Alternatively Assessing Conceptual Learning in an Emergency Clinical Environment—A Mixed Methods Design

Zhidong Zhang[1] & Alice Yang Zhang[2]

[1] The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA

[2] Department of Ophthalmology, The University of North Carolina at Chapel-Hill, North Carolina, USA

Correspondence: Zhidong Zhang, The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA.

Abstract

Modeling cognitive processes in clinical learning environments is a necessary first step towards improving learning assessment and medical practice by using an alternative assessment model. Verbal protocol and cognitive content analyses are effective methods of exploring such cognitive processes. For the purpose of simplifying the discussion, we have labeled these processes as Identification of Information, Advanced Cognition, and Medical Cognitive Action. Exploring problem solving processes with Bayesian network techniques can characterize students' dynamic learning processes quantitatively, identify differences in cognitive components at different stages of learning and better represent clinical problem solving features.

We develop a hierarchical cognitive model as a cognitive assessment tool to describe the complex cognitive network relations, which can be applied to various clinical cognitive situations. The study concludes that the cognitive model was useful in identifying students' learning trajectories by representing the different cognitive features.

Keywords: cognitive process, clinical learning environment, Bayesian network model

1. Introduction

An examination of medical student clinical learning through verbal protocol analysis can contribute to our understanding of the diverse aspects of problem solving (Lajoie, Greer, Munsie, Wilkie, Guerrera, & Aleong, 1995; Lu, 2007; Rose, Bearman, Naweed, & Dorrian, 2019). This study seeks to enhance our understanding of how medical students acquire competence in clinical problem solving by modeling their problem solving practices and learning trajectories. We are most interested in how medical students come to identify important information, what is an indicator of advanced cognition and what actions reflect medical cognitive actions in clinical problem solving. Three research questions are addressed: (1) How do medical students identify relevant information, evaluate evidence, and take action in the learning process in a clinical environment? (2) How can we establish a moment-by-moment model that can dynamically represent students' cognitive trajectories? (3) What are the features of student problem solving trajectories? Simulated clinical problem solving data will be used to answer these questions.

2. Theoretical Background

Lajoie (2003) proposed that identifying expert's knowledge, skills and competence can help construct the trajectory toward competence for specific cognitive tasks. Learning trajectories have several signposts where learning transitions can take place. Implicitly, these signposts can be represented as critical cognitive components applied in describing problem solving processes (Zhang & Frederiksen, 2007; Zhang & Lu, 2014a). The learning trajectory concept reflects the dynamic nature of learning and that there are transitional processes that lead to expertise. Zhang and Frederiksen (2007) have explored the learning trajectory concept for statistics problems that require the use of ANOVA. They developed a Bayesian network model based on the cognitive model of the domain. Other similar studies can be seen in mathematical learning assessment (Zhang, 2016; 2018) and argumentative learning (Zhang & Lu, 2014b). The procedures are applied to this investigation where we analyze clinical problem solving through a verbal protocol analysis.

3. Cognitive Skills and Sequential Levels

For some problem solving tasks, the sequence of performing certain cognitive skills is important and consequently must be identified in order to promote learning. With respect to reasoning, Henri (1992) identified five sequential skills: elementary classification, in-depth classification, inference, judgment and application of strategies. Complex problem solving has been described in terms of the plans, goals, operators, methods and strategies needed to solve problems (Hayes, 1989; Newell & Simon, 1972). Lu (2007) has explored clinical decision-making and communication in the context of groups participating in a simulated medical emergency activity. Analyses of the decision-making activities (Werner, 1995) revealed a series of sequential components: planning, data collecting, managing, and interpreting patient data. The analysis from the communicative activities revealed a sequential pattern consisting of informative, argumentative, elicitative, responsive, and directive acts. Evidence from studies of reasoning, decision making and communication can inform future research in the area of cognition.

4. Cognitive Hierarchical Model

Sequential cognitive skills can be used in representing differential learning trajectories. However, as these cognitive skills cannot be directly observed, their structure can be presented in a hierarchical model (Zhang & Frederiksen, 2007; Zhang & Lu, 2014b), in which explanatory variables obtain support from evidence variables (Figure 1). In other words, a set of explanatory variables representing a cognitive structure obtains information support from another set of explanatory or evidence variables. Horizontal nodes represent sequence and vertical nodes represent the flow of information from evidence and explanatory sources. Fig. 1 represents the relations of sequential and hierarchical features. Explanatory variable BI, B2 and B3 are sequential. The hierarchical relation has been represented, for example, from Explanatory A, Explanatory B2 to Evidence C3.
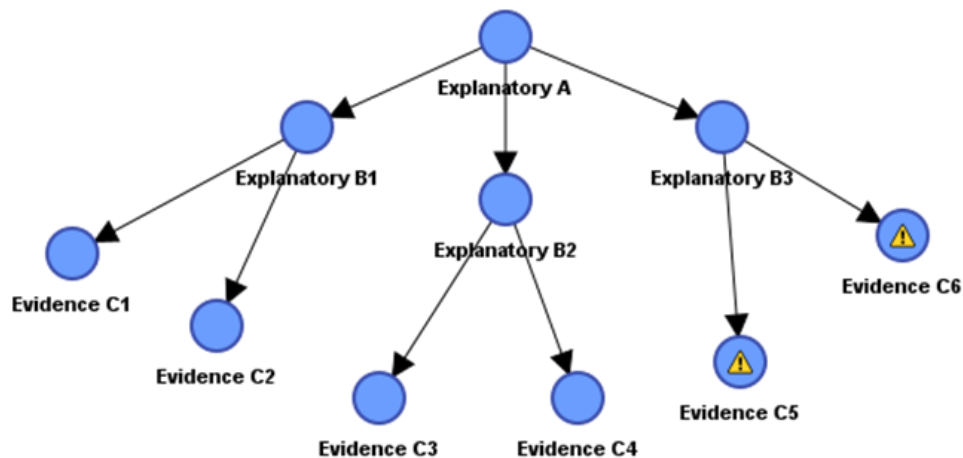

Figure 1. Concept relations of sequential and hierarchical features

5. Research Methods

5.1 Research Designs and Paradigms

This study used mixed research design, of which there are two steps. The first step is qualitative data collection, in which cognitive task analysis method was used. The second step is to quantify the collected data in the cognitive task analysis (Annett, 2000; Chipman, Schraagen, & Shalin, 2000; Seemster, Redding & Kaempf, 2000). A Bayesian network (Conrady & Jouffe, 2015) was used to quantitatively represent a cognitive model (Almond, Mislevy, Steinberg, Yan, & Williamson, 2015), which recorded both individual student and a group of students' learning trajectories. In the mixed methods research, this design was called exploratory sequential design (Creamer, 2018; Creswell & Plano Cark, 2018; DeCuir-Gunby & Schutz, 2017), which effectively represented qualitative data collected by using cognitive task analysis or other alternative qualitative coding techniques (Miles, Huberman & Saldana, 2014).

The paradigm of the study took dialectical pluralism (Creswell & Plano Clark, 2018; Goertzen, 2010; Greene & Hall, 2010). The study in exploratory design starts with the collection and analysis of qualitative data. The research study initially emphasizes the research problem which was supported by the findings of qualitative data

analysis. Thus, the researcher attempted to obtain deep understanding in constructivist perspectives. When researchers moved to the quantitative phase, the post-positivism will be used in explaining the analysis results and findings. Consequently, dialectical pluralism as a paradigmatic framework is appropriate in the study to assemble both data from the cognitive task analysis and the one represented in a Bayesian network model.

5.2 Scenario and Data

The scenario was from the study of medical problem solving (Lu, 2007) and medical student learning (Zhang & Lu, 2014b), in which a clinical case was presented to medical learners. The data were a simulated data set. A 37-year-old man with a history of alcohol abuse presents to the emergency room with sudden onset diffuse abdominal pain. Eight third-year medical students participate in the study to discuss how to approach diagnosing the case. The clinical teaching session was based on a simulated scenario in the emergency room.

5.3 Cognitive Tasks and Analysis

Cognitive tasks are critical targets from which evidence variables can be extracted. The design is based on cognitive theory and research findings (Lesgold, Lajoie, Logan, & Eggan, 1990; Hollnagel, 2003; Riera & Debernard, 2003; Elm, Potter, Gualtieri, Easters, & Roth, 2003). In this study, the statement of participants problem solving strategies were qualitative data resources. Idea unit analysis (Frederiksen, 1986; Ayala, Yin, Shavelson, & Vanides, 2002) is an analytical technique for extracting variables from cognitive tasks. Idea units have been decided based on theoretical constructs which can function as both criteria for extracting evidence variable values and model components for presenting different aspects of competency (Elm, Potter, Gualtieri, Easters, & Roth, 2003; Ayala, Yi, Shavelson, & Vanides, 2002) to identify their clinical reasoning processes. Idea units connect both evidential data and theoretical constructs.

6. Bayesian Network Theory and Model

6.1 Bayesian Networks

Bayesian networks (BNs) are also known as belief networks, which are usually represented in a directed acyclic graph (DAG) to model a learning process or a problem solving process and procedure (Culbertson, 2016; Mislevy, 1994; Pearl, 1988; Tan & Tay, 2010). Koller and Friedman (2009) state that, assuming there is a class of variables which can be designated by $x_1, x_2, \ldots x_n$ and C. The structure can be seen in figure 2. In such an example variable $x_1, x_2, \ldots x_n$ are observed. The upper level variable C means a class. It represents a concept, which is supported by all of these observed/evidence variable x_s. All of these variable x_s are conditionally independent on the variable *Class*. Thus, the relationship of all of these variables can be described in a model which factorizes as:

$$P(c, x_1, x_2 \cdots xn) = P(c)\Pi_{i=1}^{n} P(X_i|c)$$

This model can represent the joint distribution by using a group of variables, which include a prior distribution $P(C)$ and a set of $P(X_i|C)$. The model indicates that a concept, which is usually a latent variable, can be supported by a group observed variable x_s.
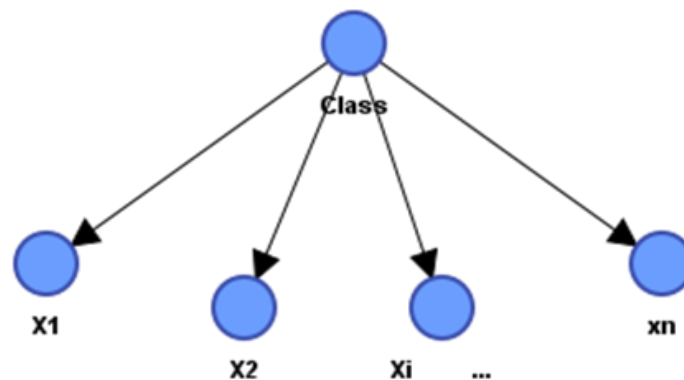

Figure 2. A Bayesian network graph

6.2 Bayesian Networks and Cognitive Features

In this study, a hierarchical Bayesian network (BN) of cognitive feature trajectories has been developed based on the analysis of clinical problem solving process in the aspect of clinical information collection and application. A BN can be used to model complicated cognitive tasks where researchers are more concerned with cognitive processes. In its graphic model, it can be described as a directed acyclic graph (DAG), and each node represents either explanatory or evidence variables and each directed edge connects two variables (Aryadoust, 2011; de la Torre & Minchen, 2014; Zhang & Frederiksen, 2007). Although explanatory variables cannot be directly observed, they can be informed by evidence variables which are obtained from the data. A BN model is instantiated by updating evidence variables.

7. Development of Cognitive Variables and Model

A cognitive BN is composed of a configuration of evidence and explanatory variables. Evidence variables collect evidence from cognitive tasks (Zhang, 2007; Zhang & Lu, 2014a). Explanatory variables are theoretical constructs. In this study, evidence variables represent observable judgments from simulated students' learning protocols, which are the records of students' problem solving process based the clinical case.

7.1 Definition of Evidence Variables

The cognitive model is organized hierarchically and the evidence variables at the bottom layer of the model are used to extract variable values from the data by the idea unit analysis (Zhang, 2007; Zhang & Lu, 2014b). There are ten evidence variables.

1) Identifying critical elements among these clinical Information: Students list related information which they believe are useful in the diagnosis and patient management process. The symptoms and clinical test results may provide important information about the patient. The students should carefully and systematically recognize the useful information. In addition, time is a sensitive factor in life and death situations ("I would check the vital signs". These students should be able to select and identify critical elements of the information in the time sensitive clinical environments.

2) Situation judgment: This variable describes students facing a situation where patients may have more than one problem or a set of syndromes "He recently had a sports-related shoulder injury and now has new abdominal pain". The students learn to organize critical symptoms and find relationships among these symptoms. In this variable, an inferential and logic chain may develop.

3) Medical decision-making: Students express their opinions and make a diagnosis. E.g., "I would think he would have tachycardia."

4) Evaluation: Students give a rationale for a judgment, opinion, or action. E.g., "I would ask for an abdominal X-ray to look for free air under the diaphragm, to see if he has a bowel perforation."

5) Medical metacognition: It is an advanced cognitive process which assesses and analyzes evaluations and judgments. For example: "I'm worried about the blood pressure, because I think he'd be tachycardic, and have a fever. But I do not think he'd have high blood pressure. "

6) Patient history: Students intend to collect more information about patient history. The purpose is to find evidence to support current problem solving and reasoning. E.g., "I would like to know his meds and what he took for the shoulder pain."

7) Physical examination: Students intend to collect more information through physical examination. E.g., "I would ask for blood pressure on both arms and I would do an abdominal exam looking for signs of peritonitis."

8) Doing laboratory test: Students intend to collect more information through laboratory test, such as "white blood cell count".

9) Giving first-order medication: This is a cognitive action that the medical student would take after a judgment. " … I would start him on intravenous broad-spectrum antibiotics"

10) Monitoring patient: This is a more detailed action that might include observations and control of patient situation. For example: "give oxygen..."

7.2 Development of Explanatory Variables

Explanatory variables consist of two layers in a hierarchical cognitive model (Zhang, 2007). It is more convenient to describe the explanatory variable from the very bottom level which is close to the evidence variables. As the model is completely established, the information of the explanatory variables will be transferred

from the top explanatory variable to the evidence variables.

7.3 The Lower and Upper Levels of Explanatory Variables

As soon as those evidence variables are defined, the lower explanatory variables have to be developed. The lower level means that the explanatory variable is closely connected to the evidence variable; but they are not observed from the data set or they have to be updated by evidence variables. Ten evidence variables have been grouped into four categories.

Recognizing Critical Element and Situation Judgment share a common feature: Identification of Information. It is obvious that the latter concept indicates the first step in medical problem solving. Medical Decision-Making, Evaluation and Medical Metacognition consist of Advanced Cognition which reflects the second step following Identification of Information. Patient History, Physical Examination, and Doing Laboratory Tests are composed of Collecting Patient Data in the past, present, and future time points. Giving First-Order Medication and Monitoring Patients together consist of Managing Patients. Collecting Patient Data and Managing Patients share a common feature: Medical Cognitive Action. Figures 2(a), (b) and (c) depict three different sub-models of which represent three cognitive process aspects.

Identification of Information, Advanced Cognition and Medical Cognitive Action consist of upper level of the model and Emergency Performance Assessment demonstrates a sequential trajectory. The trajectory is a natural cognitive process in which medical students solve a clinical emergency problem. However, for example, Medical Cognitive Action cannot be directly observed; it can be measured by instantiating the variables of Collecting Patient Data and Managing Patients. In brief, the upper level of explanatory variables represents an advanced level of clinical cognitive process.
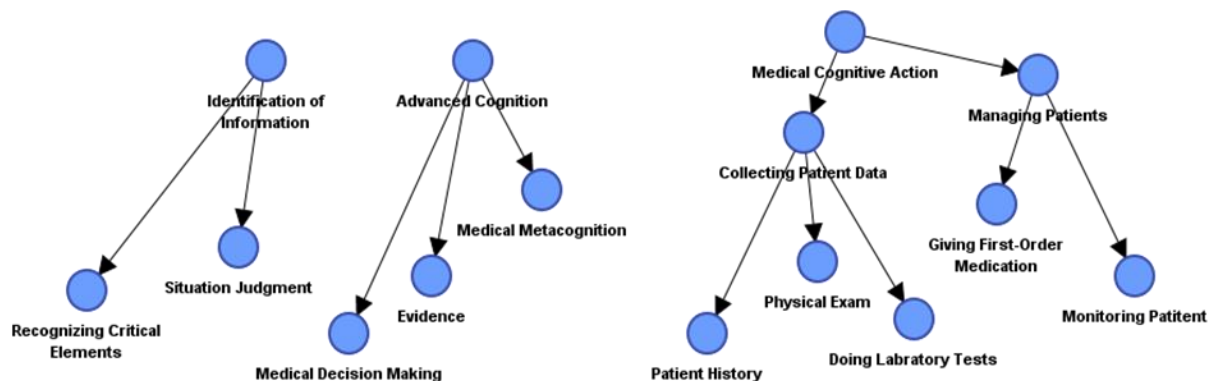

Figure 3. Cognitive Sequential models

7.4 Brief Description of the Entire Cognitive Model

The entire cognitive model is presented in Figure 4 and consists of three cognitive aspects: Identification of Information, Advanced Cognition and Medical Cognitive Action. These aspects represent clinical cognitive processes. This level in the model has played an important role in the entire hierarchical cognitive model.

These three aspects represent a general competency of clinical problem solving, which is Emergency Performance Assessment, and also called Cognitive Feature Trajectory (CFT) in the graph model. Recognizing Information directly connects to two evidence variables. Advanced Cognition is composed of three evidence variables; Medical Cognitive Action consists of two sub explanatory variables: Collecting Patient Data and Managing Patients. Collecting Patient Data is updated by three evidence variables; Managing Patients is updated by two evidence variables. Finally, Identification of Information, Advanced Cognition and Medical Cognitive Action are composed of a Cognitive Feature Trajectory called Emergency Performance Assessment in the graph model.
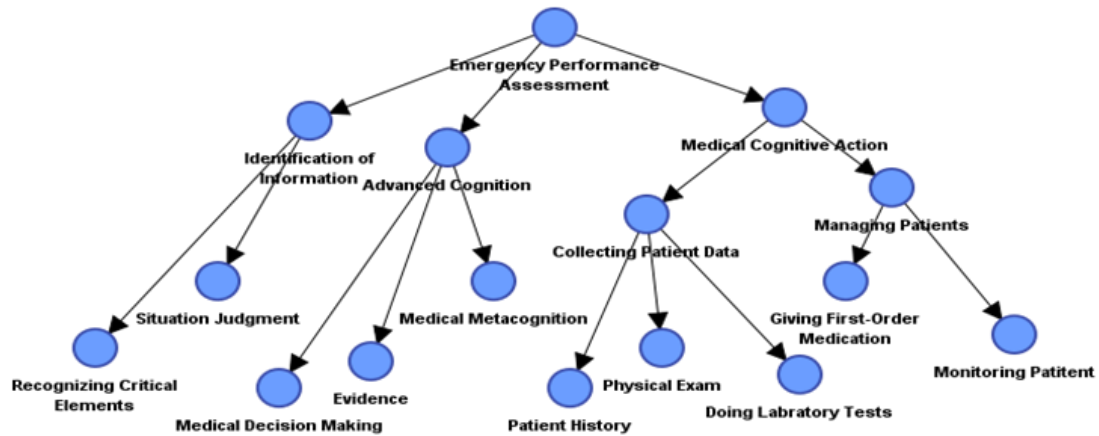
Figure 4. Hierarchical emergency performance assessment model

8. Initializing and Simulating the Bayesian Network Cognitive Model

Initializing and simulating a model are important steps to establishing a cognitive model. Initializing the model allows the model to accept instantiating values; simulating the model allows the examination and exploration of some features of the established model.

8.1 Initializing the Bayesian Network Cognitive Model

There is not much evidence of probabilities distribution in each cognitive component in the cognitive model. It is normal that any Bayesian network needs to be initialized with data which can be artificial. However, doing so does not bother the following processes. Therefore, the value of the top cognitive component is set at 0.5 as successful performance in Emergency Performance Assessment. The values of other network children (nodes), regardless of explanatory variables or evidence variables are all set at 0.7 as successful mastery of these cognitive components. The assumption of these settings is that there is no evidence to say that students express high level in Emergency Performance Assessment, so the probability of 0.5 is set in a "default" status. We have confidence to believe that students can master each sub-explanatory and evidence component at above 70% of chance after they finish a clinical cognitive process.

8.2 Simulating the Bayesian Network Cognitive Model

There are ten evidence variables which can be used to update the entire cognitive model. A random sampling method is applied to test evidence states from one true evidence to ten true evidence observations. The true status means the variable acquires a positive value based on the student's statement in the problem solving processes. In order to reduce the tedious process of the test, the random process starts from zero performance models, which means that the 10 evidence variables do not present any positive values. The student masters zero evidence variables. At the next step, the simulated data of the evidence variables demonstrates mastery of one, two and three evidence variables until complete mastery model which means the student excelled in all ten evidence variables. The mastery statuses of problem solving processes are listed in Table 1.

Table 1. Updating probabilities of random evidence combination of problem solving process

Positive Evidence	Emergency Performance Assessment	Identification of Information	Advanced Cognition	Medical Cognitive Action	Collecting Patient Data	Managing Patients
0	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
1	0.5800	0.5320	0.7000	0.5320	0.5128	0.5128
2	0.6109	0.5443	0.7106	0.6109	0.5443	0.7106
3	0.6396	0.5558	0.7206	0.6843	0.7360	0.7360
4	0.7102	0.7449	0.7449	0.7102	0.7449	0.7449

5	0.7577	0.8720	0.7613	0.7276	0.7509	0.7509
6	0.7996	0.8815	0.8815	0.7430	0.7562	0.7562
7	0.8130	0.8846	0.8846	0.7867	0.8786	0.7713
8	0.8342	0.8894	0.9470	0.7939	0.8802	0.7738
9	0.8407	0.8909	0.9478	0.8167	0.9449	0.7816
10	0.8504	0.8931	0.9490	0.8504	0.9490	0.8931

Table 1 shows relations of various instantiated evidence variables and explanatory variables. The more important column is Emergency Performance Assessment, which indicates the general level of the problem solving processes. When only one positive evidence is in the model, the Emergency Performance Assessment is 0.5800 and then increases as the evidence increases. When all 10 positive evidence fills in the model, the Emergency Performance Assessment value is 0.8504. The probabilistic range from 0.5800 to 0.8504 provides a scale to differentiate different cognitive trajectories.

8.3 Classification of the Students' Emergence Performance and Measures

The measures can be defined in different ways. One of the measures is based on the range of the entire probability and theory-driven categories. For example, if a researcher believes three categories are appropriate, e.g. lower, middle and high, the researcher can define the measure distances in three intervals: (a) $0 \leq x \leq 0.33$, (b) $0.33x \leq 0.67$, and (c) $0.67 \leq x \leq 1.00$. However, the different distributions of given specific data sets are not considered in the above measures. The categories are not sensitive to the given data sets. Therefore, the categorical validity will be weak.

The second measure can be defined in terms of the practical range of each data set. In other words, we entertain the lowest and highest values of a random variable. This measure can effectively describe the classification of a set data. If there are n categories, the classification will be defined mathematically,

$$X_c = R_p / n$$

Where X_c is the interval of the classification, R_p is the range of the probability values of a given data, and n is the number of classification. We assume the lower, middle, and high are appropriate classifications in our case. So here n is equal to 3.

The third measure can be defined in terms of the practical range of each data set and weighted B_s. Regardless of how we consider the weights from any theory and related evidence from data-drive reports, a set of weight B_s can be incorporated into the classification. Thus, the classification will be defined mathematically,

$$X_{ci} = B_i R_p / n$$

Where i is from 1 to n.

$$\Sigma B_i = 1$$

In this study, we have chosen the second measure because this is a data-driven study; we have no theories or research evidence to refer to. Therefore, the second measure is an appropriate measure in determination of the classification.

8.4 Students' Emergency Performance Assessment Categories

There are 8 participants in the study. In Table 2, Emergency Performance Assessment presents a general cognitive feature based on these students' data. Identification of Information, Advanced Cognition, and Medical Cognitive Action represent three different aspects. The probabilities beneath the four columns of cognitive terms quantitatively describe these features. Emergency Performance Assessment can be directly used to represent general "rich" level of problem solving processes, while three sub-categories emphasize different aspects. In other words, there might be similar values of Emergency Performance Assessment for each subject, but the weights of three sub-categories are probably different.

Table 2. Students' performance of the problem solving records

Subject Number	Emergency Performance Assessment	Identification of Information	Advanced Cognition	Medical Cognitive Action
S1	76.32	76.32	60.53	60.53
S2	78.55	77.09	61.42	67.92
S3	83.49	78.79	78.79	69.86
S4	87.87	89.95	89.95	65.15
S5	86.18	79.12	79.72	80.33
S6	89.3	90.27	80.79	83.35
S7	92.70	91.05	95.84	85.62
S8	92.99	91.11	95.88	87.32

The L, M and H represent lower, middle and high, three skill levels of problem solving processes. The values are probabilities that indicate the extent to which the judgment of levels can be trusted. These levels provide suggestions to researchers and students and indicate what kinds of information and at what levels they were produced, as well as what clinical information collection and application processes are useful in a problem solving process.

Table 3. Students' performance category with hierarchical scores and assessment category

Subject Number	Emergency Performance Assessment	Identification of Information	Advanced Cognition	Medical Cognitive Action
S1	76.32 (L)	76.32 (L)	60.53 (L)	60.53 (L)
S2	78.55 (L)	77.09 (L)	61.42 (L)	67.92 (L)
S3	83.49 (M)	78.79 (M)	78.79 (M)	69.86 (M)
S4	87.87 (H)	89.95 (H)	89.95 (H)	65.15 (M)
S5	86.18 (H)	79.12 (H)	79.72 (H)	80.33 (H)
S6	89.30 (H)	90.27 (H)	80.79 (H)	83.35 (H)
S7	92.70 (H)	91.05 (H)	95.84 (H)	85.62 (H)
S8	92.99 (H)	91.11 (H)	95.88 (H)	87.32 (H)

Based on the results in Table 3, Emergency Performance Assessment can be categorized into 3 general groups: L, M and H groups.

However, if we look at the whole patterns of the top four latent variables, students' performance scores can be classified into four different patterns, which can be seen in Table 4.

Table 4. Students' performance category with only assessment category

Score Pattern	Emergency Performance Assessment	Identification of Information	Advanced Cognition	Medical Cognitive Action	Student
Pattern One	L	L	L	L	S1, S2
Pattern Two	M	M	M	M	S3
Pattern Three	H	H	H	M	S4
Pattern Four	H	H	H	H	S5, S6, S7 S8

There are four patterns based on the combination of each explanatory level in Table 4.

Pattern one indicates that there are four L levels in these four latent variables. Subject S1 and S2 are in this category. Pattern two indicates that there are four M levels in these four latent variables. Subject S3 belongs to this category. Pattern three indicates that the first three latent variables are at H levels and the last latent variable, Medical Cognitive Action is at M level. Subject S4 belongs to this category. Pattern four indicates that all of four latent variables are at H levels. Subject S5, S6, S7 and S8 are in this category.

From each column of Table 4 we can recognize the different patterns of these students' performance. For example the latent variable Medical Cognitive Action is a challenge learning task. The learning pattern is L, M, M and H, comparing with other three latent variables, which are L, M, H and H. The Medical Cognitive Action consists of two sub-latent variables: Collecting Patient Data and Managing Patients. Both of these two tasks are challenging learning tasks. Thus the instructor should pay more attention to this learning task.

9. Conclusions

This study explores a cognitive process indicating how medical students recognize information, experience deep cognition, and take action in a simulated emergency medicine situation. The cognitive process describes students' clinical information collection and application process. A general explanatory variable has been defined as Emergency Performance Assessment. The Emergency Performance Assessment consists of three cognitive explanatory variables sequentially: Identification of Information, Advanced Cognition, and Medical Cognitive Action. In order to further examine the Emergency Performance Assessment, a hierarchical cognitive model has been established through the analysis of the student learning task records. The explanatory variables of the Emergency Performance Assessment cannot be observed. They are latent variables for explaining students' different features in clinical problem solving processes. Ten evidence variables are developed through the idea unit data analysis. These instantiated evidence variables are used to update the entire model.

A Bayesian network was applied to propagate the information from evidence to explanatory variables. A set of simulated network data provides a scale that allows knowing the variation of the students' levels in clinical information collection and application. Medical Cognitive Action results indicate a lower value. The other three latent variables indicate higher scores, which present four different patterns. These patterns can provide differential information to analyze different stages in a problem solving process. When facing a different task, students have to find and select related information with the clinical case. It is very challenging for students to make decisions and to take medical cognitive action. It is reasonable that students can contribute more in the emergency scenario if they have more opportunities to search for more information from available resources such as an electronic clinical library or from the clinical tutor.

In brief, the hierarchical cognitive model has effectively identified different types of students' competencies in problem solving processes. The model provides moment-by-moment differential information for both students and cognitive feature categories. The cognitive Bayesian network will be more robust in differentiating different student groups and cognitive feature categories with updating evidence by the Bayesian network learning. The cognitive construct is established by theory-driven analysis. Therefore, the construct validity will be further explored with extended clinical data.

In addition, this study takes a mixed methods design. There are two steps in this design. The first step is qualitative data collection, in which cognitive task analysis method was used. The second step is to quantify the collected data in the cognitive task analysis. Further, a Bayesian network was developed to quantitatively represent the trajectories of the student problem solving processes. This qualitative-quantitative sequential design is a practical example for medical instructors to design their students' learning and achievement assessment in medicine.

10. Limitations

This study had data from 8 students. The analyses and findings are not generalizable to all clinical learning environments. However, the Bayesian network model and cognitive task analysis are both good examples for medical instructors to think about how to evaluate student's learning in a complex cognitive environment.

References

Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian networks in educational assessment*. New York, NY: Springer. https://doi.org/10.1007/978-1-4939-2125-6

Annett, J. (2000). Theoretical and pragmatic influences on task analysis methods. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 3-23). Mahwah, NJ: Lawrence Erlbaum.

Aryadoust, V. (2011). Cognitive diagnostic assessment as an alternative measurement model. *JALT Testing & Evaluation SIG Newsletter, 15*(1), 2-6.

Ayala, C. C., Yin, Y., Shavelson, R., & Vanides, J. (2002, April). On the cognitive validity of science performance assessment task types. *Paper presented at the AERA conference*, New Orleans, LA, USA.

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* (pp. 25-37). Mahwah, NJ: Lawrence Erlbaum.

Conrady, S., & Jouffe, L. (2015). *Bayesian networks and BayesiaLab: A practical introduction of researchers.* Franklin, TN: Bayesian USA.

Creamer, E. G. (2018). *An introduction to fully integrated mixed methods research*. Thousand Oaks, CA: SAGE. https://doi.org/10.4135/9781071802823

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE.

Culbertson, M. J. (2016). Bayesian networks in educational assessment: The state of the field. *Applied Psychological Measurement, 40*(1), 3-21. https://doi.org/10.1177/0146621615590401

De la Torre, J., & Minchen, N. (2014). Cognitively diagnostic assessment and cognitive diagnosis model framework. *Psicologia Educativa, 20*, 89-97. https://doi.org/10.1016/j.pse.2014.11.001

DeCuir-Gunby, J. T., & Schutz, P. A. (2017). *Developing mixed methods proposal: A practical guide for beginning researchers.* Thousand Oaks, CA: SAGE. https://doi.org/10.4135/9781483399980

Elm, W. C., Potter, S. S., Gualtieri, J. W., Easters, J. R., & Roth, E. M. (2003). Applied cognitive work analysis: A pragmatic methodology for designing revolutionary cognitive affordance. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 357-382). Mahwah, NJ: Lawrence Erlbaum. https://doi.org/10.1201/9781410607775.ch16

Frederiksen, C. H. (1986). Cognitive models and discourse analysis. In C. R. Cooper, & S. Greenbaum (Eds.), *Studying writing: Linguistic approaches* (pp. 227-267). Beverly Hills, CA: Sage.

Goertzen, J. R. (2010). Dialectical pluralism: A theoretical conceptualization of pluralism in psychology. *New Ideas in Psychology, 28*, 201-209. https://doi.org/10.1016/j.newideapsych.2009.09.013

Greene, J. C., & Hall, J. N. (2010). Dialectics and pragmatism: Being of consequence. In A. Tashakkori, & C. Teddlie (Eds.), *Sage handbook of mixed methods in social & behavioral science* (pp. 119-143). Thousand Oaks, CA: Sage. https://doi.org/10.4135/9781506335193.n5

Hayes, J. R. (1989). *The Complete Problem Solver*. Hillsdale, NJ: Erlbaum.

Henri, F. (1992). Computer conferencing and content analysis. In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 1 15-136). New York: Springer. https://doi.org/10.1007/978-3-642-77684-7_8

Hollnagel, E. (2003). Prolegomenon to cognitive task design. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 3-15). Mahwah, NJ: Lawrence Erlbaum. https://doi.org/10.1201/9781410607775

Jonson, B. (2017). Dialectical pluralism: A metaparadigm whose time has come. *Journal of Mixed Methods Research, 11*(2), 156-173. https://doi.org/10.1177/1558689815607692

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques.* Cambridge, MA: The MIT Press.

Lajoie, S. P. (2003). Transitions and trajectories for studies of expertise. *Educational Researcher*, *32*(8), 21-25. https://doi.org/10.3102/0013189X032008021

Lajoie, S. P., Greer, J. E., Munsie, S., Wilkie, T., Guerrera, C., & Aleong, P. (1995, Dec.). Establishing an argumentation environment to foster scientific reasoning with Bio-World. *Paper presented at the International conference on computers in education*, Singapore.

Lesgold, A., Lajoie, S. P., Logan, D., & Eggan, G. M. (1990). Cognitive task analysis approaches to testing. In N. Frederiksen, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition* (pp. 325-350). Hillsdale, NJ: Lawrence Erlbaum.

Lu, J. (2007). Supporting medical problem solving with electronic whiteboard. *Unpublished doctoral*

dissertation, McGill University, Canada.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods source book.* Thousand Oaks, CA: SAGE.

Mislevy, R. J. (1994). Evidence and inference in educational assessment. *Psychometrika, 59*(4), 439-483. https://doi.org/10.1007/BF02294388

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice Hall.

Pearl, J. (1988). *Probabilistic reasoning in intelligent system: Network of plausible inference.* San Mateo, CA: Morgan Kaufmann. https://doi.org/10.1016/B978-0-08-051489-5.50008-4

Riera, B., & Debemard, S. (2003). Basic cognitive principles applied to the design of advanced supervisory systems for process control. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 255-281). Mahwah, NJ: Lawrence Erlbaum. https://doi.org/10.1201/9781410607775.ch12

Rose, J., Bearman, C., Naweed. A., & Dorrian, J. (2019). Proceed with caution: using verbal protocol analysis to measure situation awareness. *Ergonomics, 62*(1), 115-127. https://doi.org/10.1080/00140139.2018.1527951

Seemster, T. L., Redding, R. E., & Kaempf, G. L. (2000). A skill-based cognitive task analysis framework. In J. M. Schraagen, S. F. Chipman, & V. L. Shalin (Eds.), *Cognitive task analysis* (pp. 135-139). Mahwah, NJ: Lawrence Erlbaum.

Tan, C. S., & Tay, Y. H. (2010). A preliminary study of the progress on using Bayesian networks for educational assessment. *Proceedings of the 2010 IEEE conference on sustainable utilization and development in engineering and technology universiti Tunku Abdul Rahman, Malaysia, 75-77.*

Werner, M. (1995). A model for medical decision making and problem solving. *Clinical Chemistry*, *41*(8), 1215-1222. https://doi.org/10.1093/clinchem/41.8.1215

Zhang, Z. (2007). Cognitive assessment in a computer-based coaching environment in higher education: Diagnostic assessment of development of knowledge and problem-solving skill in statistics. *Dissertation Abstract International: Section A. Humanities and Social Sciences, 68*(10-A), 4271.

Zhang, Z. (2016). Assessment of matrix multiplication learning with a rule based analytical model-A Bayesian network representation. *International Education Studies, 9*(12), 182-193. https://doi.org/10.5539/ies.v9n12p182

Zhang, Z. (2018). Designing cognitively diagnostic assessment for algebra content knowledge and thinking skills. *International Education Studies, 11*(2), 106-117. https://doi.org/10.5539/ies.v11n2p106

Zhang, Z., & Frederiksen, C. H. (2007, April). Diagnostic cognitive assessment in complex domains of learning and performance: A Bayesian belief network approach. *Paper presented at the AERA annual meeting*, Chicago, IL, USA.

Zhang, Z., & Lu, J. (2014a). Assessing argumentative learning with Bayesian network models in debatable social issues. *International Education Studies, 7*(11), 120-135. https://doi.org/10.5539/ies.v7n11p120

Zhang, Z., & Lu, J. (2014b). Quantitative assessment of medical student learning through effective cognitive Bayesian representation. *International Education Studies, 7*(6), 86-97. https://doi.org/10.5539/ies.v7n6p86