

# Do medical students generate sound arguments during small group discussions in problem-based learning?: an analysis of preclinical medical students' argumentation according to a framework of hypothetico-deductive reasoning

Hyunjung Ju<sup>1</sup>, Ikseon Choi<sup>2</sup> and Bo Young Yoon<sup>3</sup>

<sup>1</sup>Innovation Center for Medical Education, Inje University College of Medicine, Busan, Korea, <sup>2</sup>Learning, Design, and Technology Program, The University of Georgia College of Education, Athens, GA, USA, and <sup>3</sup>Department of Internal Medicine, Inje University College of Medicine, Busan, Korea

**Purpose:** Hypothetico-deductive reasoning (HDR) is an essential learning activity and a learning outcome in problem-based learning (PBL). It is important for medical students to engage in the HDR process through argumentation during their small group discussions in PBL. This study aimed to analyze the quality of preclinical medical students' argumentation according to each phase of HDR in PBL.

**Methods:** Participants were 15 first-year preclinical students divided into two small groups. A set of three 2-hour discussion sessions from each of the two groups during a 1-week-long PBL unit on the cardiovascular system was audio-recorded. The arguments constructed by the students were analyzed using a coding scheme, which included four types of argumentation (Type 0: incomplete, Type 1: claim only, Type 2: claim with data, and Type 3: claim with data and warrant). The mean frequency of each type of argumentation according to each HDR phase across the two small groups was calculated.

**Results:** During small group discussions, Type 1 arguments were generated most often (frequency=120.5, 43%), whereas the least common were Type 3 arguments (frequency=24.5, 8.7%) among the four types of arguments.

**Conclusion:** The results of this study revealed that the students predominantly made claims without proper justifications; they often omitted data for supporting their claims or did not provide warrants to connect the claims and data. The findings suggest instructional interventions to enhance the quality of medical students' arguments in PBL, including promoting students' comprehension of the structure of argumentation for HDR processes and questioning.

**Key Words:** Argumentation, Hypothetico-deductive reasoning, Problem-based learning

## Introduction

Medical students need to develop clinical reasoning skills so that they can engage in accurate and timely problem solving and decision making when faced with

clinical problems in their future professional practice [1, 2]. Many medical schools have adopted a problem-based learning (PBL) method to enhance students' clinical reasoning skills, particularly hypothetico-deductive reasoning (HDR) skills, that helps students develop diagnostic reasoning abilities and deepen their know-

Received: October 19, 2016 • Revised: December 11, 2016 • Accepted: March 9, 2017

Corresponding Author: Hyunjung Ju (<http://orcid.org/0000-0001-6954-0310>)

Innovation Center for Medical Education, Inje University College of Medicine, 75 Bokji-ro, Busanjin-gu, Busan 47392, Korea

Tel: +82.51.890.6178 Fax: +82.51.890.6466 email: [hyunjung@inje.ac.kr](mailto:hyunjung@inje.ac.kr)

Korean J Med Educ 2017 Jun; 29(2): 101-109.

<https://doi.org/10.3946/kjme.2017.57>

eISSN: 2005-7288

© The Korean Society of Medical Education. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ledge in their early medical training [3,4,5].

In PBL, medical students in a small group are encouraged to engage in HDR processes by constructing and exchanging their ideas and providing reasoned arguments [6]. Such students' argumentation during PBL guides them in applying basic scientific knowledge to clinical contexts, taking a coherent approach to diagnostic inquiry, and constructing collaborative knowledge that involves causal explanations of a patient problem [6,7], which can promote the HDR process.

If HDR is one of the essential learning activities as well as learning outcomes of PBL, it is critical to find ways to diagnose and enhance students' argumentation exchanged during their small group discussions. Although there are several studies that have examined the types of verbal interactions (e.g., statements and questions) among medical students or between students and tutors during PBL sessions [8], there are very few studies that have focused on medical students' argumentation with regard to the HDR process during PBL.

Many medical schools in South Korea have implemented PBL as part of their preclinical curricula since the late 1990s [9]. However, preclinical students have experienced challenges with meaningful discussions during PBL [10]. For example, some students were unwilling to exchange their ideas due to anxiety about their lack of prior knowledge and tended to leap to a certain diagnosis of an illness, overlooking basic mechanisms underlying a patient's problem [10]. These students' tendencies might influence the quality of their argumentation according to HDR processes in PBL. In order to enhance preclinical students' argumentation in PBL, empirical research on how they actually engage in argumentation during PBL through analyses of their discussions is needed.

Thus, the purpose of this study was to analyze preclinical medical students' argumentation during their

small group discussions in PBL. In particular, this study focuses on understanding the quality and quantity of arguments generated in each phase of HDR. It is expected that the findings of this study will provide insights into developing effective strategies to foster medical students' argumentation in PBL, which would improve their HDR.

---

## Subjects and methods

### 1. Research site

This study was conducted at Inje University College of Medicine (IUCM) in the fall semester of 2014. In IUCM's 2-year preclinical curricula, 14 organ system block courses were implemented; each included a 1-week-long PBL course followed by 3- to 4-week-long lectures. Thus, preclinical students in the school did not have prior knowledge related to the given topic before each PBL course. For each PBL course, a small group of seven to eight students had 2-hour discussion sessions guided by a tutor three times a week (Monday, Wednesday, and Friday) to work on a clinical case, engaging in HDR processes. During the first session, a standardized patient was used.

### 2. Participants

We recruited participants for this study from the cohort of the first-year preclinical students (a total of 100 students) and collected their small group discussions during the sixth of seven PBL courses in the first-year preclinical curriculum. Among seven PBL courses, three were implemented in their first semester and the other four in their second semester. In the process of selecting one of the seven PBL courses for this study, the first and the last courses in each semester were eliminated to

avoid any potential influence of the students' lack of familiarity with the PBL course and time pressure due to final exams respectively. The second PBL course in the first semester was also eliminated, because we thought having one PBL course might not be enough for the students to become familiar with the PBL process, and their lack of understanding of the PBL process might influence their discussion activities. Thus, we concluded the second or third PBL courses in the second semester would be the most appropriate point to collect the data with an appropriate level of the students' experience with PBL. Communicating with the PBL administrators, we were able to gather the data from the third PBL course in the second semester, which was the sixth PBL course during the first preclinical year. The topic for the chosen PBL course was concerned with the cardiovascular system according to the curriculum, and a clinical case used was an acute myocardial infarction with a patient complaining of chest pain and dyspnea. Although a topic itself may influence the dynamics of students' discussions in general, we believe that the participants' experience with PBL and prior knowledge related to any given discussion topics would be more critical factors that need to be considered. In this regard, it is important to note that the students' prior knowledge level would share the same condition across all of the different PBL courses, since each PBL course preceded any lectures indicated above.

For the PBL course, 100 first-year preclinical students were divided into 15 small groups, and 15 tutors were randomly assigned to each group. To minimize the influence of the tutors' lack of facilitation skills on the students' discussions while representing the majority of the tutors' experience levels, we identified two out of the 15 tutors in the range between 3 and 5 years of experience with PBL tutoring. In the PBL course, the average of the tutors' experience was 4 years (standard

deviation=3.3). Thus, out of the 15 student small groups, two groups who would be facilitated by the two tutors with 3 and 5 years respectively were selected for this study. Then, a total of fifteen students in the two groups selected—one group included eight students: five males (62.5%) and three females (37.5%) and the other group contained seven students: five males (71.4%) and two females (28.6%)—were asked to participate in this study, and all submitted their informed consent.

### 3. Data collection

A set of three PBL discussion sessions guided by a tutor for each of the two small groups was audio-recorded. The data consisted of about 12 hours of audio-recordings for the two groups' PBL discussion sessions; each session lasted for about 2 hours. All of the audio-recorded discussions were transcribed verbatim.

### 4. Data analysis

#### 1) An analytical framework for examining argumentation in HDR

Toulmin [11]'s model is regarded as the most seminal argumentation model [12,13]. Toulmin [11]'s model emphasizes three primary components of an argument: (1) a claim; (2) data for supporting the claim; and (3) a warrant for justifying the transition from the claim and data. Additionally, Toulmin [11] suggested that more complex arguments can include three additional components: (1) a backing for providing a rationale for the warrant; (2) a rebuttal for presenting an exception to the claim; and (3) a qualifier for expressing limited certainty of the claim (e.g., "perhaps" and "probably"). Since constructing sound arguments basically depends on the ability to provide data and warrants for one's claims [11], this study focused on the three primary components of argumentation—a claim, data, and a warrant.

In this study, the authors developed an analytical

framework for analyzing students' argumentation according to each phase of HDR [14]. As shown in Table 1, this framework integrates the structure of argumentation, including the three key elements of an argument (a claim, data, and a warrant) based on Toulmin [11]'s model, into the six phases of HDR adapted from Barrows [3]' clinical reasoning model—problem framing, hypothesis generation, inquiry strategy, data analysis/synthesis, diagnostic decision, and therapeutic decision [14].

## 2) Unit of analysis

The unit of analysis was one argument verbally expressed by one person, including at least one of the three primary components of an argument. One argument dealt with one topic that was related to clinical reasoning for the patient's problem and was in a statement form, not a question form.

## 3) Analysis

Before analyzing each argument as a basic unit of analysis, meaningful segments of successive arguments exchanged among the participants of each small group were identified, and each segment was categorized into one of the six phases of HDR. Each argument was analyzed according to the three essential elements of argumentation (a claim, data, and a warrant) based on the analytical framework (Table 1). Then, each argument was categorized by a coding scheme of the following four types of arguments (Table 2).

The first author (an education expert) reviewed the entire transcripts of the two groups' discussions. As a result, a total of 55 discussion segments, holding a total of 561 arguments, were identified. To obtain inter-rater reliability of the first author's coding, a total of 24

Table 1. Analytical Framework for Argumentation in the HDR Process

HDR phase	Structure of argumentation		
	Data	Warrant	Claim
Problem framing	Identified information or cues	Explanation why the identified information or cues are important	Initial concept of the patient's problem from the identified information or cues considered important
Hypothesis generation	Identified information or cues recognized as important data	Pathophysiological mechanisms involved in the patient's problem	Basic mechanisms (anatomy, biochemistry, physiology, etc.) or disease entities that could be responsible for the patient's problem
Inquiry strategy	Patient's information or cues organized by generated hypotheses	Basic mechanisms underlying hypotheses entertained; information that the inquiry actions will produce	Actions or decisions on what information would be necessary
Data analysis/synthesis	Data acquired from inquiry strategies	Basic mechanisms at the appropriate level	Interpretation on significant patient data that relates to the hypotheses considered
Diagnostic decision	Rearranged significant patient data and its interpretations	Underlying responsible mechanisms involved in the patient's problem; diagnostic criteria for the most likely disease	Decision on the most likely hypothesis(es) responsible for the patient's problem
Therapeutic decision	Diagnostic decision(s) with relevant patient's data	Basic mechanisms relating to the therapeutic interventions; research into the therapeutic efficacy of the chosen treatments	Decision on the approach to the treatment of the patient's problem

Adapted from Ju and Choi. *Interdiscip J Probl Based Learn*. Forthcoming, with permission of Purdue University Press [14]. HDR: Hypothetico-deductive reasoning.

Table 2. Coding Scheme

Type of argument	Code	Description	Example
0	Incomplete	Data or warrant(s) without claim	The patient has high levels of blood urea nitrogen (data).
1	Claim only	Claim without data and warrant(s)	I think angina is likely (claim).
2	Claim-data	Claim with data but no warrant(s)	Because the patient has low blood pressure (data), he may have bleeding (claim).
3	Claim-data-warrant	Claim with data and warrant(s)	It is necessary to provide beta blockers or nitrates (claim) for the patient with acute myocardial infarction (data). The medications can lower his blood pressure. A decrease in blood pressure leads to a decrease in oxygen demand, which in turn relieves ischemic chest pain (warrant).

segments (44% of the total segments) were randomly selected, which included 195 arguments (35% of the total arguments), and these were coded by the third author (an internal medicine physician and educator) independently. Cohen [15]'s  $\kappa$  was calculated for the coding of the argument type, and a high level of inter-rater reliability ( $\kappa=0.88$ ) was obtained. Regarding the classification of the arguments according to HDR phases, 100% of agreement between the two reviewers was obtained.

To analyze students' arguments generated by each group during their group discussions, a frequency analysis was conducted according to the four types of argumentation and the six phases of HDR. The mean frequency of each type of argument constructed across the two groups was used for results.

## Results

### 1. Frequency of the different types of student arguments during the overall HDR process

A total of 280.5 arguments per group were constructed during the overall HDR process. Of the four types of arguments, as shown in Table 3, Type 1 arguments were generated most often (frequency=120.5, 43%), whereas

the least common were Type 3 arguments (frequency=24.5, 8.7%). Type 2 arguments had the second highest frequency (68.5, 24.4%).

### 2. The frequency of each type of argumentation per group during each HDR phase

Results of the frequency analysis according to each HDR phase are as follows (Table 3).

#### 1) Problem framing

This phase included 4.5 arguments per group (1.6% of the total arguments constructed during all HDR phases). We observed that the students obtained very limited information from the standardized patient during the initial interview. The most frequent type of argument was Type 1 arguments (frequency=2.5, 55.6%), but there were no Type 3 arguments.

#### 2) Hypothesis generation

A total of 75 arguments were constructed during this phase. We found that most students brainstormed disease names rather than basic mechanisms without explicitly providing data or warrants. Over half of the arguments generated were taken up by Type 1 arguments (frequency=43.5). On the other hand, Type 3 arguments accounted for 4.7% of the total arguments in this phase (frequency=3.5), which was the least common.

Table 3. Mean Frequency of Arguments during Each Phase of Hypothetico-Deductive Reasoning

Type of argument	Hypothetico-deductive reasoning phase						Total
	Problem framing	Hypothesis generation	Inquiry strategy	Data analysis/synthesis	Diagnostic decision	Therapeutic decision	
Type 0: incomplete	1.5 (33.3)	5.5 (7.3)	0	53.0 (71.1)	0	7.0 (9.7)	67.0 (23.9)
Type 1: claim only	2.5 (55.6)	43.5 (58.0)	45.5 (87.5)	5.5 (7.4)	0	23.5 (32.6)	120.5 (43.0)
Type 2: claim-data	0.5 (11.1)	22.5 (30.0)	5.5 (10.6)	9.0 (12.1)	1.0 (40.0)	30.0 (41.7)	68.5 (24.4)
Type 3: claim-data-warrant	0	3.5 (4.7)	1.0 (1.9)	7.0 (9.4)	1.5 (60.0)	11.5 (16.0)	24.5 (8.7)
Total (%) <sup>a)</sup>	4.5 (1.6)	75.0 (26.7)	52.0 (18.5)	74.5 (26.6)	2.5 (0.9)	72.0 (25.7)	280.5 (100)

Data are presented as mean frequency (%).

<sup>a)</sup>% of the total during the overall hypothetico-deductive reasoning phases.

### 3) Inquiry strategy

The total number of arguments raised during this phase was 52. Most students tended to generate routine action items (e.g., following a list of basic tests, such as blood tests and X-rays) without providing specific reasons about why the particular inquiry strategies would be necessary for the patient. Type 1 arguments were generated most often (frequency=45.5, 87.5%), whereas the frequency of Type 3 arguments was the lowest (frequency=1, 1.9%).

### 4) Data analysis/synthesis

During this phase, a total of 74.5 arguments were offered. Most students simply repeated the factual information obtained from inquiry strategies (e.g., results of lab tests) without interpreting the patient's data. Of the total, the most and least frequent types of arguments were Type 0 (frequency=53, 71.1%) and Type 1 arguments (frequency=5.5, 7.4%), respectively. We found only one or two students per group were strongly and recurrently predispose to offer a claim with justifications, which shows they might have learned how to interpret expected test results for the hypotheses through their self-directed study before the second session including this phase.

### 5) Diagnostic decision

The total number of arguments produced during this

phase was 2.5. We found that when one or two students in a group initiated arguments about a patient's diagnosis, other group members tended to agree with the ideas. Among the four types of arguments, only Type 2 and Type 3 arguments were found. This indicates the students advanced their arguments while practicing their analyses of the patient's data obtained from their inquiry and integrating related basic mechanisms toward the validation of the hypothesis entertained.

### 6) Therapeutic decision

A total of 72 arguments were constructed during this phase. We found that the students tended to have difficulty using biomedical knowledge to explain how treatment strategies considered important can improve the patient's condition. The frequency of Type 2 arguments (30, 41.7%) was the largest.

## Discussion

It is important for medical students to construct reasoned arguments during PBL so that they can build scientific and causal explanations of a patient's problem and carry out quality scientific inquiry [3,16], which in turn can enhance their HDR abilities and experiences with PBL. However, this study revealed that the students

predominantly generated claims without proper justifications during their group discussions in PBL; they often omitted data and warrants for supporting their claims. The findings in this study suggest the need for guidance in promoting medical students to build sound arguments during PBL.

## 1. Overall patterns of arguments constructed by students during each HDR phase

In this study, the largest frequency of Type 1 arguments (claim only) among four types of arguments found during most of HDR phases can be a furtherance of those studies reporting that claims were generated more frequently than evidence or warrants by high school students during scientific inquiry processes [12,17]. The students' tendency might be attributed to their lack of consideration for explicitly presenting data and warrants to support their claims [18] or their lack of knowledge about the given topic [17,18].

More importantly, in PBL, students are expected to acquire basic science knowledge (e.g., physiology, biochemistry, and pathophysiology) in the context of clinical problems, focusing on the underlying responsible mechanisms [3,16]. It is essential for students with little or no clinical experience to activate their basic science knowledge in order to understand a patient's problem [3,4,16]. However, during most of the HDR phases, the smallest frequency of Type 3 arguments (claim with data and warrant) might be interpreted as indicative of the students' deficiency in integrating their knowledge of basic mechanisms with a patient's clinical features (e.g., symptoms, signs, or clinical findings) [10]. This might have made it difficult for the students to analyze, evaluate, and manage a patient's problem mechanistically. It would be necessary for the tutors to prompt students in using and retrieving the basic science knowledge involved in the patient's case according to

HDR processes during PBL by employing appropriate tutoring strategies, such as questioning [14,16] (see the suggestion section for more detail).

Interestingly, most of Type 0 (incomplete) arguments constructed in all HDR phases were attributed to the data analysis and synthesis phase. The largest frequency of Type 0 arguments observed during this phase can be explained by the students' lack of knowledge and skills in analyzing the patient data and determining how test results were related to a more likely hypothesis [10]. Thus, to support the students' analysis of significant patient data, it may be helpful to provide them with related learning resources, such as references and reading materials.

## 2. Suggestions for enhancing medical students' argumentation in PBL

The following discusses two instructional interventions to improve the quality of medical students' argumentation during their small group discussions in PBL based on previous studies.

### 1) Understanding the structure of argumentation

One strategy for enhancing medical students' argumentation can be helping them comprehend the structure of argumentation for HDR processes [14]. Cerbin [18] argued that students' naïve conceptions of argument structures can cause weak argument construction. Medical students should be provided with instruction in understanding the nature of argumentation and the structure of argumentation contextualized in each HDR phase, including the primary components of argumentation (presented in Table 1), and the practice of constructing arguments [14]. Moreover, in PBL, tutors are supposed to facilitate students' discussions rather than directly providing them with knowledge or information [3,6,16,19]. Thus, tutors should grasp concepts of how to structure valid arguments according to HDR processes in

order to guide students' meaningful argumentation [14]. It would also be necessary that tutors have opportunities to identify the structure of argumentation for HDR processes through a training session [14].

## 2) Questioning

Questioning can also be an effective intervention for promoting students' argumentation [6,14]. Several studies found that questioning helps students produce well-reasoned arguments during problem-solving processes [13,20]. In PBL, tutors should be able to ask appropriate questions so that students can engage in high-quality argumentation, including data and warrants for supporting claims [14]. Examples of questions, such as question prompts, based on the structure of argumentation for each HDR phase should be provided for tutors to support students' argumentation according to HDR phases [14]. For example, during the hypothesis generation phase, when a student offered a claim and data without any warrant, a tutor can ask a question to encourage him to articulate a warrant, such as "What pathophysiological mechanisms might be concerned with the patient's problem?" [14,16]. In addition, it can be helpful for students to use questions for enhancing their argumentation [20]. During PBL, students' application of similar example questions provided for tutors can also be used as a self-guide for formulating coherent arguments and stimulating their peers to present quality arguments.

## 2. Limitations and future research

There may be many factors influencing the dynamics of small group discussions in PBL, such as students' PBL experience level and prior knowledge related to the given topic, and a tutor's experience level. Despite our efforts to minimize those influences, the topic itself, and the time points in which students engage in certain PBL courses, and the tutor's facilitation skills, including their statements or questions used during PBL, might result in

a different quality of the students' argumentation. Thus, to understand a broader spectrum of the students' discussions and argumentation activities during PBL, further studies should include other discussion topics using PBL across different time points (e.g., second preclinical years) and consider the tutors' style of tutoring.

According to our data coding, about one-fourth of the total arguments constructed were classified as incomplete arguments, which may mainly be due to our data analysis protocol focusing only on the statements verbally expressed by students. Some of the incomplete arguments might have implicit claims, but it was not easy to determine argumentation components unless explicitly stated. In addition, given the PBL situation in which students exchange their ideas with one another, all of their actual thinking processes may not be expressed verbally. So, there is a gap between students' stated arguments and their implicit reasoning, which will be a limitation of our current analysis. Thus, further studies may consider integrating advanced analysis methods that could capture a more in-depth flow of group discussion. Furthermore, future research needs to focus on finding ways to enhance the quality of students' argumentation in PBL based on the suggested instructional interventions.

### ORCID:

Hyunjung Ju: <http://orcid.org/0000-0001-6954-0310>;

Ikseon Choi: <http://orcid.org/0000-0001-8005-1649>;

Bo Young Yoon: <http://orcid.org/0000-0002-9635-4799>

**Acknowledgements:** This manuscript comprised part of a manuscript style dissertation the first author submitted to the University of Georgia. The authors would like to thank Dr. Byoung Doo Rhee, Vice-President of Medical and Pharmaceutical Affairs, and Dr. Jong-Tae Lee, Dean of College of Medicine, at Inje University, for their support



on this research.

**Funding:** None.

**Conflicts of interest:** None.

## References

1. Groves M. Understanding clinical reasoning: the next step in working out how it really works. *Med Educ* 2012; 46: 444-446.
2. Kempainen RR, Migeon MB, Wolf FM. Understanding our mistakes: a primer on errors in clinical reasoning. *Med Teach* 2003; 25: 177-181.
3. Barrows HS. *Practice-based learning: problem-based learning applied to medical education*. Springfield, USA: Southern Illinois University; 1994.
4. Patel VL, Arocha JF, Zhang J. Thinking and reasoning in medicine. In: Holyoak KJ, Morrison RG, eds. *The Cambridge handbook of thinking and reasoning*. New York, USA: Cambridge University Press; 2005. p 727-750.
5. Durham CO, Fowler T, Kennedy S. Teaching dual-process diagnostic reasoning to doctor of nursing practice students: problem-based learning and the illness script. *J Nurs Educ* 2014; 53: 646-650.
6. Hmelo-Silver CE, Barrows HS. Facilitating collaborative knowledge building. *Cogn Instr* 2008; 26: 48-94.
7. Frederiksen CH. Learning to reason through discourse in a problem-based learning group. *Discourse Process* 1999; 27: 135-160.
8. Visschers-Pleijers AJ, Dolmans DH, de Leng BA, Wolfhagen IH, van der Vleuten CP. Analysis of verbal interactions in tutorial groups: a process study. *Med Educ* 2006; 40: 129-137.
9. Yeo S, Chang BH. Students' perceptions and satisfaction level of hybrid problem-based learning for 16 years in Kyungpook National University School of Medicine, Korea. *Korean J Med Educ* 2016; 28: 9-16.
10. Ju H, Choi I, Rhee BD, Lee JT. Challenges experienced by Korean medical students and tutors during problem-based learning: a cultural perspective. *Interdiscip J Probl Based Learn* 2016; 10: Article 8.
11. Toulmin S. *The uses of argument*. Cambridge, UK: Cambridge University Press; 1958.
12. Jimenez-Alexandre MP, Rodriguez AB, Duschl RA. "Doing the lesson" or "doing science": argument in high school genetics. *Sci Educ* 2000; 84: 757-792.
13. Jonassen DH. *Learning to solve problems: a handbook for designing problem-solving learning environments*. New York, USA: Routledge; 2011.
14. Ju H, Choi I. The role of argumentation in hypothetico-deductive reasoning during problem-based learning in medical education: a conceptual framework. *Interdiscip J Probl Based Learn*. Forthcoming.
15. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas* 1960; 20: 37-46.
16. Barrows HS. *How to design a problem-based curriculum for the preclinical years*. New York, USA: Springer; 1985.
17. McNeill KL, Lizotte DJ, Krajcik J, Marx RW. Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *J Learn Sci* 2006; 15: 153-191.
18. Cerbin, B. The nature and development of informal reasoning skills in college students. Paper presented at: the 12th National Institute on Issues in Teaching and Learning; 1988 April 24-27; Chicago, USA.
19. Schmidt HG, Rotgans JI, Yew EH. The process of problem-based learning: what works and why. *Med Educ* 2011; 45: 792-806.
20. Chin C, Osborne J. Supporting argumentation through students' questions: case studies in science classrooms. *J Learn Sci* 2010; 19: 230-284.