

University of Texas Rio Grande Valley

**ScholarWorks @ UTRGV**

---

Teaching and Learning Faculty Publications and  
Presentations

College of Education and P-16 Integration

---

10-11-2022

## **A Mixed Assessment for the Science Learning via a Bayesian Network Representation**

Zhidong Zhang

Angelica Guanzon

Follow this and additional works at: [https://scholarworks.utrgv.edu/tl\\_fac](https://scholarworks.utrgv.edu/tl_fac)



Part of the [Education Commons](#)

---

# A Mixed Assessment for the Science Learning via a Bayesian Network Representation

Zhidong Zhang<sup>1</sup> & Angelica Guanzon<sup>1</sup>

<sup>1</sup> The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA

Correspondence: Zhidong Zhang, The College of Education and P-16 Integration, The University of Texas-Rio Grande Valley, Texas, USA. E-mail: zhidong.zhang@utrgv.edu

Received: August 11, 2022

Accepted: September 15, 2022

Online Published: November 11, 2022

doi:10.20849/jed.v6i5.1309

URL: <https://doi.org/10.20849/jed.v6i5.1309>

## Abstract

This study explored an alternative assessment model to examine Chemistry learners' progress. "The Assessment of Problem-Solving in Chemistry Learning" as a model represented students' mastery of chemistry study. The data were from journaling narratives and analyzed through cognitive task analysis. Based on the analyses, a student model was established, which represents the qualitative information in a structure, and provides a potential framework of the assessment model for the quantitative representation—a Bayesian network assessment model. The student's performance was assessed via the Bayesian network assessment model, and classified into three categories: low level, middle level, and high level. The mastery level should be at least scored at and above 90.51/100 for Declarative, Procedural, and Strategic Knowledge respectively.

**Keywords:** science learning, Bayesian network representation, student modeling, diagnostically cognitive assessment, and mixed methods design

## 1. Introduction

Developing and validating reliable assessment instruments and models are challenging processes. Many researchers and assessors think about the designs in different ways, which may be determined by assumptions and perspectivism (National Research Council, 2001; Stowe & Cooper, 2018; Toomela, 2009). However, we usually use the words, "beliefs", and "purposes" rather than assumptions and perspectivism in describing the types of assessment models. In the context of science learning, some educators want to measure what students know and what students learned in the learning tasks (National Research Council, 2001). Alternatively, other educators want to know what mistakes students make and what leads to the mistakes. In addition, the learning tasks also influence the assessment designs and models.

This study was to develop an alternative assessment through a cognitive task analysis (Hollnagel, 2003; Lesgold, Lajoie, Logan, & Eggan, 1990), and the data were from collections of chemistry learning journaling. The journaling texts were coded in terms of thematic analysis (Saldana, 2021). An assessment model was developed based on the thematic analysis, and further formed a cognitive framework. This framework consists of three critical components: declarative, procedural, and strategic (conditional) knowledge components. Three components supported the framework of the student model, and are further connected to the Bayesian network model ((Mislevy, Almond & Lukas, 2004; Zhang, 2022). Thus, a structured assessment model was developed and students' learning process, progress, and mastery can be recognized. The mixed methods design transitions qualitative or mixed data to the quantitative data representation via a Bayesian network model.

## 2. Theoretical Frameworks

Perspectivism and assumptions (Landa, Westbroek, Janssen, van Muijlwijk, & Meeter, 2020) influence how the assessment is described and defined. It is believed that the design of any assessment should target cognition or cognition-relevant elements. Knowledge, problem-solving skills, and strategic knowledge are basic components that assessment procedures and models will target (Zhang, 2007a). The perspective of measuring and assessing "what students know" with cognitive frameworks lays the foundation of modern assessment (NRC, 2001; Pellegrino, 2014).

### 2.1 Pellegrino's Triangle

The Assessment basically refers to deliberate effort with a given procedure or model to observe student learning through different means to evaluate students learning objectives. Pellegrino (2014) and other researchers (e. g., Stowe & Coper, 2018) presented that “assessment should be approached as a process of reasoning from evidence generated by students on assessment tasks has gained broad acceptance in the science education community.” (Stowe & Coper, 2018, p. 3) The assessment triangle is an important concept in Pellegrino's assessment model, which includes a) a model of student cognition and learning in the domain, b) the observation that provides evidence of students' performance, and c) an interpretation process for making sense of the evidence.

Pellegrino's Triangle model provides a very general assessment framework. The assessors holding different assumptions and perspectivism might explain the model in different ways.

### 2.2 Evidence-Centered Assessment (ECA) Design

Evidence-centered assessment design was developed at the Educational Testing Service by Mislevy, Steinberg, and Almond (2000). “This framework provides an effective structure and process for designing, producing, and delivering assessments that can be used to enhance the validity of learning assessments. The statistical mechanism of Bayesian networks connects cognitive processes and evidence from given task performances (Zhang, 2007b, p. 25). The evidence-centered assessment provides assessors and educators with an effective framework for considering how to connect between qualitative assessment information and quantitative representation.

### 2.3 Quantitative Representation of the Student Model With a Bayesian Network

Student models represent student knowledge, skills and expertise. Although they cannot be directly observed, knowledge, skills and expertise can be indirectly inferred through what students say or do which provides evidence about assessment constructs, that is, student-model variables (Zhang, 2007b; Zieky, 2014). The student modeling is to explore learning processes and extract evidence from student behavioral data for measurement of student's achievement in the context of the online environments, and establishment of a model for student mastery learning processes. The student model represents a process of data changes from a raw dataset to a structured dataset (Mislevy, Almond & Lukas, 2004).

The student model can be represented by using a quantitative construct such as Bayesian network representation. The Bayesian network represents the semantic/qualitative components and constructs the probability model separately (Koller & Friedman, 2009). Thus, the combination of a student model and a Bayesian network model provides a path to develop an assessment framework used to collect and represent the data (Zhang, 2022).

## 3. Research Methods

### 3.1 Cognitive Task Analysis

The data were from a set of simulated pilot studies that assumed the students to finish the chemical learning by journaling their understanding of the learning tasks. Cognitively we call a learning task a cognitive task.

This study seeks to analyze cognition through cognitive task analysis. Students' thinking is monitored through the use of journaling as they embark on answering questions. The cognitive component of students' understanding of chemistry concepts and ideas will be analyzed through a performance of cognitive tasks. Cognitive tasks are goal-driven (Yates, 2003) and require thinking and yield performance.

The Cognitive tasks in this study focused on the descriptions of the steps used in problem-solving in chemistry learning. Students' journal narratives were analyzed by using the three phases as classified by Linhart (1976), 1) Phase 1, the discovery of the problem / situation; 2) Phase 2, identifying the properties of the situation and trying to obtain resources that address the situation to achieve the goal as a solution process; and 3) Phase 3, confirming and substantiation of the learned concept or process and its application in the specified problem at hand. Domain-specific knowledge in students' online journal entries is identified whether they're declarative, procedural and conditional. Based on the cognitive task analysis, a structured student model was developed.

### 3.2 Journaling and Data Resources

Several cognitive tasks were included, which are in both statement and question formats:

Students will be asked to finish the following cognitive tasks:

1. Positioning the melting point capillary directly next to the temperature-sensing bulb of the thermometer is a procedure for determining a melting point. Discuss why. Explain your answers in 20 sentences using the

following vocabulary words: melting point, capillary tube, crystals, approximately, apparatus, thermometer, temperature, and degrees.

- Discuss the necessity of drying the benzoic acid crystals before finding its melting point. How would the melting point be affected if crystals are still wet? Explain your answer in 6 sentences using the following vocabulary words: crystals, remove, boiling water, dry, filter paper, benzoic acid, melting point.
- Based on your explanation, explain why a higher percentage of measurement accuracy is derived with a dry benzoic acid as compared to a wet crystal.

### 3.3 Student Model Development and Bayesian Network Representation

Students' journaling records were assessed using the three phases as categorized by Linhart (1976). Each question was coded: 1) the student understood the problem and situation; 2) the student identified the properties of the situation and tried to obtain resources, that addressed the situation to achieve the goal as a solution process; and 3) confirmed and substantiated in his / her explanation the learned concept or process and its application in the specified problem at hand.

Cognitive knowledge was identified and coded in three categories: 1) declarative knowledge, which explains presents facts; 2) procedural knowledge, which shows a collection of facts and step by step explanation of the concepts or ideas outlined in the problem; and 3) strategic knowledge, also called condition knowledge, which shows knowledge on when and where facts were derived and indicated through their writing procedures as to how they solved the problem.

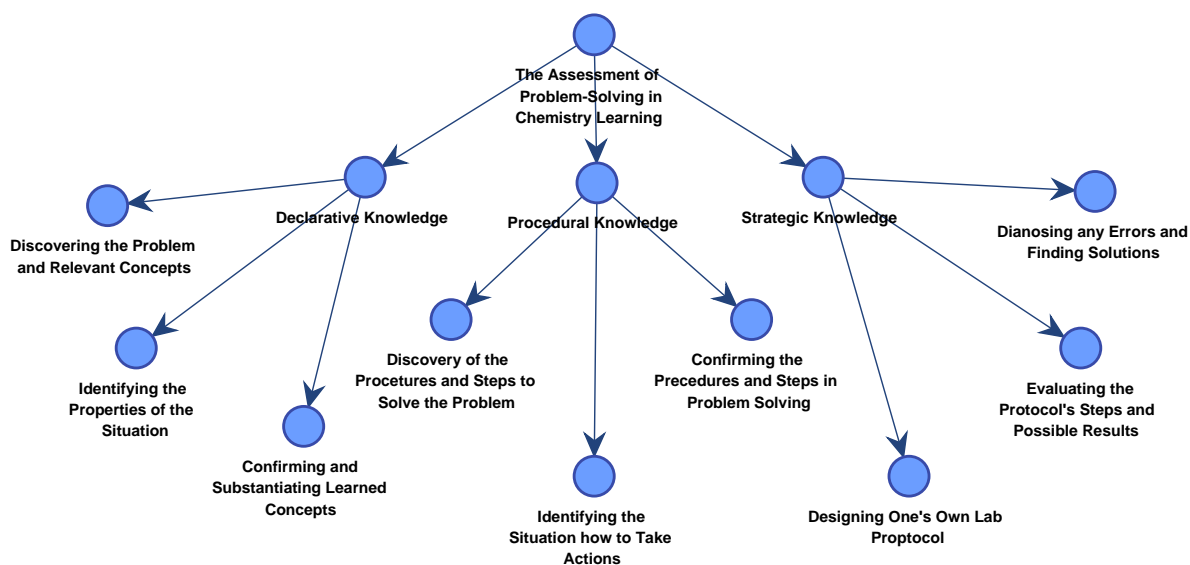


Figure 1. A Student Model and a Bayesian Representation of the Chemistry Learning

As shown in Figure 1, a student model of chemistry learning was developed, and further, the model was represented in a Bayesian network. This is a mixed representation of both the qualitative model, which was from structured cognitive task analysis and the quantitative model, which was structured by using a Bayesian network. Bayesian networks are represented a learning process and further an assessment process (Culbertson, 2016).

Koller and Friedman (2009) suggested that assuming there is a class of variables that can be designated by  $x_1, x_2, \dots, x_n$ , and  $C$ , where variables  $x_1, x_2, \dots, x_n$  are observed. The upper-level variable  $C$  means a class. It represents a concept, which is supported by all of these observed/evidence variables  $x_s$ . Thus, the relationship of all of these variables can be described in a model which factorizes as:

$$(C, x_1, x_2 \dots x_n) = (C)^n \prod_{i=1}^n P(X_i | C)$$

This model represents 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  $x_s$ .

As shown in Figure 1, there are three layers of the Bayesian network. The top one is The Assessment of Problem-Solving in Chemistry Learning, which is a concept or a latent variable. The bottom nine variables are evidential variables, which collect the data from students' journaling records.

#### 4. Results

##### 4.1 Joint Probability and Initializing the Bayesian Network

The Bayesian network needs to be initialized with subjective data (Almond, Mislevy, Steinberg, Yan, & Williamson, 2015; Jackman, 2009; Koski & Noble, 2009). The value of each variable is set at 0.67 as a successful performance, and the failed performance of each variable is set at 0.33. The network becomes stable and robust with the data increase.

There is no evidential data to propagate upwards from the bottom variables, and the model is "empty." As shown in Figure 2, the joint probabilities are: The Assessment of Problem-Solving in Chemistry Learning is 0.67 for the successful option and other variables are 0.5578 for the successful option. Figure 2 is only a part of the entire Bayesian network assessment model for illustrative purposes.

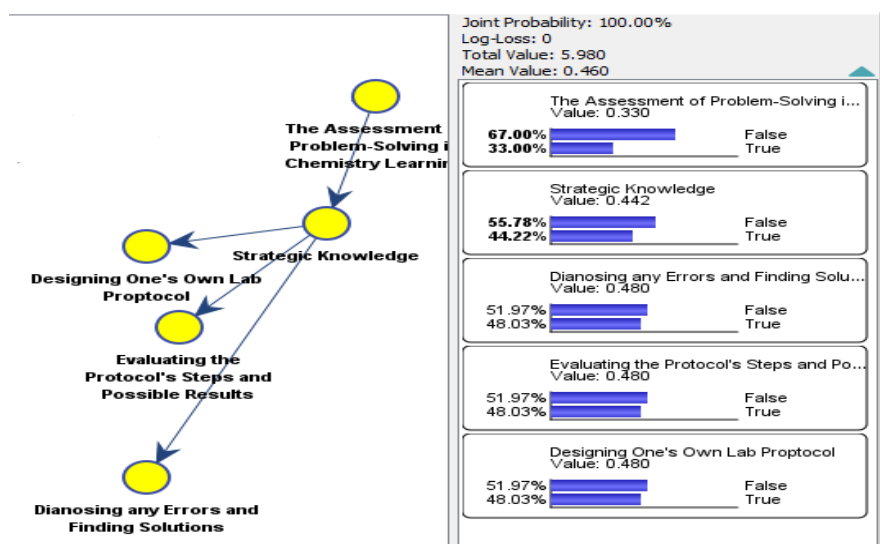


Figure 2. The initialized Bayesian network assessment

Figure 3 shows another case full of evidential data to propagate upwards from the bottom variables. The figure indicates the highest joint probabilities with both *The Assessment of Problem-Solving in Chemistry Learning* and *Strategic Knowledge* for the successful options. The *Strategic Knowledge* is 90.51% and *The Assessment of Problem-Solving in Chemistry Learning* is 71.84%. This means if a student receives the full scores of the nine evidential variables based on the "empty model," the student can receive as high as 71.84/100 at the top level—*The Assessment of Problem-Solving in Chemistry Learning*. The strategic knowledge score at 90.51/100.

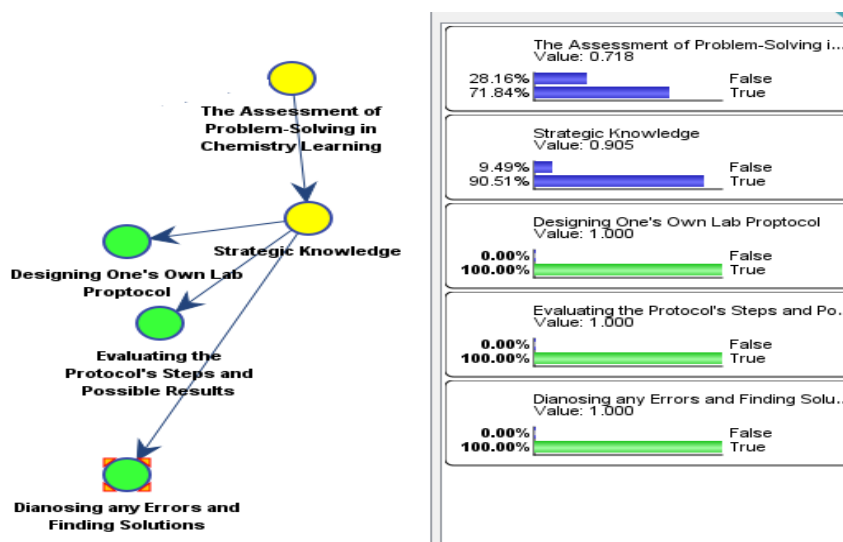


Figure 3. The full evidence model

### 4.2 Data, Evidence and Findings

In this study, the simulated data from nine chemistry students have been shown in Table 1. They completed the journaling when they learned a set of chemistry tasks. The range from 11.54 to 71.84 was established. The range is  $71.84 - 11.54 = 60.30$  and is divided into three categories: low level, middle level, and high level. Thus, 11.54—31.64 is the level one range;  $31.64^+ - 51.74$  is the level two range, and  $51.74^+ - 71.84$  is the level three range. Based on the scores in the column of “*The Assessment of Problem-Solving in Chemistry Learning,*” students 1-4 are at level one; students 5-6 are at level two, and students 7-9 are at level three.

Table 1. Students’ Performance of the Motion Learning

Student	The Assessment of Problem-Solving in Chemistry Learning	Declarative Knowledge	Procedural Knowledge	Strategic Knowledge	Assessment by Range
S1	11.54	23.04	7.17	7.17	L
S2	17.19	55.24	7.96	7.96	L
S3	22.16	83.57	8.65	8.65	L
S4	28.08	84.40	28.08	9.48	L
S5	38.32	85.83	61.68	10.91	M
S6	46.01	86.90	86.90	11.98	M
S7	53.89	88.00	88.00	35.95	H
S8	65.03	89.56	89.56	69.82	H
S9	71.84	90.51	90.51	90.51	H

### 5. Discussion

This study explored an alternative assessment model for chemistry student learning with journaling records. These were analyzed by using qualitative analysis and cognitive task analyses. A student model was developed based on these analyses. The student model can be further represented by using a quantitative construct—the Bayesian network model. The dynamic processes of the chemistry students were assessed based on the Bayesian network representation. The assessment provided an alternative model for educators and researchers to think about the assessment process and results in a different way. The journaling records can be data resources for the assessment. The data can be further explored with cognitive task analysis techniques and quantitatively

represented via a Bayesian network model. The student progress was comprehensively shared with the students and relevant stakeholders.

## 6. Limitations

This study used data from 9 students. The findings and analyses have limited generalizations. The author will collect more assessment information from other students. The proficiency student model can be varied based on different expertise. The initialized values of the variables are subjective, and these initial values can be changed based on different learning scenarios. The Bayesian network model becomes more robust with the increase in the sample size. Thus, the mastery level of the semantic explanation is relative.

## References

- Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian networks in educational assessments*. Springer.
- 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>
- Giere, R. N. (2006). *Scientific perspectivism*. University of Chicago Press
- Hollnagel, E. (2003). Prolegomenon to cognitive task design. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 3-15). Lawrence Erlbaum. <https://doi.org/10.1201/9781410607775>
- Jackman, S. (2009). *Bayesian analysis for the social science*. Wiley.
- Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. Cambridge, MA: The MIT Press.
- Koski, T., & Nobel, J. M. (2009). *Bayesian networks: An introduction*. Wiley.
- Landa, I., Westbroek, H., Janssen, F., van Muijlwijk, J., & Meeter, M. (2020). Scientific perspectivism in secondary-school chemistry education: Integrating concepts and skills in chemical thinking. *Science & Education*, 29, 1361-1388
- 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). Lawrence Erlbaum.
- Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2004). *A brief introduction to evidence-centered design* (CSE Rep. 632). Princeton, NJ: Educational Testing Service.
- Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (2000). *Evidence-centered assessment design (A submission for the NCME award for technical or scientific contribution to the field of educational measurement)*. Unpublished manuscript, Educational Testing Service in Princeton.
- National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.
- NRC. (2001). *Knowing what students know: The science and design of education assessment*. Committee on the Foundations of Assessment.
- Pellegrino, J. W. (2014). A learning sciences perspective on the design and use of assessment in education. In K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 233-252). Cambridge University Press.
- Saldana, J. (2021). *The coding manual for qualitative researchers*. Sage.
- Stowe, R. L., & Coper, M. M. (2018). Chemistry education research from personal empiricism to evidence, theory, and informed practice. *Chemical Reviews*, 118(12), 6053-6087. <https://doi.org/10.1021/acs.chemrev.8b00020>
- Toomela, A. (2009). *The psychology of scientific inquiry*. Springer.
- Yates, K. (2014). *Towards a taxonomy of cognitive task analysis methods: a search for cognition and task analysis interactions* [Data set]. University of Southern California Digital Library (USC.DL). <https://doi.org/10.25549/USCTHESES-M378>
- Zhang, Z. (2007a). 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. (2007b). *Cognitive assessment in a computer-based coaching environment in higher education: Diagnostic assessment of development of knowledge and problem-solving skill in statistics* (Doctoral Dissertation), Ottawa, ON: Library and Archives Canada (ISBN: 978-0-494-32337-3).
- Zhang, Z. (2022). A mixed methods design for assessing physics learning in the online learning environment. *Journal of Education and Development*, 6(2), 28-36.
- Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa*, 20, 79-87.

### **Copyrights**

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).