Developing a Two-Tier Diagnostic Test to assess Arterial Blood Gases learning by students with different background knowledge in Anesthesiology

Phongthara Vichitvejpaisal[1], Vorakran jiranantarat[1], Tippawan Panjamawat[2], Pinda Varasunun[3], Parichad Apidechakul[1], Suthisa Chamadol[1], Kanitha Kraiprasit[1], Phetcharee Chainchop[1]

Abstract

Objectives: To develop a two-tier diagnostic test in medical education on the analysis of arterial blood gases (ABGs) by students with different background knowledge in anesthesiology, using a concept and knowledge map to determine table of specifications together with an open discussion and a feedback-providing method.

Methods: A research and development study in which the developed diagnostic test was assessed for its efficiency by first-year residents, preclinical medical students and nurse anesthetist students who volunteered to join the project.

Results: There were four major misconceptions in ABGs: First, they could not clarify the significance and relationship of the symbols. Second, they could not remember the formulas and use them appropriately. Third, they did not understand the analytical steps and lacked knowledge for clinical interpretation. Fourth, they could not apply the logical results as a guideline for patient management. Medical and nurse anesthetist students had problems mainly on the third/fourth misconceptions and partly on the first/second misconceptions. Nevertheless, residents had problems mainly on the fourth misconception and partly on the third misconception.

The assessment of criterion-referenced test item difficulty, discrimination and reliability (internal consistency) was 0.59, 0.38 and 0.91 respectively. The item objective congruency (IOC) of the test was equal to 0.88.

Conclusions: Using a concept and knowledge map to define the table of specifications in ABGs concepts together with an open discussion and feedback-providing method helped facilitate the scope of developing a two-tier diagnostic more practical test. Teachers can assess misconceptions of students with different background knowledge in a short period of time and have guidelines to improve pedagogy in response to their eagerness for learning.

Key Words: research and development, diagnostic test, medical education, anesthesiology

[1]Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700

[2]17 Soi Therdthai 66 Banghwa Phaseechareon, Bangkok 10160

[3]Department of Education, Faculty of Education and Development Sciences, Kasetsart University, Kamphaeng Saen Campus 73140

Corresponding author:
Phongthara Vichitvejpaisal, MD, PhD (Educational Evaluation)
Faculty of Medicine Siriraj Hospital, Mahidol University, Bangkok 10700, Thailand
Fax: +66-02-803-1294, Email: Phongthara @gmail.com

Introduction

Novak (1981, 1996); Novak & Gowin (1984) showed that the concept and knowledge map could help students benefit from a bird's-eye view of the scope of subjects. Also, it would be a stepping stone to make an educational plan for a complete practical mastery of learning. In 1985, David F. Treagust described a development of a two-tier diagnostic test to measure students' misconceptions as follows:

1) The first tier was a multiple-choice content question related to propositional statements and parts of the concept map.
2) The second tier was a multiple-choice set of reasons or open-ended questionnaires for short answers related to the first tier.

A diagnostic test was a useful tool to verify students' knowledge, to easily ascertain their misconception with a proper remedial technique and to urge their achievement motive towards active learning of a selected subject. (Turns, Atman, & Adams, 2000; Tan, Goh, Chia, & Treagust, 2003)

Based on the procedure described by Treagust together with an open discussion and feedback mechanism (Heh, 2008), a two-tier diagnostic test was developed to assess misconceptions of first-year residents, preclinical medical students and nurse anesthetist students on the analysis of arterial blood gases (ABGs), which is a comprehensive guide for taking care of anesthetized patients in the operating theatre, recovery room or intensive care unit. However, it seemed reasonable to assume that they had little understanding of this topic. As a result, they craved knowledge in the atmosphere of a fascinated training schedule. We have hypothesized that a developed diagnostic test plays an important role in medical education; it works as a probe to examine students' weaknesses in learning ability and as a medication for their weariness in pedagogical techniques.

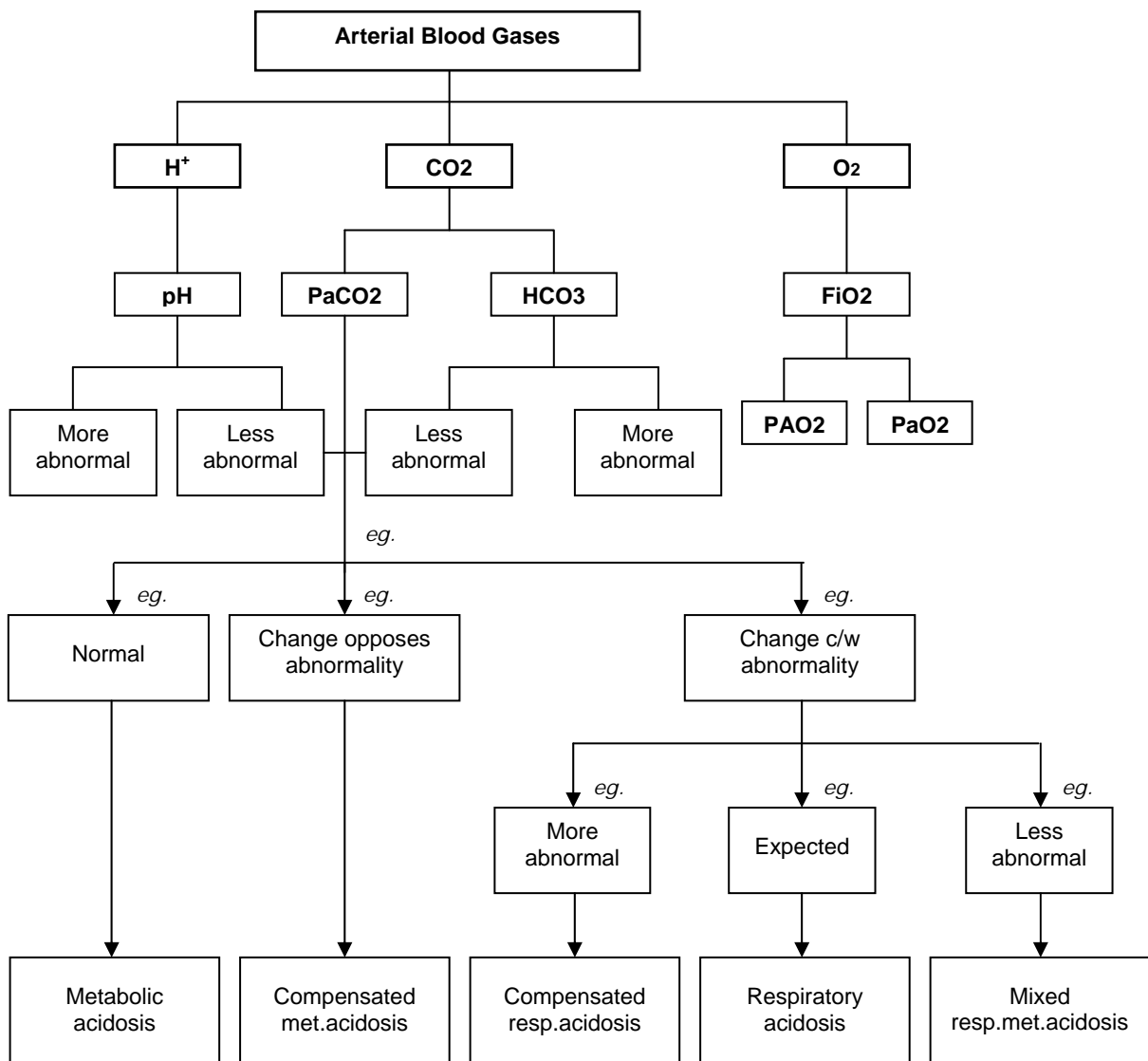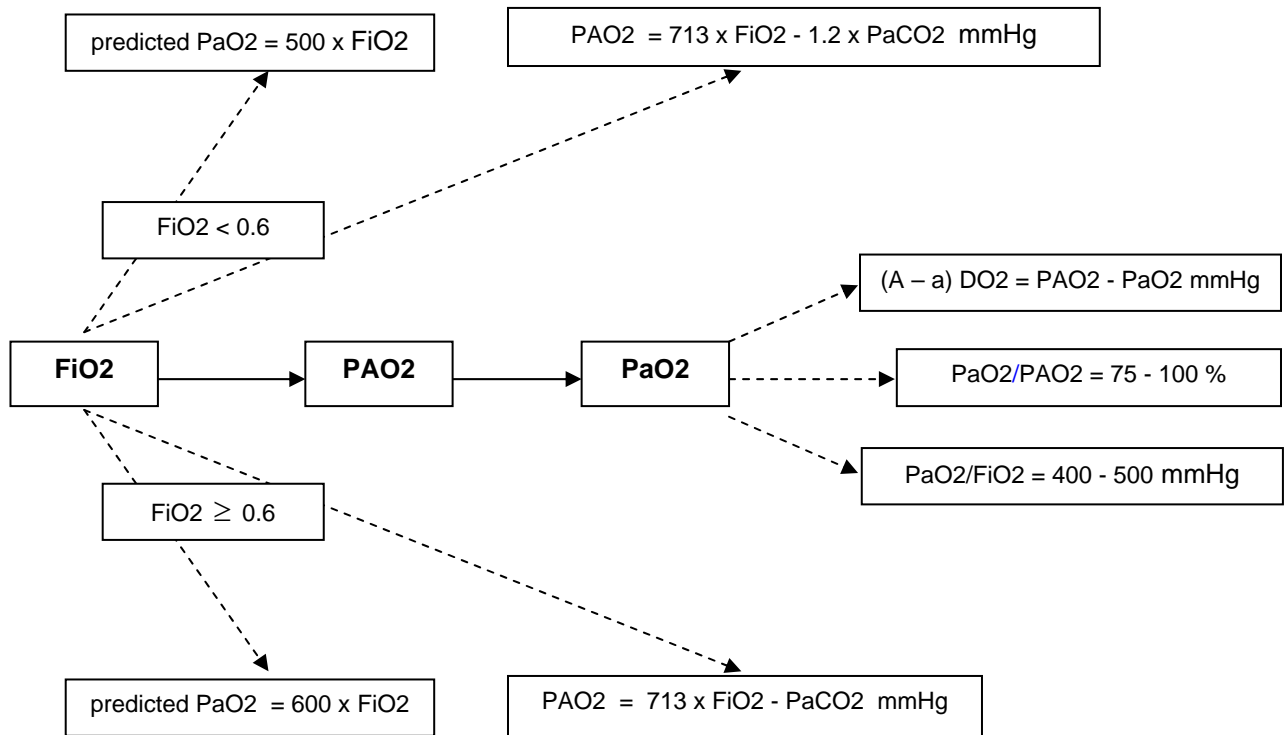*Figure 1: Concept Map: the Arterial Blood Gases Analysis*

Figure 2: Knowledge Map: The Relationship Between FiO2, PAO2, PaO2, PaCO2

predicted PaO2 = 500 x FiO2	PAO2 = 713 x FiO2 - 1.2 x PaCO2 mmHg

FiO2 < 0.6

(A – a) DO2 = PAO2 - PaO2 mmHg

FiO2 → PAO2 → PaO2

PaO2/PAO2 = 75 - 100 %

PaO2/FiO2 = 400 - 500 mmHg

FiO2 ≥ 0.6

predicted PaO2 = 600 x FiO2

PAO2 = 713 x FiO2 - PaCO2 mmHg

Methods

The sample were 22 first-year residents, 36 preclinical medical students and 23 nurse anesthetist students, Department of Anesthesiology, Faculty of Medicine Siriraj Hospital, Academic Year 2008, who completed a tutorial on the selected subject and volunteered to join this research and development study.

First tier: Analysis of the Concept (*Figure 1*) *and Knowledge Map* (*Figure 2 & 3*). This was to define the scope of ABGs concepts (table of specifications), which was then divided into three learning specification parts. Each part had at least 20% of the subjects related to each other.

- Learning specification part 1 comprised of patients' physical assessment, indications and technique of arterial puncture, equipment preparation, related complications, oxygen administration, FiO2, predicted PaO2 and alveolar air equation.
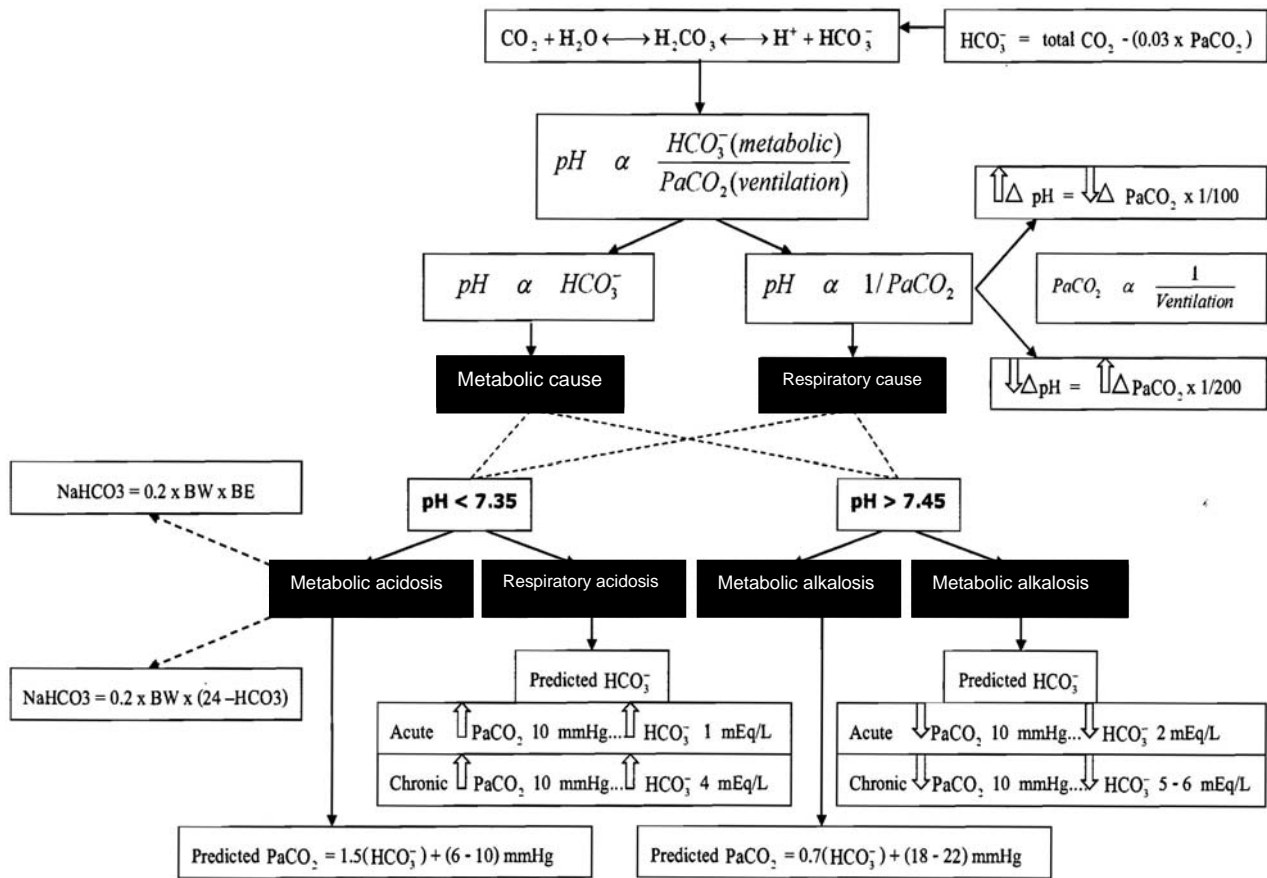
- Learning specification part 2 comprised of respiratory physiology, pulmonary pathology in any circumstances, relationship between acid-base balance and blood chemistry, prediction of acid-base disturbances, shunt estimation, predicted and measured pH, revision of predicted PaO2 and alveolar air equation.

- Learning specification part 3 comprised of pulmonary and renal pathophysiology, acute and chronic metabolic / respiratory compensation, interpretation step by step, diagnosis and guidelines to patient management, relationship between PaCO2 and HCO_3^-, electrolytes for metabolic assessment, estimation of HCO_3^-, and administration of 7.5% $NaHCO_3$.

Second tier: Activity set of the learning specifications. This was designed to find out students' misconceptions in ABGs through two paper-and-pencil tests in each set. Three sets of learning specifications took place exactly two weeks apart in which students had to pass these tests in an orderly way.

Test no.1: This consisted of fifteen multiple-choice questions requiring one best choice (30 min). The main objective was to assess the students' basic knowledge. Each test item consisted of a content question asking students to choose the outcome of a patient's real situation, which were commonly encountered by the students in their daily life and usually providing several distractors along with the correct answer. (Duit, Treagust, & Mansfield, 1996)

Figure 3: Knowledge Map: The relationship Between pH, HCO3, PaCO2

$$CO_2 + H_2O \longleftrightarrow H_2CO_3 \longleftrightarrow H^+ + HCO_3^-$$

$$HCO_3^- = total\ CO_2 - (0.03 \times PaCO_2)$$

$$pH \quad \alpha \quad \frac{HCO_3^-\ (metabolic)}{PaCO_2\ (ventilation)}$$

$$\Uparrow \Delta pH = \Downarrow \Delta PaCO_2 \times 1/100$$

$$pH \quad \alpha \quad HCO_3^- \qquad pH \quad \alpha \quad 1/PaCO_2 \qquad PaCO_2 \quad \alpha \quad \frac{1}{Ventilation}$$

Metabolic cause | **Respiratory cause**

$$\Downarrow \Delta pH = \Uparrow \Delta PaCO_2 \times 1/200$$

$$NaHCO3 = 0.2 \times BW \times BE$$

pH < 7.35 | **pH > 7.45**

$$NaHCO3 = 0.2 \times BW \times (24 - HCO3)$$

Metabolic acidosis | **Respiratory acidosis** | **Metabolic alkalosis** | **Metabolic alkalosis**

Predicted HCO_3^-

Acute	$\Uparrow PaCO_2$ 10 mmHg... $\Uparrow HCO_3^-$ 1 mEq/L
Chronic	$\Uparrow PaCO_2$ 10 mmHg... $\Uparrow HCO_3^-$ 4 mEq/L

Predicted HCO_3^-

Acute	$\Downarrow PaCO_2$ 10 mmHg... $\Downarrow HCO_3^-$ 2 mEq/L
Chronic	$\Downarrow PaCO_2$ 10 mmHg... $\Downarrow HCO_3^-$ 5 - 6 mEq/L

$$Predicted\ PaCO_2 = 1.5(HCO_3^-) + (6 - 10)\ mmHg$$

$$Predicted\ PaCO_2 = 0.7(HCO_3^-) + (18 - 22)\ mmHg$$

Test no. 2: This consisted questionnaire of ten open-end questions with regard to propositional statements requiring short answers (1 hr). The main objective was to identify students' misconceptions, which eventually served in the construction of a diagnostic test. Students had opportunities to express their understanding and reasons in response to test no.1 questions.

They were allowed to gain subject contents through an interactive web browser or available textbooks which provided them with current and relevant information. This consideration minimized the chances that students might just guess when they did not have any strongly-held conception about the question.

- At the end of test no.2, a one-hour open discussion was held for further explanation and clarification. This did not simply give feedback to students' answers to each item as mentioned by Ciofalo & Wylie (2006), but revealed the mastery of a given subject where they could find direct solutions and detailed clues. Students were supported in content understanding, problem-solving strategies, and learning success. (Bell & Cowie, 2001)

- The common misconceptions from three activity sets were analyzed for its contents and developed into a diagnostic test. This forty-question test required short answers (1 hr) was based on patients' ABGs analysis and management.

Validation and Reliability of the Test
The correctness and appropriateness of all tests (content validity and index of item objective congruence, IOC) were determined by four anesthesiologists who were not involved in the project and had at least 10 years of experience in educating residents. For additional review, test no.1 and a developed diagnostic test questions were administered to 16 second-year residents in anesthesiology. The assessment of criterion-referenced test item difficulty, discrimination and reliability (Kuder Richardson-20) was verified. The outcome was measured by scores graded by four non-physician evaluators.

Statistics

Statistical analysis was done using the Statistical Package for Social Sciences (SPSS) for Windows, release 16. Within-group and between-group test scores were compared by using one-way analysis of variance (ANOVA). A P-value < 0.05 implied statistically significant differences.

Results

Having undergone quality analysis for structural content validity, conciseness, accuracy, and clarity of the questions, test no.1 measured 0.89, 0.80 and 0.95 in IOC. The assessment of criterion-referenced test item difficulty, discrimination and reliability was 0.78, 0.89 and 0.87; 0.17, 0.11 and 0.20 as well as 0.85, 0.86 and 0.95 respectively.

Content analysis of test no.2 revealed that although the related questions had been encountered several times, the four major misconceptions in ABGs analysis were 1) they could not clarify the significance and relationship of the symbols 2) they could not remember the formulas and use them appropriately 3) they did not understand the analytical steps and lacked knowledge for clinical interpretation and 4) they could not apply the logical results as guideline for patient management. Medical and nurse anesthetist students had problems mainly on the third/fourth misconceptions and partly on the first/second misconception. Nevertheless, residents had problems mainly on the fourth misconceptions and partly on the third misconception.

For the developed diagnostic test, the IOC of the test was 0.88. The assessment of criterion-referenced test item difficulty, discrimination and reliability was 0.59, 0.38 and 0.91 respectively.

The total average test scores of all students were 25.79. The average test scores of 22 residents, 36 preclinical medical students and 23 nurse anesthetist students were 27.27, 23.00 and 23.74 respectively. There were no statistically significant differences of the mean scores between the groups. *(Table 1)*

Discussion

Test no.1 showed a low criterion-referenced test item discrimination. This might have been due in part to previous knowledge of students. As a result, it was appropriate to screen students for basic information on the selected subject and to monitor their progress. On the other hand, the developed diagnostic test was rather difficult with a high criterion-referenced test item discrimination. Therefore, it was suitable to evaluate their learning achievement.

Table 1:   Mean and Standard Deviation of the Developed Diagnostic Test Scores between the Groups

Sample groups	n	$\bar{x}$	S.D.	Levene's test	p	F	p
-Residents	22	27.27	7.60	1.407	.251	1.822	.169
-Nurse students	23	23.74	7.58				
-Medical students	36	23.00	9.46				

Misconception 1:
Students could not clarify the significance and relationship of the symbols. This implied that they could neither remember the definition nor understand the meaning of the symbols. For example, access to the ABGs analysis is on a strictly need-to-know FiO_2 basis. PaO_2 is obtained directly from blood while PAO_2 is derived from calculation. PAO_2 has a close relationship with FiO_2 by means of the alveolar air equation. Also, it reflects shunt effect via $(A - a) DO_2$ and PaO_2/PAO_2. Students should be able to distinguish the predicted pH from the measured pH. The measured pH is obtained directly from blood and varies according to $PaCO_2$ and HCO_3^- whereas predicted pH is derived from calculation and varies only according to $PaCO_2$. The difference between these entails the role of HCO_3^- and indicates the type of acid-base disturbances.

Misconception 2:
Students could not remember formulas and use them appropriately. This implied that they lacked ideas of formulas development from basic ABGs knowledge. For example, the difference in oxygen administration techniques determines the amount of FiO_2 delivered to a patient. If $FiO_2 > 0.6$, the predicted PaO_2 is $600 \times FiO_2$ and PAO_2 is equal to $713 \times FiO_2 - PaCO_2$, while if $FiO_2 < 0.6$, the predicted PaO_2 is equal to $500 \times FiO_2$ and PAO_2 is $713 \times FiO_2 - (1.2 \times PaCO_2)$. Furthermore, if $FiO_2 > 0.21$, the shunt estimation by means of $(A - a) DO_2$ is less reliable than PaO_2/FiO_2 and PaO_2/PAO_2. However, PaO_2/FiO_2 is easier to compute than PaO_2/PAO_2 because the alveolar air equation is not needed. Understandably, calculation of PAO_2 is time-consuming and error-proned.

Misconception 3:
Students did not understand the analytical steps and lacked knowledge for clinical interpretation. This implied that they did not know the process of analysis in order of priority. Normally, there followed several steps in ABGs interpretation, namely an assessment of oxygenation and ventilatory response, an existence of shunt effect, a derangement of acid-base balance and a compensatory mechanism. For example, without knowledge that FiO_2 acts as the precursor to establish the severity of the shunt effect and to reveal the pathophysiology of those clinical defects. Students preferred to focus on the oxygenation condition to the oxygen therapeutic technique and its concentration delivered to patients (FiO_2). In addition, they mainly paid attention to acid-base disturbances and rushed through

correction of defects without realizing that the body compensatory mechanism was currently working on it. As a result, the lungs and the kidney were affected with over-compensated responses and its pathological conditions approached the chronic phase with or without complications. A golden opportunity to treat patients at the proper time was wasted and a compromised situation was inevitable.

Misconception 4:
Students could not apply the logical results as the guideline for patient management. This implied that they did not understand the theoretical and practical knowledge of selected subject in details and hardly experienced patient problem-solving skills. For example, after knowing that the shunt effect really existed, they were unable to find out why the ventilation was less than perfusion or why the affected part had a marked increase in perfusion while the ventilation was normal. This might be due in part to a dramatic decrease in volume of respiration, pulmonary aspiration, atelectasis, ventilator setting, or heavy sedation on the central nervous system. Besides, over-hydration of crystalloids and colloids or over-transfusion of blood and blood components owing to poor patient management, as well as malpractice in physical examination and a misinterpretation of additional laboratory data, resulting in cardiovascular instability and failure.

There were no statistically significant difference between the average test scores of the first-year residents, preclinical medical and nurse anesthetist students. The developed diagnostic test was sufficient to assess misconceptions of students, particularly those of the same professional line but with different background knowledge.

First-year residents' achievement scores of 70% revealed that the misconceptions of 30% were due to the lack of knowledge for clinical interpretation and the inability to apply the logical results as a guideline for patient management. These problems were not unexpected because ABGs analysis, for which students needed to do their own self-study, represented a course outside the present medical curriculum (i.e., it was a hidden curriculum). Therefore most residents lacked the experience of such analysis, which arises from patient care in the operative theatre, recovery room or intensive care unit Compare this to the 60% achievement scores earned by medical and nurse anesthetist students. Their misconceptions of 40% in this group, which were similar to residents' misconceptions,

were due to the lack of enough knowledge to clarify the significance and relationship of the symbols. It was anticipated that these problems would have many different reasons. Though nurse anesthetist students graduated as registered nurses before attending a one-year specific training in anesthesiology and had work experience in some nursing care units for years, they did not comprehend ABGs interpretation and other laboratory information. Alternatively, medical students in the preclinical year never had any know-how of health care medicine except for some basic premedical knowledge that allowed them to achieve comparable scores with those of nurse anesthetist students.

The diversity and complexity of current medical information require students to constantly strive for self-learning, particularly certain subjects, not formally included in regular class hours and commonly referred to as a hidden curriculum. These subjects are increasing in degree. For further study, the online development of diagnostic test and feedback-providing mechanism seems to be highly practical to help professional students master learning. However, it needs to be well-developed, challenging, and stimulating to the curiosity of students, as well as promptly interacting with other databases. The subjects should be of their current interest, the exercises must have a complete scope of content and quality, and they must be highly independent in self-learning.

Conclusion

Using a concept and knowledge map to define the table of specifications in ABGs analysis together with open discussion and feedback-providing method, helped facilitate the scope of developing a more practical two-tier diagnostic test. Teachers can assess misconceptions of students with different background knowledge in a short time and have guidelines to improve pedagogy in response to their eagerness for learning.

References

Bell, B. & Cowie, B. (2001) The characteristics of formative assessment in science education, *Science Education,* 85, pp.536-53.

Ciofalo, J.F. & Wylie, E.C. (2006) Using diagnostic classroom assessment: one question at a time, [online] available at: http://www.tcrecord.org/printcontent.asp?contentID=12285 [accessed November 1, 2010].

Chittleborough, G.D., Mocerino, M. & Treagust, D.F. (2007) Achieving greater feedback and flexibility using online pre-laboratory exercises with non-major chemistry students, *Journal of Chemical Education,* 84,5, pp.884-8.

Duit, R., Treagust, D.F. & Mansfield, H. (1996) Investigating student understanding as a prerequisite to improving teaching and learning in science and mathematics, In: Treagust, D.F., Duit, R. & Fraser, B.J. (eds.) *Improving teaching and learning in science and mathematics*, New York: Teachers College Press, pp.1-14.

Heh, J.-S., Li, S.-C., Chang, A., Chang, M., & Liu, T.-C. (2008) Diagnosis mechanism and feedback system to accomplish the full-loop learning architecture, *Educational Technology & Society,* 11, 1, pp.29-44.

Novak, J.D. & Gowin, D.B. (1984) *Learning how to learn*, New York: Cambridge University Press.

Novak, J.D. (1981) Applying learning psychology and philosophy of science to biology teaching, *The American Biology Teacher*, 73, pp.12-20.

Novak, J.D. (1996) Concept mapping: a tool for improving science teaching and learning, In: Treagust, D.F., Duit, R. & Fraser, B.J. (eds.) *Improving teaching and learning in science and mathematics,* New York: Teachers College Press, pp.32-43.

Tan, K.-C.D., Goh, N.-K., Chia, L.-S., and Treagust, D.F., (2003) Students's understanding of acid, base and salt reactions in qualitative analysis. *School Science Review,* March 84, 308, pp.89-97.

Treagust, D.F. (1985) Diagnostic tests to evaluate students' misconceptions in science, *Paper presented at the annual meeting of the National Association for Research in Science Teaching*, Indiana: French Lick Springs.

Treagust, D.F. & Haslam, F. (1987) Diagnosing secondary students' misconceptions of photosynthesis and respiration in plants using a two-tier multiple choice instrument, *Journal of Biological Education*, 21, 3, pp.203-11.

Treagust, D.F. (1988) Development and use of diagnostic tests to evaluate students' misconceptions in science, *International Journal of Science Education*, 10, 2, pp.159-69.

Turns, J.E., Atman, C.J. & Adams, E. (2000) Concept maps for engineering education: A cognitively motivated tool supporting varied assessment functions, *IEEE Transactions on Education,* 43, 2, pp.164–73.