Conceptions of Assessment of Mainland China College Lecturers: A Technical Paper Analyzing the Chinese Version of COA-III

Li Wai Shing and Hui Sammy King Fai
The Hong Kong Institute of Education, Hong Kong

Assessment has a salient influence on student learning. With the introduction of the principle of "assessment for learning," assessment practice in different education sectors, and professional and vocational education in particular, is about to change, to adopt a more formative and authentic type of assessment. However, the change is still foreign to most of the Mainland China educators who place much emphasis on getting students to achieve good examination results. Grading and selection are two of the major functions of assessment, and those of diagnosis and guidance are developing. Therefore, before offering any concrete ideas of how assessment for learning could work in a Chinese context with its examination-oriented culture, exploring teachers' conceptions regarding assessment is meaningful. Results of a survey of 97 college lecturers of a vocational and technical institute in Hangzhou, Mainland China (Hangzhou Wanxiang Polytechnic, 杭州萬向職業技術學院) suggested their agreement that assessment improves quality of teaching and student learning and also makes schools more accountable. In contrast, the lecturers doubted whether assessment could provide valid information concerning deep learning in contrast to passing examinations. The more they agreed that assessment improves quality of teaching and student learning, the more they found it makes schools accountable; and the more they agreed that assessment provides valid information and describes student learning, the more they found it makes students accountable. The backwash effect of examination was apparent to them. Second-order factor analysis further suggested that whether they thought assessment improves quality of teaching and student learning was a different construct from whether they found that assessment provides valid information, describes student learning, and makes students and schools accountable. The paper hopes to contribute to gaining an understanding of how effective learning and assessment can be facilitated in the professional and vocational education context in Mainland China.

Authentic assessment is crucial to student learning: "[t]o the teacher, assessment is at the end of the teaching-learning sequence of events, but to the students it is at the beginning" (Biggs, 1999, p. 141). Originating from the Latin word "assidere," which means "to sit beside," the aims of assessment are to facilitate a close relationship and a sharing of experience between teachers and students (Satterly, 1989).

There are weaknesses in the current assessment system (Black & Wiliam, 1998). Tests and examinations are so overwhelming that rote and

superficial learning are over-emphasized. Tests alone do not guarantee the promotion of learning as they too often stress quantitative aspects of learning. Quantitative grading and marking are highlighted while offering descriptive feedback is usually neglected. Students are compared with one another and the prime purpose of assessment appears to emphasize competition rather than learning for its own sake. The message conveyed to students with low attainment is that "they lack ability", and therefore they look for ways to obtain higher marks not necessarily knowing how to learn better and how to demonstrate this with improved results. This is precisely the situation in Mainland China, where examinations are compulsory for all courses offered in professional and vocational education institutes. Public examinations are seen as important as they supposedly preserve the professional standards of the profession. As reported in the *Chicago Tribune*: "Chinese people have a tradition of changing their lives through examinations [...] For many, it has been the only way of changing their fate." (Dorgan, 2000, p. 15)

The purpose of this study is to explore if the belief that "examination determines fate" is entertained by college teachers. In particular this study examines the conceptions of assessment of a group of Mainland China college lecturers through administering a Chinese version of a well-validated instrument, namely *Conceptions of Assessment Questionnaire*, COA-III (Brown, 2003 & 2004). The study aims to contribute to teacher education, in professional and vocational education contexts, by providing an understanding of teachers' thinking with respect to assessment. It is anticipated that the study's findings will assist educators to develop more authentic assessment.

CONCEPTIONS AND PRACTICES OF ASSESSMENT

Research concludes that conceptions of assessment influence how teachers instruct their students (d'Ydewalle, 2000). Conceptions, as a more general mental structure, encompass beliefs, meanings, concepts, preferences, and the like (Thompson, 1992). The most typical examples are the conceptions of teaching. As commented by Kember (1997), after reviewing 13 of the foremost research studies into university academics' conceptions of teaching published from 1983 to 1994, "[t]eaching conceptions have been shown to be related to measures of the quality of student learning, so are modeled as influencing teaching approaches which in turn effect student learning approaches and learning outcomes. [...] teaching approaches are strongly influenced by the underlying beliefs of the teacher" (Kember, 1997, p. 255). How teachers conceptualize teaching influences their practice of teaching. This also applies to conceptions of assessment (Tittle, 1994; Borke, Mayfield, Marion, Flexer & Cumbo, 1997; Brown, 2003 & 2004).

To illustrate this point, it is appropriate to refer to Watkins's (1998) research into the assessment of university students in Hong Kong. He concluded "the majority of respondents (151 Hong Kong university academics) felt that they were the ones making assessment decisions about courses they were teaching" (Watkins, 1998, p. 14). The assessment methods that these academics chose (over half of them reported using individual assignments, essay examination questions, group assignments, short answer questions, and tutorial participation) are deep-rooted in their beliefs that "tertiary education should achieve higher order learning outcomes such as critical thinking, self-directed learning, and the ability to apply knowledge to novel situations" (Watkins, 1998, p. 16). However, Watkins (1988) further points out that, although the link of conceptions to practices is strong, it is still not enough for university teachers to desire to influence student learning by rewarding high order learning outcomes. This is because "[o]ver half of those interviewed claimed to desire such an outcome but felt that they were unable to achieve it" (Watkins, 1998, p. 17). This is certainly the difficulty underlying assessment practice in professional and vocational education: "the core of competency-based assessment has always been professional and vocational education and training" (Wolf, 1995, p. 31). However, internationally,

"there has been a move towards formal, detailed codification of both syllabus and assessment procedures, which operate at national (or occasionally cantonal/state) level [...] in every case – including apprenticeship – important parts of the training process take place in formal classrooms and workshops, away from the workplace" (Wolf, 1995, p. 34). Such a "move" is caused not only by deficiencies of the students and institutional factors (the assessment system), but also teachers' beliefs and judgments of what counts as valid assessment. In the light of this, before making any claims of what teachers assess and how they assess, it is necessary to explore the conceptions which they have on assessment, specifically in the professional and vocational education context in Mainland China.

Conceptions of Assessment Questionnaire (COA-III)

In measuring conceptions of assessment, the work of Brown (2003 & 2004) is helpful. Assessment is defined as "any act of interpreting information about student performance, collected through any of a multitude of means" (Brown, 2003, p. 3). Traditionally, there are three purposes of assessment: the improvement of teaching and learning, certification of student learning, and accountability of schools and teachers (Torrance & Pryor, 1998). In addition to this, Brown (2003) offers a fourth purpose: the treatment of assessment as irrelevant to the life and work of teachers and students. In his terms:

> The major premise of the improvement conception is that assessment improves students' own learning and the quality of teaching [...] This improvement has two important caveats; (a) assessment must describe or diagnose the nature of student performance and (b) the information must be a valid, reliable, and accurate description of student performance. [...] A second conception of assessment is that assessment can be used to account for a teacher's, a school's, or even a system's use of society's resources [...] The premise of the third conception of assessment is that students are held individually accountable for their learning through assessment. [...] The premise of the final conception is that assessment, usually understood as a formal, organized process of evaluating student performance, has no legitimate place within teaching and learning [...] Assessment may be rejected also because of its pernicious effects on teacher autonomy and professionalism and its distractive power from the real purpose of teaching (i.e., student learning) [...] It may also be that the degree of inaccuracy (e.g., standard error of measurement) published with any formal measurement contributes to teachers' conception of assessment as irrelevant. (Brown, 2004, pp. 304-305)

Brown (2004) further argues that the various conceptions might interact with each other and that these conceptions can lead to different practices, which are often in tension with the original purposes. Table 1 summarizes the structure of this Conceptions of Assessment Questionnaire, COA-III.

Methods of Assessment for Mainland China Professional and Vocational Education

Mainland China is the homeland of examinations. She was the first country to select her civil servants according to the results of scholastic achievement and has done so for more than one thousand years (Dore, 1976; Ma, 1993). China is still depending on public examinations for evaluative and selective purposes. The steep pyramid education system in Mainland China renders successful students in these competitive examinations promotion opportunities from one education level to another, culminating with University entrance (Wang, 1993).

The drawbacks of this type of examination system are obvious as two officials in the Ministry of Education in China reflect that:

Table 1.
Structure of the 50-item COA-III

Dimensions	Number of items	Sample item
1. *Improvement of teaching and learning:*		
• Improvement: Describe (Assessment describes student learning.)	6	Assessment is a way to determine how much students have learned from teaching.
• Improvement: Student learning (Assessment improves student learning.)	7	Assessment provides feedback to students about their performance.
• Improvement: Teaching (Assessment improves the quality of teaching.)	6	Assessment is integrated with teaching practice.
• Improvement: Valid (Assessment provides valid information.)	5	Assessment results are trustworthy.
2. *Certification of students' learning:*		
• Student accountability (Assessment makes students accountable.)	7	Assessment is assigning a grade or level to student work.
3. *Accountability of schools and teachers:*		
• School accountability (Assessment makes schools accountable.)	6	Assessment provides information on how well schools are doing.
4. *Treatment of assessment as irrelevant to the life and work of teachers and students:*		
• Irrelevance: Bad (Assessment is bad for students and teachers.)	5	Assessment forces teachers to teach in a way against their beliefs.
• Irrelevance: Ignore (Assessment is used but ignored.)	5	Teachers conduct assessments but make little use of the results.
• Irrelevance: Inaccurate (Assessment is inaccurate.)	3	Assessment results should be treated cautiously because of measurement error.

> Schools tend to focus on how to help students pass the selective examinations and how to increase promotion rates. Teachers pay closest attention to the "best" students, those that have the best prospect of entering university, while tending to neglect the students with lower test achievement. (Han & Yang, 2001, p. 8)

Unfortunately there are serious negative repercussions concerning student learning:

> This resulted in multitudinous examinations and tests in schools and also increased the learning burden on students. Since the results of examinations were linked to teachers' performance (as an indicator for assessing teachers' teaching quality), teachers emphasized teaching based on examination-oriented education. (Han & Yang, 2001, p. 7)

This conclusion was identified in a study on students' satisfaction towards teaching and learning

in Tsinghua University. Over half of the students indicated that they were dissatisfied with the teaching and learning. Over three-quarters of the students believed they were learning little during their undergraduate years (Ren, 2001). Ren (2001) goes on to comment that methods of assessment in higher education need to be improved, with more attention focusing upon those teaching and assessment methods that enhance student learning. The research likewise identified that traditional methods of assessment, objective tests, comprehension questions and memorization prevailed from Junior Secondary Entrance Examination to final year in University. In addition, research in China concerning current practices of assessment in higher education concluded that assessment methods are monotonous and textbook oriented (He & Chen, 2003). Moreover, there was a lack of formative assessment. They recommended a more authentic assessment process that would also enhance personal development, stimulate practical ability, and nurture critical and creative thinking skills as well as assisting students to take active responsibility in their own learning.

If Mainland China is to face contemporary education challenges, then current assessment regimes in professional and vocational education, in particular, require comprehensive reform. The continuation of an education system driven by the exclusive use of external examinations reinforces rote learning through the memorization of book knowledge. Unfortunately, many teachers are content with conservative methods of assessment because they know that good results from rote teaching enhance their image. Consequently, teachers are not pioneers in alternative strategies of assessment. They believe that their use of new models of assessment would invite educational authorities, parents and school heads to negatively assess the quality of their teaching (Gao, Du & Yu, 2006). These factors contribute to teachers' complex and contradictory conceptions of assessment. On the one hand, teachers know that the present practices are detrimental to their students' learning, but on the other hand, the cost to bring about innovative assessment strategies is too great because both the teachers and the students cannot afford to perform poorly in competitive scholastic achievement tests which emphasize rote learning. Despite this harsh reality, efforts to improve assessment in higher education are a priority in China (Ma, 1993; Wang, 1996; Han & Yang, 2001). The first step to achieve this goal is to change teachers' conceptions of assessment.

METHODOLOGY

A survey research method was used for this study (Fink, 1995; Munn & Drever, 1999). It was selected for three reasons. First, it allowed access to a comparatively large sample of cases within a short period of time. Second, the collection of information was generally anonymous and a high return rate was possible. Third, the use of standardized questionnaires made comparison of information possible.

A Chinese version of the 50-item Conceptions of Assessment Questionnaire, COA-III, was developed.

Brown's (2003 & 2004) 50-item COA-III was used to measure and investigate the conceptions that teachers hold on assessment. The 50 items were translated and back translated, from English to Chinese and from Chinese to English, until the Chinese wording achieved the closest match to the original English meaning. The first translated version of the instrument was piloted with 20 in-service teachers in the professional and vocational education field. These teachers were part-time Postgraduate Diploma in Education (Professional and Vocational Education) first year students of the Hong Kong Institute of Education.[1] The pilot was undertaken during the time that they were taking the module "Assessment," which was designed to provide them with the knowledge and skills needed to be critical and reflective about their assessment practices and to explore alternatives. The module also provided opportunities for participants to experience a range of assessment and feedback strategies suitable for use in

professional and vocational education. All 50 items were examined and any items that were not consistent with the scale, did not seem valid, and had the least discriminating power, were modified. There were three criteria: (i) items whose means were close to the extremes of the scale; (ii) items whose corrected item-total correlations were less than 0.30; and (iii) items whose removal increased the alpha values. Finally, a 50-item Chinese version of COA-III was developed.

The respondents were asked to indicate, using a 5-point Likert scale, how much they agreed with each of the 50 items of COA-III. The possible responses ranged from "strongly disagree" through "disagree", "no comment", "agree" to "strongly agree", with numerical values of 1 to 5 assigned for purposes of later analysis.

The questionnaire was administered to all lecturers (103 in total) of Hangzhou Wanxiang Polytechnic (杭州萬向職業技術學院 http://www.wxpoly.cn/), a vocational and technical institute in Hangzhou, Mainland China. These college lecturers traveled to Hong Kong in July 2006 to participate in an intensive 3-week summer training program, jointly organized by The Hong Kong Institute of Education and The Hong Kong Polytechnic University. This training program was designed to provide participants with up-to-date theories and practices for effective teaching and learning. There were altogether seventeen 3-hour modules, and the second author was responsible for two of these: "Instructional Design" and "Classroom Interaction Strategies." Group discussions were used in the major teaching and learning activities and participants were assessed at the end of each module, either through portfolios or reflective journals. Fieldwork took place during the period when the second author first taught the module "Instructional Design." In total, 97 questionnaires were returned which represented a response rate of 94.2%.

Table 2.
Means (and standard deviations) of the nine dimensions of COA-III

The nine dimensions of COA-III	Mean (SD)
1. *Improvement of teaching and learning:*	
· Improvement: Describe(Assessment describes student learning.)	2.86 (0.54)
· Improvement: Student learning(Assessment improves student learning.)	3.40 (0.53)
· Improvement: Teaching(Assessment improves the quality of teaching.)	3.88 (0.52)
· Improvement: Valid(Assessment provides valid information.)	2.73 (0.62)
2. *Certification of students' learning:*	
· Student accountability(Assessment makes students accountable.)	3.06 (0.51)
3. *Accountability of schools and teachers:*	
· School accountability(Assessment makes schools accountable.)	3.62 (0.56)
4. *Treatment of assessment as irrelevant to the life and work of teachers and students:*	
· Irrelevance: Bad(Assessment is bad for students and teachers.)	2.15 (0.57)
· Irrelevance: Ignore(Assessment is used but ignored.)	2.42 (0.55)
· Irrelevance: Inaccurate(Assessment is inaccurate.)	3.49 (0.67)

Note: When computing the mean, 1 = strongly disagree, 2 = disagree, 3 = no comment, 4 = agree, 5 = strongly agree.

Data was entered into the software Statistical Package for Social Sciences (SPSS) and Analysis of Moment Structures (Amos) for analysis and different techniques were employed (Arbuckle, 2003; Arbuckle & Wothke, 1999; Bryman & Cramer, 1997; Norušis, 2000). First, reliability analysis was used to measure the internal consistency of each of the nine dimensions. To show the distributions and variations of COA-III, mean and standard deviations were reported. Pearson product-moment correlations were used to measure the degree of association between the dimensions. Finally, second-order exploratory and confirmatory factor analyses were used to identify the higher order structure of COA-III (Gorsuch, 1983; Beauducel, 1997). This was important given that it was necessary to identify the pattern of thoughts embedded in the minds of the professional and vocational education Chinese teachers towards the different dimensions of conceptions measured, and thus to investigate further the illuminating feature of the scale.

RESULTS AND DISCUSSION

Mainland China College Lecturers' Conceptions of Assessment

Reliability analysis was first run for the nine dimensions of COA-III. All the dimensions were found to be internally consistent (Cronbach Alphas ranged from 0.596 to 0.760), and mean scores were computed. Distributions and variations of these mean scores are summarized in Table 2.

In general, the 97 college lecturers in this sample agreed that assessment: (i) improves quality of teaching and student learning (mean = 3.88 & 3.40); (ii) makes schools accountable (mean = 3.62); (iii) is not bad for students or teachers (mean = 2.15); and (iv) is not ignored (mean = 2.42). On the contrary, they did not agree that assessment: (i) provides valid information and describes student learning (mean = 2.73 & 2.86) and (ii) is accurate (mean = 3.49).

Pearson product-moment correlations were used to measure the degree of association between the dimensions. After careful examination, some of the correlations were found to be meaningful and illuminating, and they are reported in Table 3.

There were "moderate" correlations ($10\% < r^2 < 40\%$) between the dimensions of "assessment improves the quality of teaching and student learning" and "assessment makes schools accountable" (r = 0.373 & 0.358) and between the dimensions of "assessment provides valid information and describes student learning" and "assessment makes students accountable" (r = 0.408 & 0.416). However, there were none at all or only "weak" correlations between the 3 dimensions of "whether treatment of assessment is irrelevant to the life and work of teachers and students" and the 4 dimensions of "whether assessment is for improvement of teaching and learning" ($r^2 < 10\%$).

Second-Order Factor Analysis of the Nine Dimensions of COA-III

To identify the second-order factors of the nine dimensions of COA-III, exploratory factor analysis (EFA) was used with principal component analysis as the method for factor extraction, followed by oblique rotation. Three factors were extracted which explained 62.4% of the total variance. The use of this factor model was supported by different statistical tests, for example, a high value of KMO measure of sampling adequacy (0.631) indicated the current analysis was "meritorious" (Kaiser, 1974). Also, a large value of Bartlett's test for sphericity (135.12) rejected the hypothesis that the population correlation matrix was an identity (associated level of significance p = 0.000).

The first factor consisted of five dimensions, which were "assessment provides valid information," "assessment describes student learning," "assessment makes students accountable," "assessment improves student learning" and "assessment makes schools accountable." The second factor consisted of two dimensions, which were "assessment is bad for students" and "assessment is used but ignored." The third factor consisted of two dimensions, which were "assessment is inaccurate" and "assessment

Table 3.
Pearson product-moment correlations of the dimensions of COA-III

	Student accountability (Assessment makes students accountable.)	School accountability (Assessment makes schools accountable.)
Improvement: Describe (Assessment describes student learning.)	0.416**	0.216*
Improvement: Student learning (Assessment improves student learning.)	0.296**	0.358**
Improvement: Teaching (Assessment improves the quality of teaching.)	0.150	0.373**
Improvement: Valid (Assessment provides valid information.)	0.408**	0.302**

	Irrelevance: Bad (Assessment is bad for students and teachers.)	Irrelevance: Ignore (Assessment is used but ignored.)	Irrelevance: Inaccurate (Assessment is inaccurate.)
Improvement: Describe (Assessment describes student learning.)	0.197	- 0.133	- 0.108
Improvement: Student learning (Assessment improves student learning.)	0.069	- 0.252*	0.024
Improvement: Teaching (Assessment improves the quality of teaching.)	- 0.069	- 0.259*	0.192
Improvement: Valid (Assessment provides valid information.)	- 0.032	- 0.005	- 0.200

** Correlation is significant at 0.01 level ($p < 0.01$).
* Correlation is significant at 0.05 level ($p < 0.05$).

improves the quality of teaching". All dimensions had a high value of rotated factor loading (from 0.614 to 0.764, from 0.685 to 0.810, and from 0.518 to 0.850 respectively). The hypothesized pattern of factor loadings was tested further under confirmatory factor analysis (CFA) with maximum-likelihood as the method for factor extraction, followed by oblique rotation. The same factor structure as in the EFA was extracted, and the model was a good fit given that it was insignificant under the Goodness-of-fit test (Chi-square = 13.85, df = 12, p = 0.310). Table 4 shows the rotated factor loadings for the nine dimensions.

However, additional measures of fit in Amos suggested that the model was not a good fit. These

include a high Chi-square to df ratio (CMIN/df = 2.606, CMIN = 78.192, df = 30), unacceptable comparative fit indices (CFI = 0.610; PCFI = 0.406), a large value of Root Mean Square Error of Approximation (RMSEA = 0.129) with a small probability for the testing of the null hypothesis that RMSEA is no greater than 0.05 (PCLOSE = 0.000). All these indicated that the hypothesized factor model did not account well for the observed covariances in the data. In other words, it failed to accurately reproduce the sample correlational data.

This 3-factor model structure indicated one interesting phenomenon that the dimension "assessment improves the quality of teaching" loaded quite heavily on all three factors (larger than ± 0.40). Its high rotated factor loadings on the first two factors were reasonable: a conception that assessment improves quality of teaching on the one hand varied directly with other "proper" conceptions (like "assessment describes student learning", "assessment improves student learning", etc.) and on the other hand varied in an opposite direction with two other "negative" conceptions ("assessment is bad for students" and "assessment is used but ignored"). However, the presence of the third factor, to which this dimension and the dimension "assessment is inaccurate" loaded most heavily on, and varied in the same direction, suggested an important lesson. The underlying thoughts and conditions of the Mainland China college lecturers in this sample which revealed a belief that assessment is for improving quality of teaching, were complex and obviously affected by

Table 4.
Second-order rotated factor loadings of the nine dimensions of COA-III

The nine dimensions of COA-III	Factor 1	Factor 2	Factor 3
Improvement: Valid (Assessment provides valid information.)	0.764		
Improvement: Describe (Assessment describes student learning.)	0.731		
Student accountability (Assessment makes students accountable.)	0.719		
Improvement: Student learning (Assessment improves student learning.)	0.642		
School accountability (Assessment makes schools accountable.)	0.614		
Irrelevance: Bad (Assessment is bad for students and teachers.)		0.810	
Irrelevance: Ignore (Assessment is used but ignored.)		0.685	
Irrelevance: Inaccurate (Assessment is inaccurate)			0.850
Improvement: Teaching (Assessment improves the quality of teaching.)	0.440	- 0.465	0.518
Eigenvalue (% of variance explained)	2.9 (32.1)	1.5 (16.8)	1.2 (13.5)

Note: Only values of 0.40 or above are shown in the table.

Table 5.
Second-order rotated factor loadings of the six dimensions of COA-III

The six dimensions of COA-III	Factor 1	Factor 2
Student accountability(Assessment makes students accountable.)	0.858	
Improvement: Describe(Assessment describes student learning.)	0.788	
Improvement: Valid(Assessment provides valid information.)	0.634	
School accountability(Assessment makes schools accountable.)	0.445	
Improvement: Teaching(Assessment improves the quality of teaching.)		- 0.951
Improvement: Student learning(Assessment improves student learning.)		- 0.782
Eigenvalue (% of variance explained)	2.8 (46.9)	1.0 (17.1)

Note: Only values of 0.40 or above are shown in the table.

many concerns. Among those, the doubt they had about the accuracy of examinations was critical.

Accordingly, there were further alternatives to explore the second-order factor structure of COA-III. When only the six "proper" dimensions were run, the extracted factor structure was valid and defensible under different measures of fit. Two factors were extracted, which explained 64.0% of the total variance. A high value of KMO measure of sampling adequacy (0.762) indicated the current analysis was "meritorious" to "marvelous" (Kaiser, 1974), and a large value of Bartlett's test for sphericity (96.49) rejected the hypothesis that the population correlation matrix was an identity (associated level of significance p = 0.000). The first factor consisted of four dimensions, which were "assessment makes students accountable," "assessment describes student learning," "assessment provides valid information" and "assessment makes schools accountable." The second factor consisted of two dimensions, which were "assessment improves the quality of teaching" and "assessment improves student learning". All dimensions had a high value of rotated factor loading (from 0.445 to 0.858 and from - 0.782 to - 0.951 respectively). Table 5 shows the rotated factor loadings for these six dimensions.

CFA extracted the same factor structure as in EFA, and the model was a good fit (Chi-square = 1.52, df = 4, p = 0.824). Additional measures of fit in Amos also suggested that the model was a good fit. These include a low Chi-square to df ratio (CMIN/df = 1.045, CMIN = 12.540, df = 12), acceptable comparative fit indices (CFI = 0.994; PCFI = 0.568), a small value of Root Mean Square Error of Approximation (RMSEA = 0.022) with a large probability for the testing of the null hypothesis that RMSEA is no greater than 0.05 (PCLOSE = 0.611). All these indicated that the hypothesized factor model did account well for the observed covariances in the data and could accurately reproduce the sample correlational data.

Validity of Assessment System of Mainland China Professional and Vocational Education

Results of the distributions and variations of mean scores of the nine dimensions of COA-III indicated that the college lecturers in this sample were holding very contradictory conceptions of assessment. On the one hand, they agreed that assessment improves quality of teaching and student learning and makes schools accountable, assessment is not bad for students or teachers, and assessment is not ignored. On the other hand, they

found assessment could not provide valid information or describe student learning and, most astonishingly, assessment is inaccurate. This is certainly the reality of the assessment system of Mainland China which puts too much emphasis on examination. Generally speaking, in order to guarantee the professional standard of practitioners in the workplace, curriculum planners and developers have to make sure that the programs/courses they are offering do equip students with the necessary knowledge and skills required in the field. To remove all unwanted bias and subjective judgments over the performance of students, examination, the so-called "objective" assessment system, should therefore be adopted. Pass rates and public examination results become an important criterion of judging the professional status and accountability of an institute. Frontline college lecturers need to ensure that the curriculum content of the modules they are teaching match the standards and examination syllabuses issued by relevant government bodies.[2] Whether or not students do well in examination therefore indicates the ability of teachers to transmit the required knowledge and skills to students. The quality of teaching and student learning is also reflected in examination, specifically how "talented" the students are in reproducing such knowledge and skills. Given that examination is a part of life over which teachers and students have no control, although they doubt whether it can provide valid information and describe student learning, they rationalize it as good, as the motivator for improvement. The moderate correlations between "assessment improves the quality of teaching and student learning" and "assessment makes schools accountable" further indicate the influence of examination on teaching and learning and school accountability.

However, being recognized as academics who are influential to the development of professional and vocational education in Mainland China, the college lecturers in this sample question the accuracy of examination as the main method of assessment. They also question the theory that examination provides more valid information about student learning and makes the students more accountable, specifically when they graduate and join the workforce. Improving the validity of the assessment system of Mainland China professional and vocational education is critical. Validity of assessment is defined as the extent to which an assessment measures what it claims or purports to assess (Garrett, 1966; Zeller, 1990). Certainly, the validity of examination is in doubt because it is "summative" in nature, and the "content" of this system does not adequately and convincingly indicate how much students have learned. Examination results therefore are not at all trustworthy. It is even worse when the purpose of examination is not clear, the grading criteria and procedures are not relevant or transparent, and constructive and timely feedback is not provided. As reminded by Wiliam (1993), content validity should not only be concerned with the test items but also with the answers elicited. Therefore, it is the performance which students are able to demonstrate in the workplace that counts as valid. This is precisely the idea of performance assessment, which is defined as a measure of assessment that is based on authentic tasks such as activities, exercises, or problems that require students to show what they can do (McBrien & Brandt, 1997, pp. 78-79). On a conceptual level, assessment should be "formative" in nature, aiming to see how well the students are learning, identifying their strengths and weaknesses, and the problems that they may be encountering. The "context" in which assessment is placed should be authentic and should concentrate on the demonstration and application of functional knowledge and skills in the real world. On a practical level, the performance to be assessed should be "qualitative" in nature, requiring students to apply, synthesize, evaluate and even create from the given knowledge and skills. Learning is not static but dynamic in nature, and therefore the performance that assessment elicits should reflect how students make good sense of and go beyond given knowledge and skills (see, for example, the qualitative learning outcomes in Biggs' (1999) SOLO Taxonomy). The assessment procedures and the setting and communication of assessment criteria should "involve" the students more. The

assessment system of Mainland China's professional and vocational education must link up the performance that is of concern to students in the real world, and unless a more valid system is being recognized, any claim of accountability of students will be difficult.

Conditions for Assessment to Improve Quality of Teaching and Student Learning

The extracted 2-factor model structure for the six "proper" conceptions is informative to the understanding of the distinct pattern of thoughts that the Mainland China college lecturers in this sample have towards assessment. It is evident that whether they thought assessment is for improving quality of teaching and student learning was a very different conception from whether they found assessment could provide valid information, describe student learning, and make students and schools accountable. In other words, there was a clear distinction between a "functional" understanding of assessment for teaching and learning and an "evaluative" understanding of assessment. Following the previous analysis of how college lecturers understand and rationalize examination as a motivator for improvement, the term "functional" refers to the situation in which assessment urges teachers to train their students better in order to enhance performance in examinations. This is very different from the common perception that assessment should aim to identify students' strengths and weaknesses, and thus facilitate student learning. The term "evaluative" refers to the analysis of how well the assessment results reveal what they intend to measure. College teachers in the professional and vocational education field in Mainland China tended to weigh these two forms of understanding on a different basis. That is to say, the evaluation of how well the assessment results revealed what students had learned and what schools had achieved, did not necessarily result in better student performance or mean they knew their strengths and weaknesses.

Thus, apart from the last suggestion of having a more valid assessment system which focuses on performance and outcomes of student learning, teachers in the field should concentrate more on the assessment process to nurture the conditions for assessment in order to improve the quality of teaching and student learning. This certainly matches with the trend of "assessment for learning," which refers to the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there (Assessment Reform Group, 2002). Among the ten known principles, the following are significant and should draw more attention to: (i) assessment for learning being part of effective planning of teaching and learning; (ii) assessment for learning focusing on how students learn; and (iii) assessment for learning being regarded as a key professional skill for teachers.

Looking in more detail at the first principle, teachers should be sensitive to their planning of teaching and learning activities and assessment tasks, which envisage the opportunities for both students and teachers to obtain useful information about progress towards learning goals. Authentic tasks, which require students to demonstrate what they have learned cognitively and practically, as aligned with course objectives and teaching and learning activities, serve as validation. With the second principle, teachers should communicate with students their learning evidence and make them aware of not only "what" they have learned but also "how" they are learning. Portfolio – "log-book entry" of one's construction of personal journal items – is useful, because it gives a clear picture of one's development over time, monitors one's own thinking, and reflects on what one has considered and learned (Biggs & Tang, 1998). With the third principle, teachers should be equipped with the knowledge and skills to plan for assessment, observe learning, analyze and interpret evidence of learning, give feedback to students, and support students in self-assessment. Continuous professional development is therefore necessary.

CONCLUSION

Amid the contradictory conceptions of assessment among teachers in the study, they do

understand that assessment methods have to be changed if the quality of learning and teaching is to be enhanced. However, there are still some hindrances for this change to be fully realized. In summary, on the basis of the evidence up to this stage, two findings are constructive to effective learning and assessment in the professional and vocational education context in Mainland China. These are (i) an urge to change the belief and judgment of what counts as valid assessment and (ii) the development of an "assessment for learning" culture through professional development and training. While both academics and officials in the Ministry of Education are still trying to nurture a positive environment for the changes envisioned, teacher educators must be in the frontline to bring about new initiatives to the assessment system and to show that the idea of assessment for learning is the key for quality education throughout the whole education system.

NOTES

[1]Information about this 2-year part-time Postgraduate Diploma in Education (Professional and Vocational Education) can be retrieved from: http://www.ied.edu.hk/acadprog/postgrad/prog/pgde_prog.htm#3.

[2] This is reflected in the assessment task for the module "Instructional Design" which the second author taught for the intensive 3-week summer training program, the "Curriculum Evaluation" journal. Almost all of these college lecturers indicated that the objectives and contents of the modules they taught corresponded to government standards, and the quality was assured by examination.

REFERENCES

Arbuckle, L.J. (2003). Amos 5.0 update to the Amos user's guide. Chicago, IL: Marketing Department, SPSS Inc.: SmallWaters Corporation.

Arbuckle, L.J., & Wothke, W. (1999). Amos 4.0 user's guide. Chicago, IL: Marketing Department, SPSS Inc.: SmallWaters Corporation.

Assessment Reform Group (2002). *Assessment for learning: 10 principles*. Assessment Reform Group. Retrieved August 30, 2007, from http://teaching-resource.co.uk/teachers/afl.htm.

Beauducel, A. (1997). Transformation-matrix-Search and -Identification (Trasid): A new method for oblique rotation to simple structure. *Methods of Psychological Research Online*, *2*(2). Retrieved Jan 11, 2007, from http://www.mpr-online.de//issue3/art10/article.html.

Biggs, J. (1999). *Teaching for quality learning at university*. Buckingham, England: Society for Research into Higher Education & Open University Press.

Biggs, J., & Tang, C. (1998). Assessment by portfolio: Constructing learning and designing teaching. In P. Stimpson & P. Morris (Eds.), *Curriculum and assessment for Hong Kong: Two components, one system* (pp. 443-462). Hong Kong: Open University of Hong Kong Press.

Black, P., & Wiliam, D. (1998). Inside the black box. *Del Delta Kappan*, *80*(2), 139-148.

Borke, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching & Teacher Education*, *13*(3), 259-278.

Brown, G.T.L. (2003, November). *Teachers' instructional conceptions: Assessment's relationship to learning, teaching, curriculum, and teacher efficacy*. Paper presented at the Joint Conference of the Australian and New Zealand Associations for Research in Education (AARE/NZARE). Auckland: University of Auckland.

Brown, G.T.L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy and Practice*, *11*(3), 301-318.

Bryman, A., & Cramer, D. (1997). *Quantitative data analysis with SPSS for windows: A guide for social scientists*. London: Routledge.

Dore, R.P. (1976). *The diploma disease: Education, qualifications and development.* London: Allen & Unwin.

Dorgan, M. (2000, July 9). In China, examinations determine fate. *Chicago Tribune*, Section 1, p. 15.

d'Ydewalle, G. (2000). Sensation/Perception, information processing, attention. In K. Pawlik & M.R. Rosenzweig (Eds.), *International handbook of psychology* (pp. 79-99). London: Sage.

Fink, A. (1995). *The survey kit, volume 1: The survey handbook*. London: Sage.

Gao, W., Du, Y.H., & Yu, X. (2006). Higher education examination system reform and the nurturing of innovative abilities. *Journal of Inner Mongolia Normal University*, *19*(5), 433-445.

Garrett, H.E. (1966). *Statistics in psychology and education*. London: Longman.

Gorsuch, R.L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Han, M., & Yang X.W. (2001). Educational assessment in China: Lessons from history and future prospects. *Assessment in Education: Principles, Policy and Practice*, *8*(1), 5-10.

He, Z., & Chen, W.J. (2003). Enhancing higher education examination reform. *Petroleum Education Bimonthly*, June, 39-41.

Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, *39*, 31-36.

Kember, D. (1997). A reconceptualisation of the research into university academics' conceptions of teaching. *Learning and Instruction, 7*(3), 255-275.

Ma, J.K. (1993). Strengthen the scientific research on examinations, revive the enterprise of examinations in China. *Chinese Education and Society*, *26*(5), 30-42.

McBrien, J.L., & Brandt, R.S. (1997). *The language of learning: A guide to education terms*. Alexandria, VA: Association for Supervision and Curriculum Development.

Munn, P., & Drever, E. (1999). *Using questionnaires in small-scale research: A teachers' guide*. Edinburgh, Scotland: Scottish Council for Research in Education.

Norušis, M.J. (2000). *SPSS 10.0: Guide to data analysis*. Upper Saddle River, NJ: Prentice Hall.

Ren, X.M. (2001). Teaching evaluation: A neglected aspect of higher education. *Journal of Yuxi Normal College*, *17*(5), 53-56.

Satterly, D. (1989). *Assessment in schools* (2nd ed.). Oxford: Blackwell.

Thompson, A.G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 127-146). New York: Macmillan.

Tittle, C.K. (1994). Toward an educational psychology of assessment for teaching and learning: Theories, contexts, and validation arguments. *Educational Psychologist*, *29*, 149-162.

Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Buckingham, England: Open University Press.

Wang, G. (1996). Educational assessment in China. *Assessment in Education: Principles, Policy and Practice*, *3*(1), 75-89.

Watkins, D. (1998). Assessing university students in Hong Kong: How and why? In D. Watkins, C. Tang, J. Biggs, & R. Kuisma (Eds.), *Assessment of university students in Hong Kong: How and why, assessment portfolio, students' grading* (pp. 5-27). Hong Kong, ESEP: City University of Hong Kong.

Wiliam, D. (1993). Validity, dependability and reliability in national curriculum assessment. *The Curriculum Journal*, *4*(3), 335-350.

Wolf, A. (1995). The case of competence-based assessment. In *Competence-based assessment* (pp. 30-52). Buckingham, England: Open University Press.

Zeller, R.A. (1990). Validity. In H.J. Walberg & G.D. Haertel (Eds.), *International encyclopedia of educational evaluation* (pp. 251-259). Oxford: Pergamon Press.

Request PDF | Conceptions of assessment of Mainland China college lecturers: A technical paper analyzing the Chinese version of COA-III | Assessment has a salient influence on student learning. With the introduction of the principle of assessment for learning," assessment practice in | Find, read and cite all the research you need on ResearchGate.Â In Chinese context, Shing and Fai (2007) applied a survey method to disclose 97 college lecturers' conceptions of assessment in Mainland China. The outcomes of the study indicated that lecturers showed agreement with the idea that assessment advances both the quality of teaching and students learning. Statistics show China has become the largest source of international students thanks to its growing pool of potential candidates. In total, 5,194,900 Chinese students have studied abroad over the last 40 years, and 1,454,100 students are currently enrolled in overseas higher education institutions. In 2017, while developed countries and regions, such as the US and Western Europe, remained the most popular destinations, Belt and Road (B&R) countries have begun to harvest the growing wave of Chinese students seeking to learn abroad. In total, 66,100 students, including 3,679 on government sp AbstractAs China continues to involve teachers in the implementation of an assessment for learning or formative assessment policy, a clearer understanding of how they conceive of the purposes and functions of assessment is necessary. This paper synthesises eight interview and survey studies, which have examined how diverse samples of practicing teachers in China have described the nature and purpose of assessment. Making use of inductive analyses and factor analytic techniques, variations in the constructs identified in teachersâ€™ thinking are identified and aligned across the study methods.Â Conceptions of assessment of Mainland China college lecturers: A technical paper analyzing the Chinese version of COA-III.