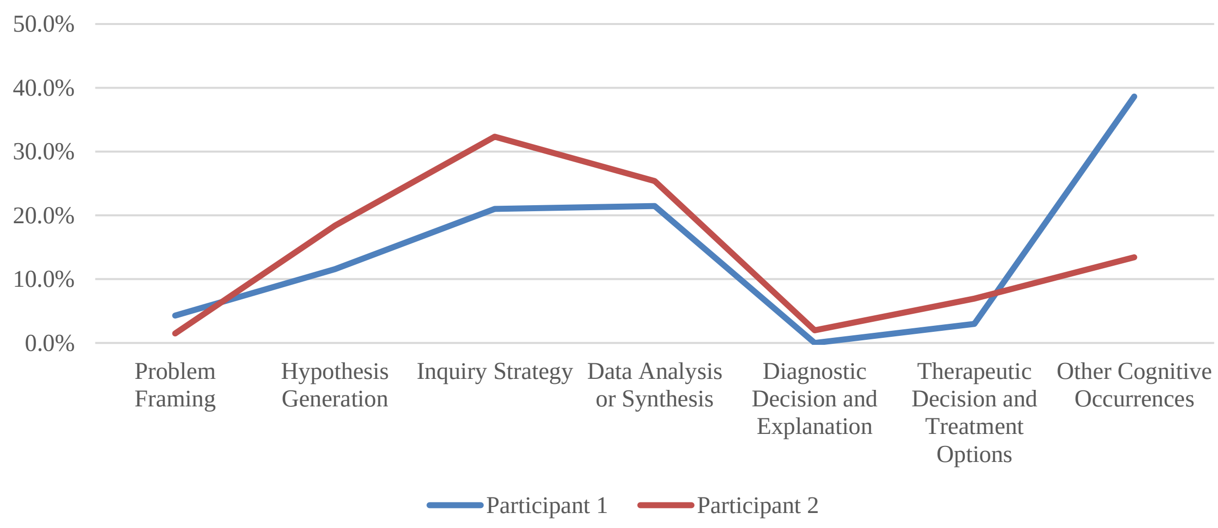


Figure 4.8. A graphical representation of clinical reasoning processes in condition 2 (CBA)

Figure 4.8 shows that each participant's clinical reasoning process pattern was similar. For condition 2 and unlike the previous result from condition 1, more clinical reasoning processes were observed for participant 2 than for participant 1. The inquiry strategy phase was the most frequent clinical reasoning process for both participants, and the data analysis or synthesis phase was the next, similar to results from condition 1 (OSCE). Based on the graphical representation, the result also discovered that two participants followed similar cognitive processes in condition 2 (CBA).

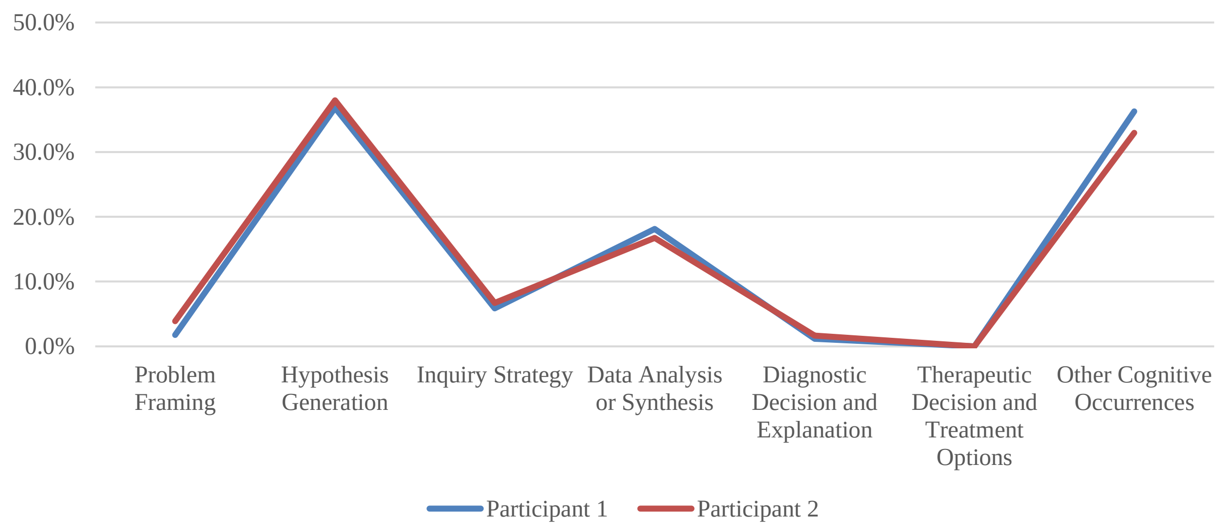


Figure 4.9. A graphical representation of clinical reasoning processes in condition 3 (MEQ)

As presented in Figure 4.9, both participants' clinical reasoning process patterns were almost identical. In condition 3, the hypothesis generation phase was the most frequent clinical reasoning process for both participants, and the data analysis or synthesis phase was the next most frequent. Also shown in Figure 4.9 is that the two participants followed similar cognitive processes in condition 3 (MEQ).

Along with all three graphical representation results, the proportional frequencies of all types of cognitive occurrences for each participant were similar in this study.

Comparison of Clinical Reasoning Processes (HDR) and Other Cognitive Occurrences among Clinical Assessments

A cross-case analysis was conducted to compare clinical reasoning processes with other cognitive occurrences, which respectively afford cognitive aids or distractions to solving problems. This cross-case analysis was also conducted for each condition as a result of the second research question, which led to an understanding of how the different types of clinical assessments promote different clinical reasoning and other cognitive occurrences. Research question 2 is provided below.

Research Question 2. How do medical students think differently to solve clinical diagnostic problems in the three different types of assessments?

Similar to the first section of this chapter, based on the analyzed results from the previous research question, all meaningful cognitive occurrences of participants' narratives were coded using hypothetico-deductive reasoning as the clinical reasoning processes, and uncoded data were classified under other cognitive occurrences. The frequencies of the two participants' cognitive occurrences were averaged in order to compare the differences in clinical reasoning among the three different clinical assessments. Because the result from research question 1.4 revealed similarities between each participant's proportional frequencies in all types of cognitive occurrences for all three conditions, the average frequency data may represent the proportional frequencies of each clinical reasoning process for each condition.

Under the second research question, two main analyses were conducted to compare medical students' clinical reasoning processes and other cognitive occurrences among different types of clinical assessments. This section provides (1) identified percentages of each research participant's clinical reasoning processes based on hypothetico-deductive reasoning in each condition, (2) average data of clinical reasoning from all cases, and (3) average data from all cases of other cognitive occurrences in each condition.

Clinical Reasoning Processes (Hypothetico-deductive Reasoning)

Aggregated result of clinical reasoning processes. A total of six cases were analyzed in the study: two participants in three different conditions each. In order to examine the differences of medical students' clinical reasoning processes in the three conditions, with each condition a different type of clinical assessments, a cross-case analysis was conducted. The results provide an overview of the clinical reasoning processes coded using hypothetico-deductive reasoning (see Table 4.28). Because this analysis was conducted with the participants' reconstructed narrative of occurrences based on naturalistic decision making through stimulated recall interviews, the number of occurrences for each case was different, depending on the participants' recalling of the occurrences. Due to this limitation of the research method, only percentages were used to compare the differences of reasoning processes in the three conditions.

Problem framing. As shown in Table 4.28, participant 1 displayed similar results in the problem framing phase of condition 1 (OSCE, 4.07%) and condition 2 (CBA, 4.29%). For participant 2, the problem framing process occurred most in condition 1 (OSCE, 6.98%) than in other assessments. The average percentages of problem framing in each condition for both participants were 5.53% (OSCE), 2.89% (CBA), and 2.83% (MEQ).

Hypothesis generation. Both participant 1 and 2 showed the most frequent hypothesis generation process occurring in the MEQ among the three conditions (36.84%, 37.99%). The average percentages of hypothesis generation in each condition from both participants were 5.36% (OSCE), 15.00% (CBA), and 37.42% (MEQ).

Table 4.28.
Identified percentages of clinical reasoning in three different types of clinical assessments

Hypothetico-deductive Reasoning		Condition 1 OSCE	Condition 2 CBA	Condition 3 MEQ
Clinical Reasoning	Participant 1	79.67%	61.37%	63.74%
	Participant 2	69.44%	86.57%	67.04%
	Average	**74.56%**	**73.97%**	**65.39%**
Problem Framing	Participant 1	4.07%	4.29%	1.75%
	Participant 2	6.98%	1.49%	3.91%
	Average	**5.53%**	**2.89%**	**2.83%**
Hypothesis Generation	Participant 1	4.07%	11.59%	36.84%
	Participant 2	6.64%	18.41%	37.99%
	Average	**5.36%**	**15.00%**	**37.42%**
Inquiry Strategy	Participant 1	39.84%	21.03%	5.85%
	Participant 2	28.90%	32.34%	6.70%
	Average	**34.37%**	**26.69%**	**6.27%**
Data Analysis or Synthesis	Participant 1	20.33%	21.46%	18.13%
	Participant 2	17.94%	25.37%	16.76%
	Average	**19.14%**	**23.42%**	**17.45%**
Diagnostic Decision and Explanation	Participant 1	5.69%	0.00%	1.17%
	Participant 2	5.32%	1.99%	1.68%
	Average	**5.50%**	**0.99%**	**1.42%**
Therapeutic Decision and Treatment Options	Participant 1	5.69%	3.00%	0.00%
	Participant 2	3.65%	6.97%	0.00%
	Average	**4.66%**	**4.98%**	**0.00%**

Inquiry strategy. Inquiry strategy process is one of the most frequent reasoning processes in condition 1 (OSCE) and condition 2 (CBA). For participant 1, 39.84% of all cognitive occurrences were in inquiry strategy for condition 1, and 21.03% of all cognitive occurrences were in inquiry strategy for condition 2. For participant 2, 32.34% of all cognitive occurrences were in inquiry strategy for condition 2, and 28.90% of all cognitive occurrences were in inquiry strategy for condition 1. The average percentages of inquiry strategy for each condition from both participants were 34.37% (OSCE), 26.69% (CBA), and 6.27% (MEQ).

Data analysis or synthesis. The data analysis or synthesis process was also one of the most frequently occurring reasoning processes among all three conditions. For participant 1, data analysis or synthesis comprised: 21.46% of all cognitive occurrences in condition 2 (CBA), 20.33% of all cognitive occurrences in condition 1 (OSCE), and 18.13% of all cognitive occurrences in condition 3 (MEQ). For participant 2, data analysis or synthesis comprised: 25.37% of all cognitive occurrences in condition 2 (CBA), 17.94% of all cognitive occurrences in condition 1 (OSCE), and 16.76% of all cognitive occurrences in condition 3 (MEQ). The average percentages of data analysis or synthesis in each condition from both participants were 19.14% (OSCE), 23.42% (CBA), and 17.45% (MEQ).

Diagnostic decision and explanation. For participant 1, diagnostic decision and explanation process comprised: 5.69% of all cognitive occurrences in condition 1 (OSCE), 3.00% of all cognitive occurrences in condition 2 (CBA), and none was detected in condition 3. For participant 2, diagnostic decision and explanation process comprised: 5.32% of all cognitive occurrences in condition 1 (OSCE), 1.99% of all cognitive occurrences in condition 2 (CBA), and 1.68% of all cognitive occurrences in condition 3 (MEQ). The average percentages of diagnostic decision and explanation in each condition from both participants were 5.50% (OSCE), 0.99% (CBA), and 1.42% (MEQ).

Therapeutic decision and treatment options. Clinical reasoning in the form of therapeutic decision and treatment options for participant 1 was 5.69% of all cognitive occurrences in condition 1 (OSCE). For participant 2 in condition 2 (CBA), therapeutic decision and treatment options was 6.97% of all cognitive occurrences. Neither participant engaged in therapeutic decision and treatment options in condition 3 (MEQ). The average percentages of therapeutic decision and treatment options in each condition from both participants were 4.66% (OSCE) and 4.98% (CBA).

Total percentages of clinical reasoning. As shown in Table 4.28., among all identified clinical reasoning in the three different types of assessments, participant 1 displayed 79.67% of the cognitive occurrences from the participant's reconstructed narratives were identified as clinical reasoning in condition 1 (OSCE). 61.37% of the cognitive occurrences were categorized as clinical reasoning in condition 2 (CBA), and 63.74% in condition 3 (MEQ). Participant 2 showed slightly different results than participant 1. 86.57% of participant 2's cognitive occurrences were coded as clinical reasoning for condition 2 (CBA). For participant 2, condition 1 (OSCE) and condition 3 (MEQ) saw, 69.44% and 67.04% identified as clinical reasoning, respectively. The average percentages for clinical reasoning across all conditions from both participants were 74.56% (OSCE), 73.97% (CBA), and 65.39% (MEQ).

Graphical representations to display the differences of the clinical reasoning processes in the three different types of clinical assessments is shown below in Figures.

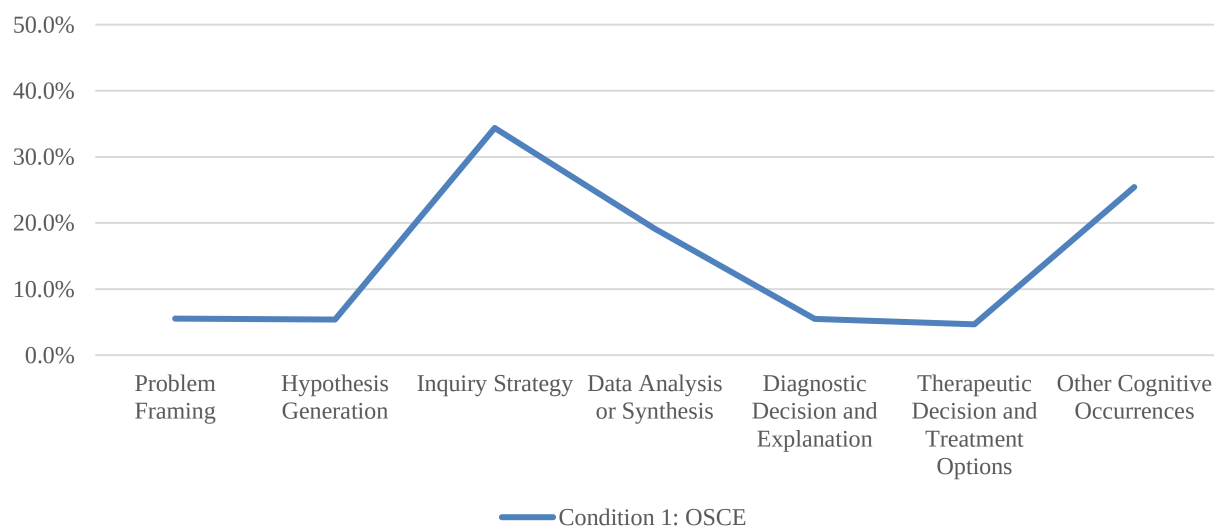


Figure 4.10. A graphical representation of clinical reasoning processes in condition 1

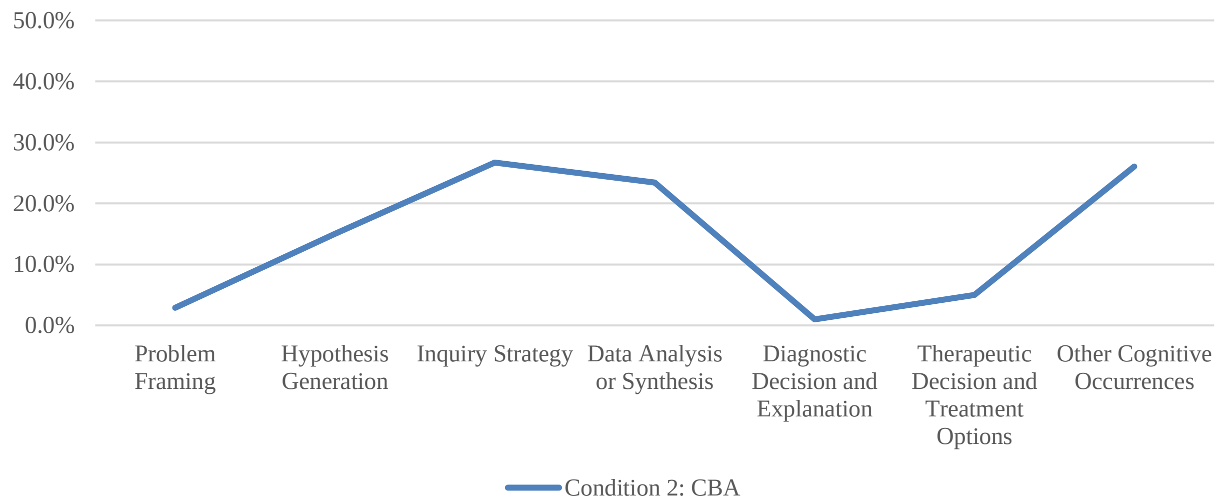


Figure 4.11. A graphical representation of clinical reasoning processes in condition 2

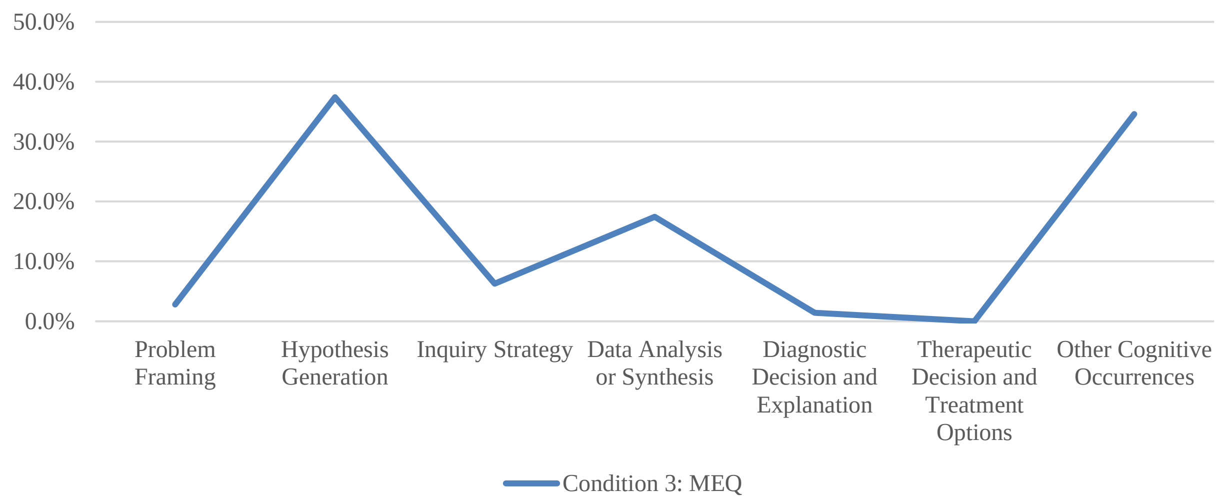


Figure 4.12. A graphical representation of clinical reasoning processes in condition 3

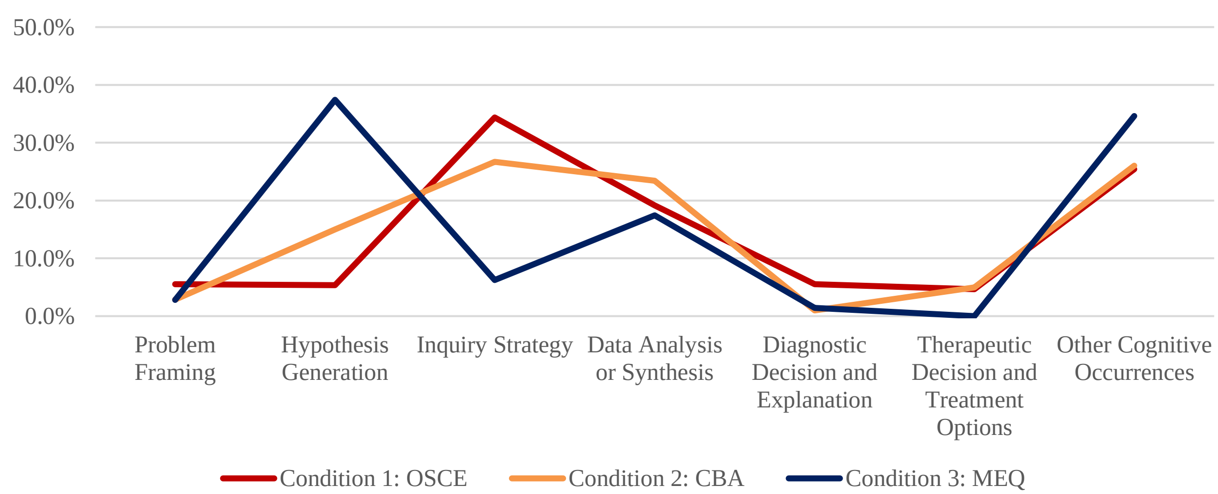


Figure 4.13. A graphical representation of clinical reasoning processes among conditions

Other Cognitive Occurrences

In order to understand the differences among other cognitive occurrences detected in the three different types of clinical assessments, the analyzed data from the cognitive occurrences of the participants' reconstructed narratives for other cognitive occurrences are presented. Due to the differences of other cognitive occurrences in each condition, the results from an individual condition are provided in this section.

Other cognitive occurrences – condition 1 (OCSE). The aggregated percentage data of each other cognitive occurrences theme from condition 1 (OSCE) is provided in Table 4.29.

Table 4.29.
Percentages of other cognitive occurrences occurring in condition 1 (OSCE)

Other Cognitive Occurrences		Percentage in entire cognitive occurrences	Percentage in other cognitive occurrences only
Test-taking Strategy	Participant 1	**15.45%**	**76.00%**
	Participant 2	**12.29%**	**40.22%**
Inattentive Action	Participant 1	7.32%	36.00%
	Participant 2	0.33%	1.09%
Off-protocol Behavior	Participant 1	5.69%	28.00%
	Participant 2	8.31%	27.17%
Educated Guessing	Participant 1	2.44%	12.00%
	Participant 2	3.65%	11.96%
Point-seeking/hunting	Participant 1	**4.88%**	**24.00%**
	Participant 2	**18.27%**	**59.78%**
Unnecessary Behavior	Participant 1	4.07%	20.00%
	Participant 2	0.66%	2.17%
Checklist Awareness	Participant 1	0.81%	4.00%
	Participant 2	17.61%	57.61%
Total	Participant 1	**20.33%**	**100.00%**
	Participant 2	**30.56%**	**100.00%**

The result showed that two major themes were identified from condition 1 (OSCE): test-taking strategy and point-seeking/hunting. The total other cognitive occurrences detected from participant 1 and 2 were 20.33% and 30.56%, respectively. The two participants showed different other cognitive occurrences based on the major themes. The most other cognitive occurrences from participant 1 was test-taking strategy (15.45%, 76.00%). On the other hand, the most other cognitive occurrences from participant 2 was point-seeking/hunting (18.27%, 59.78%).

Other cognitive occurrences – condition 2 (CBA). The aggregated percentage data of each other cognitive occurrences theme from condition 2 (CBA) is presented in Table 4.30.

Table 4.30.
Percentages of other cognitive occurrences occurring in condition 2 (CBA)

Other Cognitive Occurrences		Percentage in entire cognitive occurrences	Percentage in other cognitive occurrences only
Test-taking Strategy	Participant 1	**23.18%**	**60.00%**
	Participant 2	**8.97%**	**66.66%**
Test-taking Thinking	Participant 1	12.45%	32.22%
	Participant 2	3.97%	29.63%
Off-protocol Behavior	Participant 1	8.58%	22.22%
	Participant 2	2.00%	14.81%
Educated Guessing	Participant 1	2.15%	5.56%
	Participant 2	2.98%	22.22%
Point-seeking/hunting	Participant 1	**10.73%**	**27.78%**
	Participant 2	**1.00%**	**7.41%**
Unnecessary Constraints	Participant 1	**4.72%**	**12.22%**
	Participant 2	**3.48%**	**25.93%**
Total	Participant 1	**38.63%**	**100.00%**
	Participant 2	**13.43%**	**100.00%**

In condition 2 (CBA), the result showed that three major themes were detected: test-taking strategy, point-seeking/hunting, and unnecessary constraints. The total other cognitive occurrences identified from the participant 1 and 2 were 38.63% and 13.43%, respectively. Participant 1 presented more other cognitive occurrences than participant 2 in all three major themes. The most other cognitive occurrences from both participants was test-taking strategy (60.00%, 66.66%) in condition 2. Participant 1 showed point-seeking/hunting (27.78%) more than participant 2 (7.41%). However, for unnecessary constraints, participant 2 (25.93%) showed more than participant 1 (12.22%).

Other cognitive occurrences – condition 3 (MEQ). The aggregated percentage data of each other cognitive occurrences theme from condition 3 (MEQ) is presented in Table 4.31.

Table 4.31.
Percentages of other cognitive occurrences occurring in condition 3 (MEQ)

Other Cognitive Occurrences		Percentage in entire cognitive occurrences	Percentage in other cognitive occurrences only
Test-taking Strategy	Participant 1	**22.22%**	**43.55%**
	Participant 2	**18.99%**	**57.61%**
Test-taking Thinking	Participant 1	7.60%	20.97%
	Participant 2	3.91%	11.86%
Off-protocol Behavior	Participant 1	8.19%	22.58%
	Participant 2	12.29%	37.29%
Educated Guessing	Participant 1	6.43%	17.74%
	Participant 2	2.79%	8.48%
Point-seeking/hunting	Participant 1	**9.36%**	**25.81%**
	Participant 2	**5.03%**	**15.25%**
Unnecessary Constraints	Participant 1	**4.68%**	**12.90%**
	Participant 2	**8.94%**	**27.12%**
Total	Participant 1	**36.26%**	**100.00%**
	Participant 2	**32.96%**	**100.00%**

Three major themes were detected in condition 3 (MEQ): test-taking strategy, point-seeking/hunting, and unnecessary constraints. The total other cognitive occurrences identified from participant 1 and 2 were 36.26% and 32.96%, respectively. The most other cognitive occurrences for both participants was test-taking strategy (43.55%, 57.61%). The two participants showed different other cognitive occurrences for point-seeking/hunting and unnecessary constraints. Participant 1 showed more point-seeking/hunting (25.81%) than participant 2 (15.25%). However, participant 2 (27.12%) was more distracted by unnecessary constraints in condition 3 than participant 1 (12.90%).

CHAPTER 5

FINDINGS, DISCUSSION, AND CONCLUSIONS

This study was conducted to identify medical students' cognitive processes in solving problems from three different types of assessments (the OSCE, CBA, and MEQ) and to understand how these assessments promote different kinds of problem-solving by investigating the differences in clinical reasoning processes among the assessments. Based on the research questions provided in chapter 1, chapter 2 involved reviewing related literature to initially understand diagnostic reasoning models, different types of clinical examinations, and assessing real-world problem-solving. Data were collected and analyzed to discover the similarities and differences of the cognitive processes between each assessment, and detailed methods for data collection and analysis were provided in chapter 3. Chapter 4 presented the results from data analysis, including the frequencies and percentages of clinical reasoning and other cognitive occurrences identified from the reconstructed narratives of occurrences, visualizations of the cognitive processes, representative quotes of clinical reasoning and other cognitive occurrences, and comparisons of clinical reasoning and other cognitive occurrences among the three different types of assessments.

This chapter provides key findings from the results and discusses implications and limitations of the study before providing conclusions and future directions for research.

Findings

In order to identify medical students' cognitive processes while solving problems in different types of clinical assessments, a qualitative case study was designed to elicit research participants' thoughts in solving problems through stimulated recall interviews. A total of three conditions were prepared (the OSCE, CBA, and MEQ) as different types of clinical assessments, and two students participated in all three conditions. Data from the stimulated recall interviews on the research participants' performances in the three different types of assessments were collected and analyzed according to the specified protocol in the study. With the interview data, the smallest phrases or sentences representing a meaningful cognitive occurrence was used as the unit of analysis, and the data analysis was conducted with the following steps: data classification, data coding, un-coded data classification, data re-organization into chronological order, and data visualization. This section provides a brief summary of the results and key findings.

Summary of Results

The results were provided according to the research questions in chapter 4. With research question 1, each analyzed result from the six cases (two participants in three conditions) and the proportional frequencies of all types of cognitive occurrences for each participant were similar. Then, a cross-case analysis was conducted to understand the differences of clinical reasoning and other cognitive occurrences among the three conditions for research question 2.

First, hypothetico-deductive reasoning processes as clinical reasoning and other cognitive occurrences themes detected in the three different types of assessments were presented from each case under research question 1. All meaningful cognitive occurrences were coded through hypothetico-deductive reasoning processes: problem framing, hypothesis generation, inquiry strategy, data analysis or synthesis, diagnostic decision and explanation, and therapeutic decision and treatment options. A summary of each result of the six cases is provided below.

Case 1.1 showed that total 123 cognitive occurrences were identified. 79.67% (98) of the cognitive occurrences were coded as clinical reasoning, and 20.33% (25) of the cognitive occurrences were detected as other cognitive occurrences during the OSCE. The most frequent cognitive process as clinical reasoning in case 1 was inquiry strategy (49, 39.84%). 36.00% (9) of other cognitive occurrences were detected as inattentive action, which is when the examinee pretends to be doing a physical examination to a standardized patient or is thinking about the next steps to take in the exam instead of actually observing the results of performing the physical examination.

Case 1.2 had a total of 301 cognitive occurrences that were coded. 69.44% (209) of the cognitive occurrences were identified as clinical reasoning in the OSCE, and 30.56% (92) of cognitive occurrences were detected as other cognitive occurrences. The most frequent clinical reasoning process in case 1.2 was, like case 1.1, inquiry strategy. More than half of the other cognitive occurrences in case 1.2 were checklist awareness (53, 57.61%), which is a behavior where the examinee focuses on seeking hints from the test and/or hunting for the correct answer only.

Case 2.1 identified a total of 233 cognitive occurrences. 61.37% (143) of the cognitive occurrences were coded as clinical reasoning, and 38.63% (90) of cognitive occurrences were detected as other cognitive occurrences during the CBA. The most frequent clinical reasoning process was data analysis or synthesis (50, 21.46%); inquiry strategy (49, 21.03%) was also a frequent process. 32.22% (29) of other cognitive occurrences were detected as test-taking thinking, which is when the examinee displays reasoning or thinking related to testing situations rather than solving problems, and 27.78% (25) were point-seeking/hunting behavior similar to the results of case 1.2.

Case 2.2 had a total of 201 cognitive occurrences. 86.57% (174) of the cognitive occurrences were identified as clinical reasoning, and 13.43% (27) of cognitive occurrences were detected as other cognitive occurrences during the CBA. The most frequent clinical reasoning process was inquiry strategy (65, 32.34%), followed by data analysis or synthesis (51, 25.37%). For other cognitive occurrences in case 2.2, test-taking thinking (8, 29.63%) and unnecessary constraints (7, 25.93%) were frequent other cognitive occurrences.

Case 3.1 identified a total of 171 cognitive occurrences. 63.74% (109) of clinical reasoning were coded as hypothetico-deductive reasoning processes, and 36.26% (62) of the cognitive occurrences were detected as other cognitive occurrences during the MEQ. The most frequent clinical reasoning in case 3.1 was hypothesis generation (63, 36.84%). 25.81% (16) of other cognitive occurrences were detected as point-seeking/hunting, followed by off-protocol behavior (14, 22.58%) and test-taking thinking (13, 20.97%).

Case 3.2 identified a total of 179 cognitive occurrences. 67.04% (120) of the cognitive occurrences were coded as clinical reasoning, and 32.96% (59) other cognitive occurrences were detected during the MEQ. The most frequent clinical reasoning was hypothesis generation (68, 37.99%). In terms of other cognitive occurrences in case 3.2, off-protocol behavior (22, 37.29%) was the most frequent. Followed by unnecessary constraints (16, 27.12%).

In order to identify the similarities of the clinical reasoning processes that occurred between the two participants in each condition, a comparative analysis was conducted through graphical representations. The results revealed that the proportional frequencies of all types of cognitive occurrences for each participant were similar. The inquiry strategy phase was the most frequent clinical reasoning process in condition 1 and condition 2, and the hypothesis generation phase was the most frequent clinical reasoning process in condition 3. In particular, both participants' clinical reasoning process patterns were almost identical in condition 3.

Second, a cross-case analysis was conducted to compare clinical reasoning processes and other cognitive occurrences among the three different types of clinical assessments according to research question 2. Under the second research question, two main analysis results were presented: Comparisons of the average percentage results of clinical reasoning processes among clinical assessments and comparisons of cognitive occurrences in each condition. A summary of the cross-case analysis results is provided below.

Among the three different types of assessments, the average percentages of all identified clinical reasoning in condition 1 (OSCE) was 74.56%; the average percentage in condition 2 (CBA) was 73.97%; the average percentage in condition 3 (MEQ) was 65.39%. In the problem framing phase, the average percentages of clinical reasoning in each condition were 5.53% (OSCE), 2.89% (CBA), and 2.83% (MEQ). In the hypothesis generation phase, the average percentages of clinical reasoning in each condition were 5.36% (OSCE), 15.00% (CBA), 37.42% (MEQ). In the inquiry strategy phase, the average percentages of clinical reasoning in each condition were 34.37% (OSCE), 26.69% (CBA), and 6.27% (MEQ). In the data analysis or synthesis phase, the average percentages of clinical reasoning in each condition were 19.14% (OSCE), 23.42% (CBA), and 17.45% (MEQ). In the diagnostic decision and explanation phase, the average percentages of clinical reasoning in each condition were 5.50% (OSCE), 0.99% (CBA), and 1.42% (MEQ). In the therapeutic decision and treatment options phase, the average percentages of clinical reasoning in each condition were 4.66% (OSCE) and 4.98% (CBA), and there no therapeutic decision and treatment options were detected in condition 3 (MEQ).

Additionally, due to the differences in each condition of the other cognitive occurrences, the aggregated results in each condition were presented.

In condition 1 (OSCE), two major themes of other cognitive occurrences were detected: test-taking strategy and point-seeking/hunting. Under test-taking strategy, inattentive action, off-protocol behavior, and educated guessing were found, and under point-seeking/hunting, unnecessary behavior and checklist awareness were identified. The two participants showed different amounts of other cognitive occurrences based on the major themes. The total other cognitive occurrences detected from participant 1 and 2 were 20.33% and 30.56%, respectively. The most other cognitive occurrences from participant 1 was test-taking strategy (15.45%, 76.00%); on the other hand, the most other cognitive occurrences from participant 2 was point-seeking/hunting (18.27%, 59.78%).

In condition 2 (CBA), three major themes of other cognitive occurrences were identified: test-taking strategy, point-seeking/hunting, and unnecessary constraints. Under test-taking strategy, test-taking thinking, off-protocol behavior, and educated guessing were detected. The two participants also showed different amounts of other cognitive occurrences based on the major themes. The total other cognitive occurrences identified from participant 1 and 2 were 38.63% and 13.43%, respectively. The most other cognitive occurrences from both participants was test-taking strategy (60.00%, 66.66%).

In condition 3 (MEQ), three major themes of other cognitive occurrences were coded the same as the CBA: test-taking strategy, point-seeking/hunting, and unnecessary constraints. Under the test-taking strategy, test-taking thinking, off-protocol behavior, and educated guessing were detected. The total other cognitive occurrences identified from participant 1 and 2 were 36.26% and 32.96%, respectively. The most other cognitive occurrences from both participants was test-taking strategy (43.55%, 57.61%).

Key Findings

Medical students' cognitive processes while they solve clinical problems in three different types of assessments were identified through stimulated recall interviews. First, clinical reasoning processes in each condition were coded and presented based on the hypothetico-deductive reasoning process. Second, any un-coded data from the reconstructed narratives of the participants were categorized as other cognitive occurrences.

Clinical reasoning (hypothetico-deductive reasoning). Based on the results, the frequency patterns found in each assessment reflect that each assessment may measure the components of clinical reasoning in a way congruent with the assessment's suggested intentions. Furthermore, the results show the specific components of clinical reasoning that are present in the different types of assessments and the degree to which each assessment addresses those specific components of clinical reasoning. For example, the clinical reasoning featured in the OSCE is primarily inquiry strategy, the CBA primarily features inquiry strategy and data analysis or synthesis, and the MEQ primarily features hypothesis generation. The result revealed that participants in condition 1 (OSCE) and condition 2 (CBA) use clinical reasoning more than in condition 3 (MEQ). The participants' hypothetico-deductive reasoning identified as clinical reasoning accounted for 74.56% of all cognitive occurrences in condition 1. Also, the participants' identified clinical reasoning accounted for 73.97% of all cognitive occurrences in condition 2. The participants' clinical reasoning accounted for 65.39% of all cognitive occurrences in condition 3. This result indicates that the clinical competence examination in the form of the OSCE and the multimedia case-based assessment promote participants' clinical reasoning more than the clinical knowledge and reasoning examination in the form of the MEQ.

The result is further interpreted in the following way by comparing the three different types of assessments for each hypothetico-deductive reasoning (HDR) process.

Problem framing. The percentages of problem framing were 5.53%, 2.89%, and 2.83% for the OSCE, CBA, and MEQ, respectively. As presented in Table 4.28, clinical reasoning for this HDR process was more present in condition 1 (5.53%) than in the other two conditions (2.89%, 2.83%). According to the hypothetico-deductive reasoning process, the problem framing phase includes perceiving a variety of cues, listening to patients' initial complaints, and asking questions to form an initial concept of the patients' problems (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). The participants encountered an actual patient only in condition 1 and had to perceive relevant information by asking questions to begin their problem-solving. In direct contrast, the other two conditions had all related information provided to the participant in the form of a video or written scenario without active clinical reasoning. Due to the nature of the OSCE, the participants displayed more clinical reasoning in condition 1 rather than condition 2 and 3 for this process of the HDR.

Hypothesis generation. Clinical reasoning from the participants in condition 3 accounted for 37.42% of all cognitive occurrences. The hypothesis generation phase includes all reasoning related to generating as many hypotheses as possible and ranking the hypotheses based on the cues and data from patients (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). One of the test items in the MEQ was to draw a concept map that included all possible processes from the cue to the disease based on the pathophysiological systems, anatomical locations, and so on, all linking to the clinical presentation provided in the test. Moreover, other test items in the MEQ were also related to generating hypotheses. This result showed that the concept map question facilitated the participants use of clinical reasoning for hypothesis generation in the MEQ.

For similar reasons, hypothesis generation as clinical reasoning in condition 2 accounted for 15.67% of all cognitive occurrences. In the multimedia case-based assessment, each page of the assessment had a test item that asked participants to formulate clinical hypotheses or to list possible clinical hypotheses and to explain their decision. This test item also facilitated the participants to generate relevant hypotheses through clinical reasoning.

On the other hand, only 5.36% of cognitive occurrences were identified as hypothesis generation process in condition 1 because the participants tended to have only one or two hypotheses: They had to narrow down their hypotheses using the data they could perceive and reach a final diagnosis rapidly due to the time pressure and point-seeking/hunting behavior. Both participants described this behavior in the interview.

> *"I asked about radiating pain, and the patient said the pain extended from her chest to chin. This was a typical answer. When I learned about angina, I learned that the typical patterns were pain from chest to chin or from chest to shoulder" (N22-N25 in case 1.1)*

> *"So, I somewhat thought I needed to meet those test requirements quickly and move on … I thought, if the patient's pain was in the center of her chest, most medical students typically think the diagnosis will be related to the circulatory system: angina" (N80, N82-N83 in case 1.2)*

Inquiry strategy. The inquiry strategy phase was the most frequent clinical reasoning process in both conditions 1 and 2. This clinical reasoning from the participants in condition 1 accounted for 34.37% of all cognitive occurrences; and in condition 2, inquiry strategy accounted for 26.69% of all cognitive occurrences. Inquiry strategy includes asking questions to rule out hypotheses, taking medical, family, and social histories, conducting physical examination, and so on (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). The result suggests that the inquiry strategy phase may be a core process of clinical reasoning, and the participants showed the same result in conditions 1 and 2, but not in condition 3. Due to the nature of the MEQ, the paper-based examination does not require physical interactions and communication between examinees and patients to perform medical treatment. In the CBA, there was also no physical interactions and communication for the participants to perform medical treatment, as the OSCE was the only assessment in this study that featured standardized patients. However, the CBA included related test items that had the examinees describe what they needed to do regarding history taking, physical examinations, and/or laboratory tests to assess medical students' clinical reasoning regarding inquiry strategy.

Data analysis or synthesis. The result revealed that data analysis or synthesis is a clinical reasoning process that is prominent processes in all conditions. This kind of clinical reasoning from the participants in condition 1 accounted for 19.14% of all cognitive occurrences; 23.42% in condition 2; and 17.45% in condition 3. The data analysis or synthesis phase includes obtaining any data from patients and/or laboratory findings that help physicians make decisions (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). Also, the ongoing summary of clinical hypotheses is included in this phase. Based on the result, the participants used their clinical reasoning on data analysis or synthesis phase similarly in all three different types of assessments. Thus, the participants were asked to analyze the data given in the conditions, and these three conditions include an appropriate portion of test items that affords this type of clinical reasoning.

Diagnostic decision and explanation. The result showed that the diagnostic decision and explanation was the least frequent clinical reasoning process in condition 2 and 3. This type of clinical reasoning accounted for 0.99% of all cognitive occurrences in condition 2 and 1.42% in condition 3. In particular, there were no diagnostic decision and explanation detected from participant 1 in condition 2. Diagnostic decision and explanation includes evaluating each hypothesis with data analysis results and making a decision on an initial diagnosis (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). Furthermore, explaining the initial diagnosis to patients was included in this phase. The reason for the lack of this clinical reasoning process in conditions 2 and 3 is that these two assessments focused on forming an initial diagnosis from the clinical presentation and given information, further data, such as laboratory results or other important examinations, beyond physical examination data were not provided on the tests.

Therapeutic decision and treatment options. The result supports there being no meaningful difference between condition 1 (4.66%) and 2 (4.98%) for this type of clinical reasoning. However, there was a big difference between condition 3 and conditions 1 and 2. No clinical reasoning was identified for the therapeutic decision and treatment options in condition 3 from both participants. The therapeutic decision and treatment options includes deciding appropriate management plans, such as surgery or medication plans, providing possible treatment options, and so on (Barrows, 1994; Barrows & Tamblyn, 1980; Ju & Choi, 2018). Because the test items in the MEQ does not include the therapeutic decision and treatment options process, the process was not detected from the participants.

Similarity of clinical reasoning processes between two participants. Based on the result from research question 1, the study found that the two participants had similarities in their clinical reasoning processes. In condition 1 and 2, both participants showed a similar clinical reasoning pattern while solving clinical problems; in condition 3, participants' clinical reasoning processes were almost identical. Based on the result, each clinical assessment promotes select clinical reasoning processes. For example, the OSCE promotes inquiry strategy and data analysis or syntheses more than other processes, the CBA promotes hypothesis generation, inquiry strategy, and data analysis or synthesis more than other processes, and the MEQ promotes hypothesis generation and data analysis or synthesis over other processes. And these patterns were identified from both participants.

Other Cognitive Occurrences. In order to understand the differences of clinical reasoning processes occurring in the three different types of clinical assessments, other cognitive occurrences were also identified to examine what other types of thinking occurred instead of clinical reasoning while medical students solve clinical problems in each assessment.

In condition 1, two major themes of other cognitive occurrences were detected: test-taking strategy and point-seeking/hunting. Although the patterns of other cognitive occurrences from the two participants were different, five other cognitive occurrences, including inattentive action, off-protocol behavior, educated guessing, unnecessary behavior, and checklist awareness, which fell under the major two themes, were constantly detected as distractors to participants' clinical problem-solving. In particular, there were some other cognitive occurrences identified in condition 1 due only to the nature of the OSCE. In the case of inattentive action, participant 1 tended to think of the next steps while the person performed different clinical procedures or pretended to perform physical examinations.

Moreover, both participants explained that many hints or cues from the test itself were revealed to them, which owes to the limitations of the clinical scenario and/or standardized patients utilized in the OSCE. This is related to the limitations of the OSCE because in its current form, the OSCE can only cover a limited number of clinical scenarios using standardized patients. Also, the students easily ignore clinical contexts in testing because of the fact that the standardized patients are not real patients but healthy actors. Unnecessary behaviors were also detected due to students considering the other cognitive occurrences that could lead to higher scores rather than solving clinical problems in the OSCE. Because the students were aware of being graded and anticipated the grading rubric while they performed clinical treatments during the exam, they tended to ask additional questions or perform additional physical examinations to cover the checklist, even if they thought these behaviors were unnecessary.

In condition 2, three major themes of other cognitive occurrences were detected: test-taking strategy, point-seeking/hunting, and unnecessary constraints. Although patterns of other cognitive occurrences from the two participants were different, five other cognitive occurrences, including test-taking thinking, off-protocol behavior, educated guessing, point-seeking/hunting, and unnecessary constraints, were constantly detected as distractors to participants' clinical problem-solving. Similar to the results in condition 1, test-taking strategy and point-seeking/hunting were detected for both participants. The theme of unnecessary constraints was added for condition 2 due to the nature of the multimedia case-based assessment. The participants were distracted by the test design or test item design; for example, participant 1 did not know the case videos are rewindable and thus missed important information from the video cases. Because of this issue, participant 1 had to solve problems by guessing the patient's information that was missed.

Moreover, both participants struggled with deciding the scope of their answers according to the test items. Four sections were developed based on the clinical procedures in the CBA, and the participants were able to proceed to the next section only when they had finished the previous section. Due to this inability of test navigation in the CBA, the participants were not able to know what situations would happen in the next video unlike they can actively ask questions or perform clinical procedure according to their decisions.

Condition 3 had the same three major themes of other cognitive occurrences as condition 2: test-taking strategy, point-seeking/hunting, and unnecessary constraints. Both participants showed that test-taking strategy was the most frequent other cognitive occurrences. One of the test items in the MEQ was to draw a concept map that included all possible processes from the cue to the disease based on the pathophysiological systems, anatomical locations, and so on, regarding the clinical presentation provided in the test. Although the test item helped the participants generate related hypotheses for solving problems, both participants were distracted by writing down all possibilities within a limited time and with limited knowledge.

Based on clinical reasoning processes and other cognitive occurrences identified and detected in the three different types of clinical assessments, the results support the fact that each clinical assessment promotes a particular set of reasoning processes and affords distractions that prevent medical students from solving problems clinically.

Discussion

The findings of this study showed that the clinical reasoning patterns of both participants were highly similar in each clinical assessment. The participants presented different patterns of clinical reasoning among the three different types of clinical assessments. This section discusses the implications of the study based on the results provided. The remainder of the section discusses the limitations of the study.

Implications of the Study for Research and Practice

Use of the research design as another way of checking test validity. This study was designed to identify medical students’ cognitive processes in solving diagnostic problems in three different types of clinical assessments and to understand the differences of cognitive processes among the three clinical assessments using stimulated recall interviews. The findings revealed that each clinical assessment tested in the study promotes particular reasoning processes or affords distractions that prevent clinical problem-solving. Based on the findings, the research design that was utilized to identify medical students' cognitive processes using the stimulated recall protocol may be used to check test validity, in particular for supporting construct validity. Construct representation, one of the fundamental features of construct validity, is "concerned with identifying the theoretical mechanisms that underlie task performance" through cognitive-process analysis by decomposing the task into the processes, strategies, and knowledge (Embretson, 1983, p. 180; Messick, 1995). The method used in this study may indicate an alternative way of identifying cognitive processes that underlie test performances as a form of construct representation to support construct validity.

Moreover, the test scores that medical students receive from clinical assessments and the clinical reasoning used to solve diagnostic problems in the assessments may not be connected to each other. In other words, what students receive as their scores and what students think during problem-solving may be different due to the other cognitive occurrences that have been found in this study. For example, numerous other cognitive occurrences were detected in the phase of physical examination in the OSCE. Even if the students receive a high score from the physical examination phase because their performances displayed correct procedure, some of the students' performances may not be indicative of clinical reasoning but test-taking strategy or point-seeking/hunting. This issue may be a significant threat to the tests' construct validity. This study was intended to explore an initial understanding of what medical students think while they are solving clinical problems in three different types of clinical assessments; therefore, there was a limitation of examining the differences between the medical students' test scores and clinical reasoning cognitive process. In a future study, the research design utilized in this study may be used to explore alternative ways of checking construct validity.

Redesign of clinical assessments, including the multimedia case-based assessment. Another implication of the study is that it provides design guidelines for the OSCE, MEQ, and the multimedia case-based assessment as it was used in this study. Based on the findings, the results showed that each assessment promotes different clinical reasoning processes. Although the OSCE and MEQ have been developed and widely implemented for assessing medical students' diagnostic reasoning abilities, only a small portion of the entire process of hypothetico-deductive reasoning was identified in the OSCE and MEQ, or certain processes were not presented while the students solved diagnostic problems. The multimedia case-based assessment (CBA) was developed and implemented alongside the other two clinical assessments, and the CBA provided a pattern similar to that of the clinical reasoning process in the OSCE. Assuming that this points to the multimedia case-based assessment as providing a higher degree of authenticity in terms of providing real-world context in a testing environment than the MEQ, the CBA may be considered as a reasonable and affordable way to assess clinical reasoning.

Each type of assessment has its advantages regarding development, implementation, administration, grading, and cognitive affordances of clinical reasoning. Moreover, each assessment strongly promotes a certain clinical reasoning process. Therefore, the finding may provide design guidelines for revising each clinical assessment to promote a balanced way of assessing clinical reasoning abilities through all three clinical assessments.

Limitations of the Study

The sample size. Two medical students were selected and participated based on their GPA and previous OSCE results, and three different types of clinical assessments were used for the study. Based on the research design of a cross-case study, the research findings were drawn from the unique results of each case result. Although a total of six cases were collected and analyzed for the results, this study utilized a small sample size of medical students from one school and one clinical presentation (chest pain). Moreover, the selected participants had high GPAs and qualifying results from previous OSCEs to eliminate bias from a lack of knowledge. Different levels of clinical knowledge and clinical performance might lead to different results in the study. Although the reason for selecting two high-GPA students was eliminate bias from a lack of knowledge, the selected two participants may provide skewed results. Therefore, different sampling methods and different clinical presentations for the assessments may lead to different results, and this study's findings may not be generalizable. For future research on the findings here, a validation with a larger sample size and different GPA levels of participants should be conducted.

Furthermore, themes of other cognitive occurrences from each clinical assessment were detected. The findings revealed that some patterns of other cognitive occurrences were similar to each other in each assessment; however, the percentages of other cognitive occurrences detected from each participant were different. If more data from different participants can be collected and analyzed, more solid findings based on other cognitive occurrences can be provided in future studies.

The collected data were based on the narratives of participants' recalls. This study used the participants' narratives to identify their clinical reasoning and other cognitive occurrences for data analysis. All data were collected based on the participants' recalls through stimulated recall interviews. Even though the data were collected by asking questions related to directly their performances on clinical assessments, the data completely relied on the participants' recalls. One bias may be that their narratives reflected only positive occurrences as their status as high-performing medical students; also possible is they missed some important cognitive occurrences because they did not catch their performances in their performance videos. For example, when solving clinical problems in the OSCE, comprehensive and intensive cognitive processes are required to generate hypotheses. The depth of cognitive processes, however, might not always have been detectable from the participants' narratives and therefore were not coded as hypothesis generation. The CBA and MEQ, on the other hand, had test items that explicitly asked examinees to generate hypotheses.

Therefore, identified cognitive occurrences from the participants' recalled narratives might be biased. In particular, the stimulated recall interviews were designed to extract the participants' thought processes when they encountered clinical problems during the assessments. The participants needed to explain or mention what they thought, saw, heard, or did regarding their performances. However, in some cases, cognitive occurrences were initially coded as clinical reasoning or other cognitive occurrences from collected data, including the participants' thoughts regarding their performances in the interviews as they were watching their performances. While watching the video of themselves, they could assess their own performance or do post-assessment thinking, such as what they should have done or could have done differently, instead of revealing what they were actually thinking during the moment they were doing the exam. Although the data were eliminated if the data were clearly detectable as post-assessment additions during data cleaning, any undetected data or other narrative data that was not from the retrospective interviews may be a serious threat to the data reliability. For future research, revisions of interview questions in the stimulated recall and interview protocol will be needed.

Test-retest bias based on the research design. In this study, three different types of clinical assessments were used for data collection and analysis. In order to avoid test-retest biases, each assessment was given to the participants in one-week intervals. Moreover, each assessment was developed and implemented with a different final diagnosis from the same clinical presentation and considered to be a different assessment. However, a participant revealed in the interview that they assumed the purpose of this research was a comparison among a fixed clinical presentation but with variable final diagnosis, and they tried to skip or cut short their clinical reasoning to jump to the diagnosis. In order to avoid this test-retest bias, which is a threat to reliability of the research, more time will be needed to collect data among assessments for the future research.

Conclusions

The purpose of the study was to identify medical students' cognitive processes in solving diagnostic problems and to compare how they think differently in three different types of clinical assessments, including multimedia case-based assessment (CBA). In the context of medical education, the objective structured clinical examination (OSCE) and modified essay question (MEQ) have been widely used to assess medical students' diagnostic reasoning. However, little attention has been given to the types of thinking that medical students actually engage in during these assessments.

A cross-case study was employed with two 4th year medical students in three different clinical assessments, the OSCE, CBA, and MEQ. The findings revealed that participants in the

OSCE (74.56%) and the CBA (73.97%) use clinical reasoning more than in the MEQ (65.39%). In particular, the inquiry strategy phase of hypothetico-deductive reasoning (HDR) was the most frequent clinical reasoning process in both the OSCE and CBA. Whereas the data analysis or synthesis phase also occurred frequently in all assessments, diagnostic decision and explanation and therapeutic decision and treatment options phases were rarely found in the MEQ. Moreover, other cognitive occurrences were detected in all assessments (25.44%, 26.03%, and 34.61%; the OSCE, CBA, and MEQ, respectively). Two main other cognitive occurrences were test-taking strategy and point-seeking/hunting, and sub-themes were identified, such as inattentive action, off-protocol behavior, educated guessing, unnecessary behavior, and checklist awareness. This result indicates that the participants were more distracted by other cognitive occurrences in the MEQ than in the OSCE and CBA.

In summary, the findings of the study suggest that each clinical assessment used in this study may strongly promote select diagnostic reasoning processes and that the clinical reasoning process may be distracted by other cognitive occurrences afforded by the assessment. As a result, suggestions for future research are provided. This study's research method may be used to validate clinical assessments for assessing medical students' diagnostic reasoning, and the results of this study can inform the design of clinical assessments as well as the multimedia case-based assessment for medical education.

REFERENCES

Accreditation Council for Graduate Medical Education (ACGME), & American Board of Medical Specialties (ABMS). (2000). *Toolbox of assessment methods* (Version 1.1). Retrieved from http://njms.rutgers.edu/culweb/medical/documents/ToolboxofAssessmentMethods.pdf

Anderson, D. C., Harris, I. B., Allen, S., Satran, L., Bland, C. J., Davis-Feickert, J. A., … Miller, W. J. (1991). Comparing students' feedback about clinical instruction with their performances. *Academic Medicine, 66*(1), 29-34.

Allen, B. S., Otto, R. G., & Hoffman, B. (2000). Case-based learning: Contexts and communities of practice. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 443-471). New York, NY: Macmillan/Gale Group.

Barnes, L. B., Christensen, C. R., & Hansen, A. J. (1994). *Teaching and the case method: Instructor's guide* (3rd ed.). Boston, MA: Harvard Business Publishing.

Barrows, H. S. (1985). *How to design a problem-based curriculum for the preclinical years*. New York, NY: Springer.

Barrows, H. S. (1993). An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Academic Medicine*, *68*(6), 443-451.

Barrows, H. S. (1994). *Problem-based learning applied to medical education*. Springfield, IL: Southern Illinois University.

Barrows, H. S., & Tamblyn, R. M. (1980). *Problem-based learning: An approach in medical education*. New York, NY: Springer Publishing Company.

Beers, P. J., Boshuizen, H. P. A., Kirschner, P. A., Gijselaers, W., & Westendorp, J. (2008). Cognitive load measurements and stimulated recall interviews for studying the effects of information and communications technology. *Educational Technology Research & Development. 56*, 309-328.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347-364.

Carpenter, J. L. (1995). Cost analysis of objective structured clinical examinations. *Academic Medicine, 70*, 828-833.

Chi, M. T. H., & Glaser, R. (1985). Problem solving ability. In R. J. Sternberg (Ed.), *Human abilities: An information processing approach*. New York, NY: W. H. Freeman.

Chipman, S. F., Schraagen, J. M., & Shalin, V. L. (2000). Introduction to cognitive task analysis. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (eds), *Cognitive task analysis.* Mahwah, NJ: Lawrence Erlbaum Associates.

Cho, J., Kim, J., Park, H., & Hwang, I. (2011). Correlation of CPX scores with the scores on written multiple-choice examinations on the certifying examination for family medicine in 2009 to 2011. *Korean Journal of Medical Education, 23*(4), 315-322.

Choi, E., & Sunwoo, S. (2009). Correlations of clinical assessment tools with written examinations. *Korean Journal of Medical Education, 21*(1), 43-52. doi: 10.3946/kjme.2009.21.1.43

Choi, I. & Lee, K. (2008). A case-based learning environment design for real-world classroom management Problem Solving. *Tech Trends, 52*(3). 26-31.

Choi, I., & Lee, K. (2009). Designing and implementing a case-based learning environment for enhancing ill-structured problem solving: classroom management problems for prospective teachers. *Educational Technology Research & Development*, *57*(1), 99-129.

Choi, I., Lee, S., & Kang, J. (2009). Implementing a case-based e-learning environment in a lecture-oriented anesthesiology class: Do learning styles matter on complex problem solving over time? *British Journal of Educational Technology*. *40*, 933-947.

Choi, I., Hong Y., Park, H., & Lee, Y. (2013). Case-based learning for anesthesiology: Enhancing dynamic decision-making skills through cognitive apprenticeship and cognitive flexibility. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, & N. Winters (Eds.), *Handbook of design in educational technology* (pp. 230-240). New York, NY: Routledge.

Choi, S., & Koh, S. (2008). Introduction of Objective Structured Clinical Examination (OSCE) to residency examination. *Korean Journal of Medical Education*, *20*(4), 343-349.

Coderre, S., Mandin, H., Harasym, P. H., & Fick, G. H. (2003). Diagnostic reasoning strategies and diagnostic success. *Medical Education, 37*. 695-703.

Crandall, B., Klein, G., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. Cambridge, MA: MIT Press.

Creevy, K. E., Cornell, K. K., Schmiedt, C., Park, H., Rong, H., Radlinsky, M. G., & Choi, I. (2017). Impact of expert commentary and student reflection on veterinary clinical decision-making skills in an innovative electronic-learning case-based platform. *Journal of Veterinary Medical Education*. 1-13.

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage Publications.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Croskerry, P. (2009). A universal model of diagnostic reasoning. *Academic Medicine, 84*(8), 1022-1028.

Danielson, J. A., Mills, E. M., Vermeer, P. J., Preast, V. A., Young, K. M., Christopher, M. M., . . . Bender, H. S., (2007). Characteristics of a cognitive tool that helps students learn diagnostic problem solving. *Educational Technology Research and Development, 55*(5), 499-520.

Dabbagh, N., & Dass, S. (2013, November). *CATS: A tool for identifying the cognitive affordances of learning technologies*. Paper presented at the 2013 Association for Educational Communication & Technology Conference, Anaheim, CA. Abstract retrieved from http://www.aect.org/pdf/proceedings13/2013i/13_10.pdf

de Vaus, D.A. (2001). *Research design in social research*. Thousand Oaks, CA: Sage.

Edwards, I., Jones, M., Carr, J., Braunack-Mayer, A., & Jensen, G. M. (2004). Clinical reasoning strategies in physical therapy. *Physical Therapy, 84*, 312-330.

Elliott, S. N. (1994). *Creating meaningful performance assessments: Fundamental concepts*. Reston, VA: The Council for Exceptional Children.

Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.

Embretson, S. (1983). Construct validity: Construct representation versus nomothetic span. *Psychological Bulletin, 93*, 179-197.

Eseryel, D., Ifenthaler, D., & Ge, X. (2013). Validation study of a method for assessing complex ill-structured problem solving by using causal representations. *Educational Technology Research and Development, 61*, 443-463.

Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgement, and social cognition. *Annual Review of Psychology, 59*, 255-278.

Feinstein, A. R. (1973). An analysis of diagnostic reasoning. I. The domains and disorders of clinical macrobiology. *Yale Journal Biology Medicine*, *46*(3), 212-232.

Feldman, R. (2009). Thinking, reasoning, and education. In H. Siegel (Ed.), *The Oxford handbook of philosophy of education* (pp. 67–82). Oxford, NY: Oxford University Press.

Feletti, G. I., & Smith, E. K. (1986). Modified essay questions: Are they worth the effort? *Medical Education*, *20*, 126-132.

Feltovich, P. J., Johnson, P. E., Moller, J. H., & Swanson, D. B. (1984). The role and development of medical knowledge in diagnostic expertise. In W. J. Clancey & E. H. Shortliffe (Eds)., *Readings in medical artificial intelligence: The first decade* (pp. 275-319). Reading, MA: Addison-Wesley.

Fredreiksen, N. (1994). The integration of testing with teaching: Applications of cognitive psychology in instruction. *American Journal of Education, 102*, 527-564.

Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language research*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Gaver, W. W. (1991). Technology affordances. *In proceedings of the CHI Conference on Human Factors in Computing Systems*, New York, NY: ACT Press, 79-84.

Gibson, E. J., & Pick, A. D. (2000). *An ecological approach to perceptual learning and development*. New York, NY: Oxford University Press.

Gibson, J. J. (1977). The theory of affordances. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing: Toward an ecological psychology* (pp. 67-82). Hillsdale, NJ: Erlbaum.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336-342.

Grossman, P. L. (1992). Teaching and learning with cases: Unanswered questions. In J. H. Shulman (Ed.), *Case methods in teacher education* (pp. 227-239). New York, NY: Teachers College Press.

Grotzer, T. A. & Perkins, D. N. (2000). A taxonomy of causal models: The conceptual leaps between models and students' reflections on them, paper presented at the *National Association of Research in Science Teaching (NARST)*, New Orleans, LA, 28 April to 1 May.

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research & Development, 52*(3), 67-86.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Gaithersburg, MD: Advanced Analytics, LLC.

Han, J., Park, H., Eo, E., Yoo, K., Lee, D., & Jung, W. (2004). An OSCE for summative assessment after clinical clerkship – experience in Ewha Medical School. *Korean Journal of Medical Education, 16*(1), 33-40.

Harden, R. M., Stevenson, M., Downie, W. W., & Wilson, G. M. (1975). Assessment of clinical competence using Objective Structured Examination. *British Medical Journal, 1*, 447-451.

Hardin, L. E. (2002). Research in medical problem solving: A Review. *Journal of Veterinary Medical Education, 30*(3), 227-232.

Hart, D. (1994). *Authentic assessment: A handbook for educators*. Menlo Park, CA: Addison-Wesley Publishing Company.

Hartson, H. R. (2003). Cognitive, physical, sensory, and functional affordances in interaction design. *Behaviour & Information Technology, 22*(5), 315-338.

Higgs, J., & Jones, M. (2000). Clinical reasoning in the health professions. In J. Higgs & M. Jones (Eds.), *Clinical reasoning in the health professions* (pp. 3-14). Woburn, MA: Butterworth-Heinemann.

Hogan, T. P. (2007). *Educational assessment: A practical introduction*. Hoboken, NJ: John Wiley & Sons, Inc.

Hur, Y., Kim, S., & Park, S. (2007). The correlation between CPX and written examination scores in medical students. *Korean Journal of Medical Education, 19*(4), 335-341.

Hutchins, E. (2010). Cognitive ecology. *Topics in Cognitive Science, 2*(4), 705-715.

Hwang, J., & Jeong, H. (2011). Relationship between the content of the medical knowledge written examination and clinical skill score in medical students. *Korean Journal of Medical Education, 23*(4), 305-314.

Jonassen, D. H. (1997). Instructional Design Models for Well-Structured and Ill-Structured Problem-Solving Learning Outcomes. *Educational Technology Research and Development, 45*(1), 65-94.

Jonassen, D. H. (2011). *Learning to solve problems. A handbook for designing problem-solving learning environments*. New York: Routledge.

Ju, H., & Choi, I. (2018). The role of argumentation in hypothetico-deductive reasoning during problem-based learning in medical education: A conceptual framework. *Interdisciplinary Journal of Problem-Based Learning, 12*(1). http://doi.org/10.7771/1541-5015.1638

Kaptelinin, V. (n.d.). Affordancse. In M. Soegaard & R. F. Dam (Eds.), The encyclopedia of human-computer interaction (2nd ed.) (Ch. 44). Retrieved from https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/affordances

Kassirer, J. P., & Gorry, G. A. (1978). Clinical problem solving: A behavioral analysis. *Annals of Internal Medicine, 89*, 245-255.

Kassirer, J. P. (1989). Diagnostic reasoning. *Annals of Internal Medicine, 110*, 893-900.

Khan, S. & VanWynsberghe, R. (2008). Cultivating the under-mined: Knowledge mobilization through cross-case analysis. *Forum: Qualitative Social Research, 9*, 1-21.

Kim, J., Lee, K., Choi, K., & Lee, D. (2004). Analysis of the evaluation for clinical performance examination using standardized patients in one medical school. *Korean Journal of Medical Education, 16*(1), 51-62.

Kim, M. (2009). Adoption of clinical skills examination: The challenge of National Health Personnel Licensing Examination Board. *Korean Journal of Medical Education, 21*(3), 215-216.

Kim, S., & Hur, Y. (2005). How are we going to evaluate medical professionalism? *Korean Journal of Medical Education, 17*(1), 1-14.

Kitchener, K. S. (1983). Cognition, metacognition, and epistemic cognition: A three-level model of cognitive processing. *Human Development, 26*, 222-232.

Knox, J. D. E. (1989). What is… a modified essay question? *Medical Teacher, 11*(1), 51-57.

Koh, S. & Park, H. (2009). The correlation between the scores of written examinations, the clinical clerkship examination, the clinical skill assessment, and the graduation examination of the medical students. *Korean Journal of Medical Education, 21*(4), 347-352. doi: 10.3946/kjme.2009.21.4.347

Kramer, A. W. M., Jansen, J. J. M., Zuithoff, P., Dusman, H., Tan, L. H. C., Grol, R. P. T. M., & van der Vleuten, C. P. M. (2002). Predictive validity of a written knowledge test of skills for an OSCE in postgraduate training for general practice. *Medical Education, 36*, 812-819.

Kwon, I., Kim, N., Lee, S., Eo E., Park, H., Lee, D., … Ryu, K. (2005). Comparison of the evaluation results of faculty with those of standardized patients in a clinical performance examination experience. *Korean Journal of Medical Education, 17*(2), 173-183.

Lee, K., & Choi, I. (2008). Learning classroom management through Web-based case-instruction: Implications for early childhood teacher education. *Early Childhood Education Journal, 35*, 495-503.

Ludmerer, K. M. (1985). *Learning to Heal: The development of American medical education*. New York, NY: Basic Books Inc.

Lyle, J. (2003). Stimulated recall: A report on its use in naturalistic research. *British Educational Research Journal, 29*(6), 861-878.

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design (1st ed.).* Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Martin, J., Lloyd, M., & Singh, S. (2002). Professional attitudes: Can they be taught and assessed in medical education? *Clinical Medicine, 2*(3), 217-223.

Maxwell, J. A. (2013). *Qualitative research design* (3rd ed.). Thousand Oaks, CA: SAGE Publications, Inc.

McGrenere, J., & Ho, W. (2000, May). *Affordances: Clarifying and evolving a concept*. Paper presented at the Graphics Interface 2000, Toronto, Canada. Abstract retrieved from https://www.cs.ubc.ca/~joanna/papers/GI2000_McGrenere_Affordances.pdf

Meng, K., Kang, B., & Lee, S. (1994). Analysis of Korean National Medical Licensing Examination question items of 1992-1993 on their levels of cognitive domain, types of multiple choice questions and the contents of medical knowledge tested. *Korean Journal of Medical Education, 5*(2), 11-17.

Merriam, S. B. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass Inc., Publishers.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.

Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*(9), S63-S67.

Newble, D. (2004). Techniques for measuring clinical competence: Objective Structured Clinical Examinations. *Medical Education, 38*, 199-203.

Norman, D. (1999). Affordances, conventions, and design. *Interactions*, 38-42. doi: 10.1145/301153.301168

Norman, D. (2013). *The design of everyday things*. New York, NY: Basic Books.

Norman, G. R., Trott, A. L., Brooks, L. R., & Smith, E. K. M. (1994). Cognitive differences in clinical reasoning related to postgraduate training. *Teaching and Learning in Medicine, 6*, 114-120.

Norman, G., Young, M., & Brooks, L. (2007). Non-analytical models of clinical reasoning: The role of experience. *Medical Education, 41*(12), 1140-1145.

Ogle, J., Bushnell, J., & Caputi, P. (2013). Empathy is related to clinical competence in medical care. *Medical Education, 47*, 824-831.

Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance and other formats*. Hingham, MA: Kluwer Academic Publishers.

Palmer, E. J., Duggan, P., Devitt, P. G., & Russell, R. (2010). The modified essay question: Its exit from the exit examination? *Medical Teacher, 32*(7), 300-307.

Park, H. (2004). Clinical application of objective structured clinical examination (OSCE). *Korean Journal of Medical Education, 16*(1), 13-23.

Park, H. (2008). Clinical skills assessment in Korean medical licensing examination. *Korean Journal of Medical Education, 20*(4), 309-312.

Park, K., Chung, W., Hong, D., Lee, W., & Shin, E. (2009). Relationship between the clinical performance examination and associated variables. *Korean Journal of Medical Education, 21*(3), 269-277.

Park, K., Park, H., Kim, J., & Hwang, H. (2015). Predictors of clinical practice examination scores of fourth year medical students after core clinical clerkships. *Korean Journal of Medical Education, 27*(4), 275-282.

Park, W., Lee, S., Kim, E., Kim, Y., Kim, S., Shin, J., & Lee. Y. (2005). Correlation of CPX scores with the scores of the clinical clerkship assessments and written examinations. *Korean Journal of Medical Education, 17*(3), 297-303.

Patel, V. L., Arocha, J. F., & Zhang, J. (2012). Medical reasoning and thinking. In K. J. Holyoak, R. G. Morrison (Eds)., *The oxford handbook of thinking and reasoning* (pp. 736-754). New York, NY: Oxford University Press.

Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science, 10*, 91-116.

Patel, V. L., Groen, G. J., & Norman, G. R. (1993). Reasoning and instruction in medical curricula. *Cognition and Instruction, 10*(4), 335-378.

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Payton, O. D. (1985). Clinical reasoning process in physical therapy. *Physical Therapy, 65*, 924-928.

Pirnay-Dummer, P., Ifenthaler, D., & Spector, J. M. (2010). Highly integrated model assessment technology and tools. *Educational Technology Research and Development, 58,* 3-18.

Reed, E. S. (1996). *Encountering the world: Toward an ecological psychology*. New York, NY: Oxford University Press.

Remmen, R., Scherpbier, A., Denekens, J., Derese, A., Hermann, I., Hoogenboom, R., … Bossaert, L. (2001). Correlation of a written test of skills and a performance based test: A study in two traditional medical schools. *Medical Teacher, 23*(2), 29-32. doi:10.1080/0142159002005541

Reznick, R. K., Blackmore, D., Cohen, R., Baumber, J., Rothman, A., Smee, S., … Berard, M. (1993). An Objective Structured Clinical Examination for the licentiate of The Medical Council of Canada: From research to reality. *Academic Medicine, 68*(10), S4-S6.

Schmidt, H. G., Norman, G. R., & Boshuizen, H. P. A. (1990). A cognitive perspective on medical expertise: Theory and implications. *Academic Medicine, 65*, 611-621.

Segers, M., Dochy, F., & Cascallar, E. (2003). *Optimising new modes of assessment: In search of qualities and standards.* Dordrecht, Netherlands: Kluwer Academic Publishers.

Shaw, R. E., Effken, J. A., Fajen, B. R., Garrett, S. T., & Morris, A. (1997). An ecological approach to the on-line assessment of problem-solving paths: Principles and applications. *Instructional Science, 25*(2), 151-166.

Shin, J., Lee, S., & Park, H. (2005). Standardized patients' accuracy in recording checklist items during clinical performance examinations. *Korean Journal of Medical Education, 17*(2), 197-203.

Shulman, J. H. (1992). Toward a pedagogy of cases. In J. H. Shulman (Ed.), *Case methods in teacher education* (pp. 1-30). New York, NY: Teachers College Press.

Sloan, D. A., Donnelly, M. B., Schwartz, R. W., & Strodel, W. E. (1995). The Objective Structured Clinical Examination: The new gold standard for evaluating postgraduate clinical performance. *Annals of Surgery, 222*(6), 735-742.

Small, P. A. Jr., Stevens, C. B., & Duerson, M. C. (1993). Issues in medical education: Basic problems and potential solutions. *Academic Medicine*, *68*(10), S89-S98.

Smee, S. (2003). Skill based assessment. *British Medical Journal, 326,* 703-706. doi: 10.1136/bmj.326.7391.703

Spector, J. M. (2006). A methodology for assessing learning in complex and ill-structured task domains. *Innovation in Education and Teaching International, 43*(2), 109-120.

Spiro, R. J., Collins, B. P., Ramachandran, A. R. (2006). Modes of openness and flexibility in cognitive flexibility hypertext learning environments. In B. Khan (Ed.), *Flexible Learning in an Information Society* (pp. 18-25). Hershey, PA: Idea Group.

Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: SAGE Publications, Inc.

Stake, R. E. (2006). *Multiple case study analysis*. New York, NY: Guilford Press.

Stillman, P. L., & Swanson, D. B. (1987). Ensuring the clinical competence of medical school graduates through standardized patients. *Archive Internal Medicine, 147*(6), 1049-1052.

Swing, S. R. (2007). The ACGME outcome project: Retrospective and prospective. *Medical Teacher, 29*, 648-654.

Tervo, R. C., Dimitrievich, E., Trujillo, A. L., Whittle, K., Redinius, P., & Wellman, L. (1997). The Objective Structured Clinical Examination (OSCE) in the clinical clerkship: An overview. *South Dakota Journal of Medicine, 50*, 153-156.

Thorndike, R. M., & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8^{th} ed.). Upper Saddle River, NJ: Pearson Education Inc.

Voss, J. F. (1988). Problem solving and reasoning in ill-structured domains. In C. Antaki (Ed.), *Analyzing everyday explanation: A casebook of methods* (pp. 74-93). Thousand Oaks, CA: SAGE Publications.

Vu, N. V., Marcy, M. M., Colliver, J. A., Verhulst, S. J., Travis, T. A., & Barrows, H. S. (1992). Standardized (simulated) patients' accuracy in recording clinical performance check-list items. Medical Education, 26(2), 99-104. doi: 10.1111/j.1365-2923.1992.tb00133.x

Warren, W. H., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *J Exp Psychol Hum Percept Perform, 13*(3), 371-383.

Williams, R. G., Barrows, H. S., Vu, N. V., Verhulst, S. J., Colliver, J. A., March, M., & Steward, D. (1987). Direct, standardized assessment of clinical competence. *Medical Education, 21*, 482-489.

Williams, S. M. (1992). Putting case-based instruction into context: Examples from legal and medical education. *The Journal of the Learning Sciences, 2*(4), 367-427.

Young, M. F., Kulikowich, J. M., & Barab, S. A. (1997). The unit of analysis for situated assessment. *Instructional Science, 25*(2), 133-150.

Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed.). Thousand Oaks, CA: Sage.

Yin, R. K. (2009). *Case study research: Design and methods* (4th ed.). Thousand Oaks, CA: SAGE Publications, Inc.

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: Sage Publications, Inc.

Zukow-Goldring, P., & Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention. *Neurocomputing, 70*, 2181-2193. doi: 10.1016/j.neucom.2006.02.029