A Mixed Methods Design for Assessing Physics Learning in the Online Learning Environment

Zhidong Zhang[1]

[1] 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

Abstract

This study explored a Bayesian assessment model for physics students in motion learning. The simulated data was applied in examination of the Bayesian assessment model, The study used a mixed-methods design. The exploratory sequential model was developed based on a motion learning student model, which was a structured data collection template. The combination of the student model and the Bayesian network model provided an assessment tool for assessing physics students' learning in a dynamic process. The study reported that there were three different patterns for a physics student motion learning: lower performance, middle performance, and higher performance. In each pattern, the students may have different performance combinations of the twelve bottom components. These are shown in Figure 4 and used to collect students' performance data.

Keywords: physics learning, motion learning, mixed-methods design, exploratory sequential model, student model

1. Introduction

Developing an achievement assessment method for online learning is a crucial endeavor and a challenging aim for psychometric and measurement researchers. The alternative assessment models are very effective to enhance STEM students' progress, and to diagnose the problems encountered in the learning processes. Traditional forms of assessment have failed to satisfy the expectation of diagnostic feedback. Both Classical Test Theory (CTT) and Item Response theory (IRT) locate test-takers on a trait scale by providing only an overall score of their proficiency level in the target domain (Rupp, Templin, & Henson (2010). This assessment can reveal detailed information to the users of a test, and disclose the test-taker's weaknesses and strengths in the pre-specified sub-skills of the physics learning.

1.1 The Research Problem

There are principally three reasons why we believe the development of an achievement assessment method is a very demanding goal. First, although online learning provides students with useful tools for acquiring knowledge and developing problem-solving skills, and strategies, there does not exist an off-the-shelf test procedure to examine students' mastery and progress of knowledge, skills, or strategies in the learning environment. Second, different behavioral and cognitive representations in an assessment structure are needed regarding students' mastery of knowledge, skills and strategies. Third, online learning with dynamic and performance characteristics requires a dynamic network structure to represent a mastery process of knowledge acquisition and development of problem-solving skills.

1.2 The Objectives of the Study

This study modeled a dynamic process of learning behaviors with a hierarchical Bayesian approach and a student model (Mislevy, Almond & Lukas, 2004) with the data from the student learning (Conrady & Jouffe, 2015; Zhang & Zhang, 2020) as physics students solved physics problems in the environment of an online learning system. When physics students solved physics problems via the online learning system, their learning behaviors were represented in a student model (Schraagen, Chipman & Shalin, 2000). A hierarchical Bayesian model was applied to represent the student learning with the student model, and then students' performances were assessed via the hierarchical Bayesian model. Methodologically, this is a mixed-methods design.

This study had three major objectives: (1) identifying a student model that represents mastery of knowledge and skills in solving physics problems, (2) modeling students' learning behaviors using a Bayesian network model, and (3) examining the students' learning behavior and progress with the simulated data via the hierarchical Bayesian model as a behavioral model.

2. Theoretical Background

2.1 Online Learning

Online learning systems have rapidly developed in complex domains such as engineering and the sciences (Vanlehn, Niu, Siler, & Gertner, 1998; Warnakulasooriya, Palazzo, & Pritchard, 2005, 2007). These online systems have not only fostered knowledge and problem-solving skills (Corbett & Anderson, 1995), but also provided opportunities for students to explore problem-solving strategies (Jensen, 2001; Pearl, 1988, 2000; VanLehn, 1989, 2001). Thus, assessing students' learning has become an increasingly challenging issue (Pellegrino, Chudowsky, & Glaser, 2001). The development of the student model provides a possibility to solve this problem.

2.2 Student Modeling

The student modeling is to explore learning processes and extract evidence from student behavioral data for measurement of student 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).

2.3 MasteringPhysics

MasteringPhysics (Lee, Palazzo, Warnakulasooriya, & Pritchard, 2008) is an online problem-based tutorial system developed at the Massachusetts Institute of Technology (MIT). It provides students with a mastery learning environment that allows students to develop related physics knowledge and problem-solving skills by solving physics problems. Students consult with the module *Tutorials* to receive hints and feedback when they encounter difficulties during problem-solving.

When the students are involved in physics learning and complete their homework via the MasteringPhysics tutorial system, physics students start to choose physics problems from the *Problem Library* module. They choose physics problems initially from different topics. These problems are divided into three types based on learning consideration: skill-building, self-tutoring, and end-of-chapter problems. Finally, difficulty indices of the physics problems are reported as a student chooses a given item.

3. Methods

3.1 The Mixed Methods in Assignment Design

There are diverse mixed methods models, which may be applied to the alternative assessment (Zhang, 2007b; Zhang, & Zhang, 2020). Creswell and Plano Clark (2018) introduced six mixed methods designs, which were applied to the research designs and data collections of the cognitively diagnostic assessment (Zhang, 2007a, 2007b): a) convergent parallel design, b) explanatory sequential design, c) exploratory sequential design, d) embedded design, e) transformative design, and f) multiphase design. Among these mixed methods designs, the exploratory sequential design is the best candidate to apply to the cognitively diagnostic assessment. The model consists of two modules: a qualitative module and a quantitative module.

3.2 The Mixed Model and Student Model

The exploratory sequential design is the best model which consists of two modules: a qualitative module and a quantitative module. This model allows assessors or researchers to collect data in any qualitative or analytical data environment, such as coding into categories, thematic analysis, assertion development, and cognitive task analysis (Miles, Huberman & Saldana, 2014). The coding schemata and cognitive task analysis modes provide a possibility to transfer the data from the raw data to a given model such as the student model. In order to quantify the student model, the Bayesian network model is used to connect the student model (Mislevy, Almond & Lukas, 2004).

The Bayesian network represents the semantic/qualitative components and constructed 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, in which the exploratory sequential model is used to collect and represent the data, and correspondingly, the student model mirrors the qualitative representation and the Bayesian network model constructs the quantitation representation. Figure 1 indicates that the qualitative data are represented in the student model and further the student model is quantitatively represented by the Bayesian network model.

3.3 The Data Structure

The data structure is a student model in physics learning. The student model should represent the student's current state of knowledge. That includes a representation of the knowledge, concepts, and skills the student has acquired either fully or partially. It also includes the representation of a given student's special skills and needs. Moreover, there should be a mechanism to represent misconceptions, bugs, or erroneous information that the student might have acquired. (ScienceDirect, 2021, p. 1)
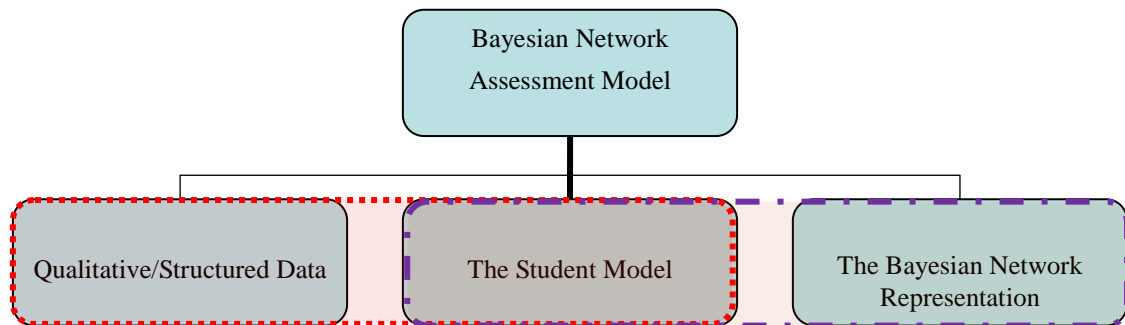

Figure 1. Qualitative data, student model and Bayesian network representation

One of the learning tasks in physics, learning is to describe *the motion*, which is the topic of MasteringPhysics. There may be several different student models of the topic; one model was selected as a student model. As shown in Figure 2, the structure of the motion concept map is the data template. From the Bayesian network perspective, learning the Bayesian network was from the students' perception or performance on the cognitive task—motion learning.
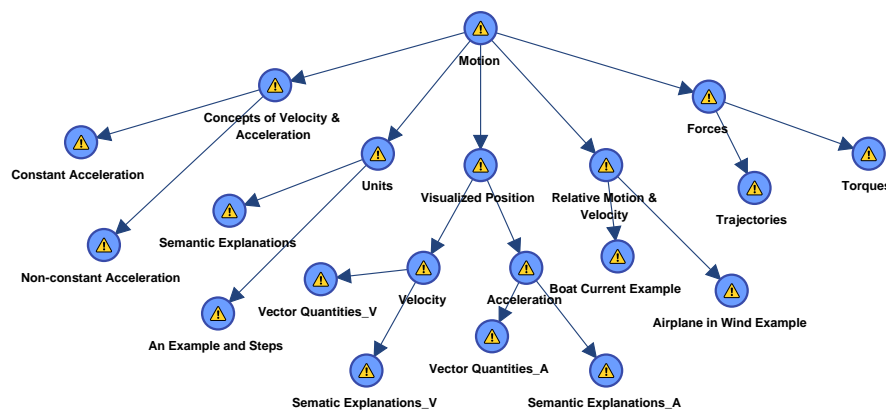

Figure 2. The student model of the motion learning

3.4 Bayesian Network Theory and Model

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 (Pearl, 1988). Koller and Friedman (2009) state that, assuming there is a class of variables which can be designated by x_1, x_2, ... x_n, and C. The structure can be seen in figure 3. In such an example variable x_1, x_2, ... 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:

$$P(c, x_1, x_2 \cdots x_n) = P(c) \, \Pi^n_{i=1} \, P(X_i \,|c)$$

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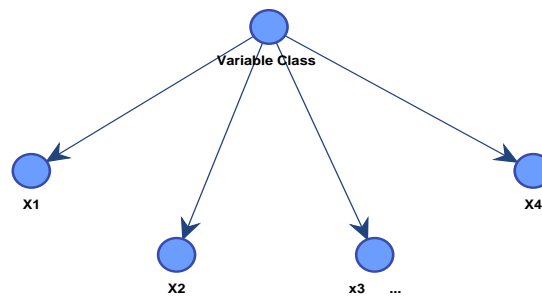

Figure 3. A Bayesian network graph

As soon as a learning behavioral construct is developed, how to develop a statistical model becomes most important. The statistical model represents the learning behavioral construct elegantly in a quantitative way. The hierarchical Bayesian approach is a powerful technique that can characterize a behavioral process in a network representation (Mislevy, Almond, & Lukas, 2004; Mislevy & Gitomer, 1996). Hierarchical structure can be updated by inputting data into the model at any time. Further, the errors and deficiencies of students' performance can be clearly diagnosed by entering the individual student's data into the model and then examining the updated model status. Thus, the hierarchical Bayesian models are dynamic and diagnostic in providing assessment evidence and representing students' progress based on the established student model (Zhang, 2007b; Zhang, Lu, & Wiseman, 2008).

As shown in Figure 4, the learning of motion includes five components: Concepts of Velocity and Acceleration, Units, Visualized Position, Relative Motion and Velocity, and Force. When the data of the student performance enters the bottom layers of the components, the Bayesian network model is updated with new data, which report assessment information and student progress level. These bottom layers of the components include Constant Acceleration, Non-constant Acceleration, Semantic Explanation to the Units, An Example and Steps of the Units, Vector Quantitative V, Semantic Explanations V, Quantitative_A, Semantic Explanations_A, Boat Current Example, Airplane in Wind Example, Trajectories of Forces and Torques of Force.
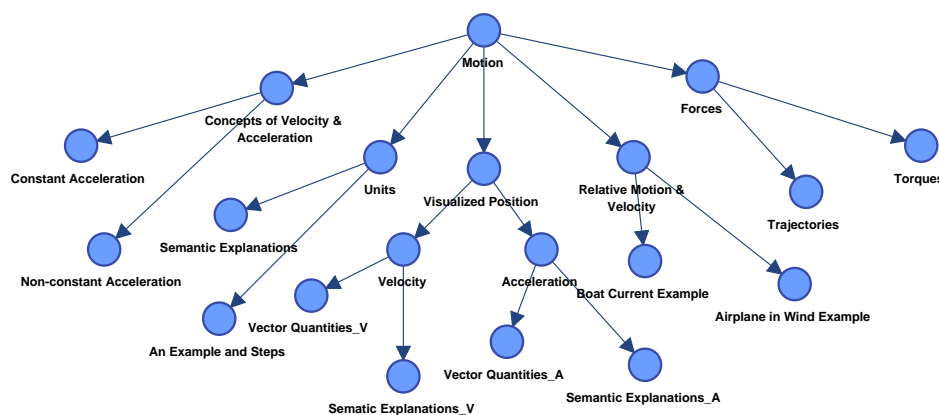

Figure 4. The Bayesian network assessment model of the motion learning with initiative probabilities

3.5 Joint Probability and Initializing the Bayesian Network Model

There is not much evidence of probability distribution in each component in the assessment 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 component is set at 0.7 as successful performance

in Variable Motion. The values of other network nodes, regardless of explanatory variables or evidence variables are also set at 0.7 as successful mastery of these components. The assumption of these settings is that there is no evidence to say that students express a high level in the physics learning process, so the probability of 0.7 is set at the "expecting" status. We have the confidence to believe that students can master each sub-component and evidence component at above 70% of chance after they finish the physics learning process.

A joint probability is the probability of two events occurring together. If they were even A and event B, the probability of the intersection of event A and a B may be written p (A ∩ B). For example, we focus on the joint probability of the event that Motion, Concepts of Velocity & Acceleration, Units, Visualized Position, Relative Motion & Velocity, and Forces success together. Figure 5 presents the joint probability of these variables.
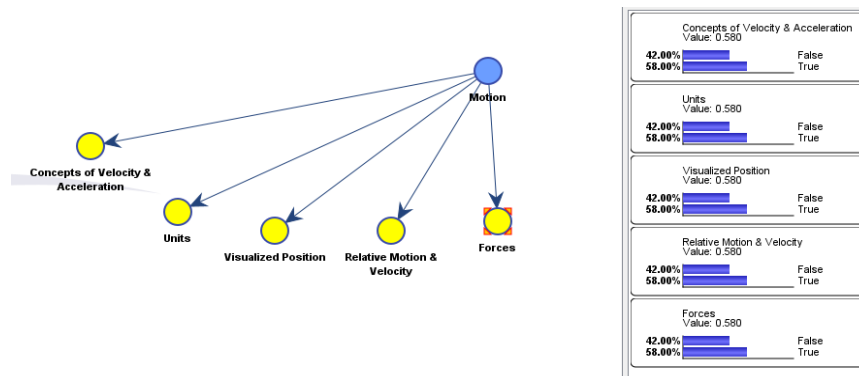

Figure 5. The initialized Bayesian network assessment model

4. Data Resources and Evidence

The data are simulated from a group of physics students who learned physics via the *MasteringPhysics* tutorial system in the study. The student model was developed in this pilot study and further, the Bayesian Network Assessment Model was applied to measure and report students' performance in motion learning. Figure 6 is a simplified evidence model for the illustration purpose. If variable Concepts of Velocity & Acceleration, Units, Relative Motion & Velocity, and Forces gain successful points; Visualized Position gain a failed point, the score of Variable Motion is 96.74%, which is very high. Figure 6 is only for illustration purposes. In fact, the evidence should be from the evidential variables in Figure 4. These evidential variables are Constant Acceleration, Non-constant Acceleration, Semantic Explanations, An Example and Steps, Vector Quantities V, Semantic Explanation V, Vector Quantities A, Semantic Explanations A, Boat Current Example, Airplane in Wind Examples, Trajectories, and Torques. The varieties of the evidence combination of these evidential variables provide the quantitative information for all of the latent variables and the top variable, Motion.
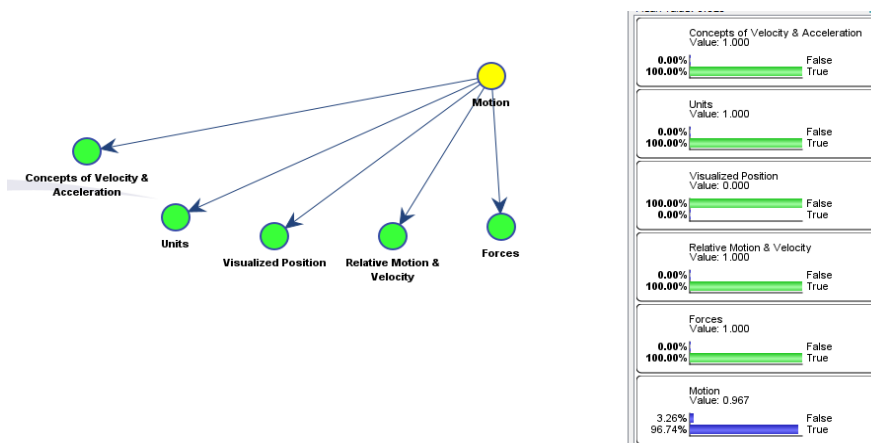

Figure 6. A simplified evidence Bayesian network model

In this pilot study, the simulated data for nine physics students in motion learning have been shown in Table 1. The range was established as follows: 85.00-88.00 is a lower performance range, 88^+-90 is a middle range, and 90^+ is the higher range. Based on the ranges, the students were classified into three groups. Students, S1, S2, S3, and S4 were classified into the lower performance group; Students S5 and S6 were classified into the middle group; and students S7, S8, and S9 were classified into the higher performance group.

Table 1. Students' Performance of the Motion Learning

Student	Concepts of Velocity & Acceleration	Units	Visualized Position	Relative Motion & Velocity	Forces	Motion	Assessment by Range
S1	75.96	86.02	71.56	75.96	61.96	85.18	L
S2	76.13	76.13	71.73	76.13	76.13	85.72	L
S3	86.62	76.78	72.43	76.78	76.78	87.88	L
S4	76.34	76.34	75.39	76.34	76.34	86.44	L
S5	77.35	77.35	73.04	87.04	87.04	89.75	M
S6	77.14	77.14	79.29	86.89	77.14	89.06	M
S7	87.50	87.50	77.06	87.50	77.98	91.81	H
S8	78.11	87.59	80.22	87.27	87.59	92.23	H
S9	87.87	87.87	80.58	87.87	87.87	93.48	H

5. Conclusions

This study explored an alternative assessment model for physics students in the motion learning model. The data was collected in physics students' learning Motion process by using a mixed-methods design. The exploratory sequential model was developed in the physic learning environment. The student model from a student motion learning was taken as a structured data collection template. The fusion of the student model and the Bayesian assessment model provided an assessment tool for assessors to assess physics students' learning in a dynamic process.

The results reported that there were three different patterns for physics student motion learning: lower performance, middle performance, and higher performance. The score ranged from 85.18 to 93.48. We also found that the scores of the components became robust and stable. Except for the component of Concepts of Velocity & Acceleration, other components: Units, Visualized Position, Relative Motion & Velocity, Forces, and Motion all kept increasing values from student 8 to student 9. Student 9 received the full successful scores. Thus, the Bayesian assessment model not only reported dynamic assessment results but also diagnosed the mastery point, which was like a cut-off score to inform what was the level of the mastery. This model suggested that a student should master 11/12 components.

6. Scholarly Significance of the Study

This study established a mixed-methods design in an alternative assessment model. The exploratory sequential design was recognized. The data can be any qualitative or structured data sets. The Bayesian network assessment model was developed based on the student model, which was a "hinge" to link the structured data set and the Bayesian network model.

The evidence-centered design has been applied in the alternative assessment for more than 2 decades. Qualitative dataset as the evidence was a meaningful practice for educators in the learning and assessment of the STEM fields.

Mastery is an important concept in science learning such as MasteringPhysics. However, this is the first time to quantify the cut-off score for physics learning. Such a mastery learning cut-off score is a model-based and data-driven determination.

Mixed-methods designs, structured datasets, student models, and Bayesian network assessment models suggest many potential applications of the alternative assessment in relevant STEM fields.

7. Limitations

This assessment model was developed with only 9 students' data in physics learning. The conclusions were limitedly generalized to different sample groups. The learning task was also structure-specific. The exact models for different learning tasks in physics are not expected.

References

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

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*, 253-278.

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

Jensen, F. V. (2001). *Bayesian networks and decision graphs*. New York, NY: Springer-Verlag.

Koller, D., & Friedman, N. (2009). *Probabilistic graphic model: Principles and techniques.* Cambridge, Mass: MIT.

Lee, Y., Palazzo, D., Warnakulasooriya, R., & Pritchard, D. E. (2008). Measuring Student learning with item response theory. *Physical Review Special Topics—Physics Education Research*, *4*, 0101102-0101108.

Miles, M. B., Huberman, A. M., & Saldana, J. (2014). *Qualitative data analysis: A methods sourcebook.* Thousand Oak, CA: Sage.

Mislevy, R. J., & Gitomer, D. H. (1996, June). *The role of probability-based inference in an intelligent tutoring system* (CSE Tech. Rep. 413). Princeton, NJ: Educational Testing Service.

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.

Pearl, J. (1988). *Probabilistic reasoning in intelligent system: Networks of plausible inference*. San Francisco, CA: Morgan Kaufmann.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. NY: Cambridge.

Pellegrino W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Rupp, A. A., Templin, J., & Henson, R. A. (2010). *Diagnostic measurement: Theories, methods and applications.* New York, NJ: The Guilford Press.

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

ScienceDirect. (2021). *Student model.* Retrieved from https://www.sciencedirect.com/topics/psychology/student-model

VanLehn, K. (1989). Problem solving and cognitive skill acquisition. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 527-580). Cambridge, Mass: The MIT Press.

VanLehn, K. (2001, April). *OLAE: A Bayesian performance assessment for complex problem solving*. Paper presented at the National Conference on Measurement in Education, Seattle, WA.

Vanlehn, K. Niu, Z., Siler, S., & Gertner, A. (1998). Student modeling from conventional test data: A Bayesian approach without priors. *Proceedings of the 4th ITS'98 conference* (pp. 434-443). Heidelberg, Berlin: Springer-Verlag.

Warnakulasooriya, R., Palazzo, D., & Pritchard, D. E. (2005). Evidence of problem-solving transfer in web-based Socratic tutor. *Proceedings of the Physics Education Research Conference*, 41-43.

Warnakulasooriya, R., Palazzo, D., & Pritchard, D. E. (2007). Time to completion of web-based physics problem with tutoring. *Journal of the Experimental Analysis of Bahavior, 88*, 103-113.

Zhang, Z., & Zhang, Y. (2020). Alternatively assessing conceptual learning in an emergency clinical environment: A mixed methods design. *Journal of Education and Development*, *4*(2), 27-37.

Zhang, Z. (2007a, Jan.). *Application of a Bayesian belief network to diagnostic cognitive assessment in studies in statistics learning and performance.* Presented at seminar, Faculty of Education, McGill University, Montreal, QC.

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*. Unpublished doctoral dissertation, McGill University, Montreal, QC, Canada.

Zhang, Z., Lu, J., & Wiseman, J. (2008, May). *Understanding and assessing students' clinical reasoning in the context of simulated emergency medical case*. Paper presented at the annual meeting of the Association of Faculties of Medicine of Canada, Montreal, QC.