Exploring Assessment Literacy

HUANG Jing[a],*; HE Zonghui[a]

[a]Lecturer, School of Foreign Languages, China West Normal University, Nanchong, China.
*Corresponding author.

Abstract

Language assessment literacy represents an important aspect of teachers' professional knowledge. Research of this kind can serve the dual purpose of informing the nature and scope of teacher education reforms and the specific direction of professional development initiatives for pre-service and in-service teachers. This article attempted to explore the concept of assessment literacy and its training components through related literature review.

Key words: Assessment literacy; Assessment training components; Assessment modes

INTRODUCTION

Assessing student performance is one of the most critical aspects of the job of a classroom teacher. It impacts nearly everything that teachers do. Research on mainstream classroom teaching has revealed that day-to-day assessment of student learning is unquestionably one of teachers' most demanding, complex and important tasks (Shulman, 1986). According to Stiggins (1999), "The quality of instruction in any classroom turns on the quality of the assessments used there" (p.20). Teachers'

inadequate knowledge and capacity in assessment can cripple the quality of education (Popham, 2009). Cui (2008) argues that teachers should obtain their assessment knowledge and skills before they begin teaching. Coombe, Al-Hamly and Troudi (2009) emphasizes that without a higher level of assessment literacy, teachers will be unable to help students attain higher level of academic achievement. Popham (2004) even advocates that teachers' assessment illiteracy is professional suicide.

A set of assessment competencies for teachers have been identified and endorsed in virtually every set of standards of teacher competence developed recently in the U.S.A., including those developed by the National Education Association (NEA), the American Federation of Teachers (AFT), the Council of Chief State School Officers (CCSSO), the National Council for Accreditation of Teacher Education (NCATE), and the National Board of Professional Teacher Standards (NBPTS) (Wise, 1996).

In the field of English language teaching, TESOL partnered with the National Council for Accreditation of Teacher Education (NCATE) created the TESOL/NCATE Standards for ESOL teacher education. Assessment constitutes one of the five knowledge domains within these standards. In Europe, the Common European Framework of Reference and the European Portfolio for Modern Languages are requiring language teachers to adopt new ways of assessing language ability.

Clearly, there is widespread global recognition that language assessment literacy represents an important aspect of teachers' professional knowledge.

1. TEACHERS' KNOWLEDGE ABOUT LEARNING AND ASSESSMENT

Historically, as views of learning have changed, concepts of teaching and assessment have shifted concomitantly. Shepard (2000) presented a framework to illustrate this shift. The cognitive revolution brought new ways of

thinking about learning, which Shepard labeled the social-constructivist framework, drawn from the cognitive, constructivist, and sociocultural theories of the twenty-first century. This framework holds that learning is socially and culturally constructed; learning is more than associations and recall, and includes higher-order thinking, such as problem solving and discourse practices. Classroom assessment, based on this view of learning, addresses learning processes as well as outcomes, and includes explicit expectations, challenging tasks, and student responsibility (Shepard, 2000) by providing useful evidence for teachers to adapt their instruction and for students to improve their learning.

Discussion of language assessment literacy needs to be considered with reference to current assessment developments. During the last decades or so, new models of assessment and evaluation emerged in England, the United States, and Canada, which offered unique lens to understand the multifaceted nature of assessment literacy.

For example, Black and William's (1998) seminal work in England entitled *Inside the Black Box* was the first to explicitly highlight the central importance of formative assessment for improving student achievement. Collectively, their studies revealed how teachers can teach well and also get good test scores when they emphasize such things as questioning techniques, feedback without grades, peer assessment, self-assessment, and the formative use of summative tests as instructional strategies (Black, Harrison, Lee, Marshall, & William, 2004; Black & William, 1998). In essence, teachers need to create learning environments where students and teachers are active assessors during classroom instructional strategies. The latter is in stark contrast to the traditional view where assessments are primarily utilized at the end of an instructional unit or course of study.

In the United States, Stiggins (2002) has argued for new ways to think about assessment because over reliance on summative assessment approaches makes it virtually impossible for teachers to adapt teaching and learning to meet individual student needs. For him, assessment for learning (formative) must be balanced with the traditional assessment of learning (summative) so that teachers can feed information back to students in ways that enable them to learn better.

In Canada, Earl (2003) and Earl and Katz (2004) extended the work of Black et al. and Stiggins to advocate for synergy among assessment of learning (summative), assessment for learning (formative), and assessment as learning (the assessment is not graded but acts as a meta-cognitive learning tool). The latter is a sub-set of assessment for learning and occurs when students personally monitor what they are learning and use the feedback from this monitoring to make adjustments, adaptations, and even major changes in what they understand.

What is common in all these visions is teachers must recognize different purpose of assessment and use them accordingly. Clearly, assessment literate teachers must be able to design and administer more than summative end-of-unit tests and exams if they are to realize improvements in schools (Green & Mantz, 2002; Sheppard, 2000). The previously noted models suggest that teachers view assessment as pedagogy so that it is integrated into their best instructional strategies. Essentially, teachers need to shift their paradigm to understand how assessment can drive instruction and positively impact student learning and performance.

2. THE CONCEPT OF ASSESSMENT LITERACY

The term *assessment literacy* has been employed by several writers in recent years to describe what the constituency of language teachers and instructors needs to know about assessment matters (Inbar-Lourie, 2008; Malone, 2008; Stiggins, 1997; Stoynoff & Chapelle, 2005). But the concept of assessment literacy could be also expanded to describe the level of knowledge, skills, and understanding of assessment principles and practice that is increasingly required by other test stakeholder groups, depending on their needs and context (e.g., among educational advisors or government officials, policy planners and decision makers, the media, and the general public).

Others choose not to formally define assessment literacy, but rather to describe the characteristics of those who possess it as illustrated in Table1.

Table 1
Definition of Teacher Assessment Literacy (Characteristics)

Researchers/ organizations	Definitions
Sadler (1998)	Superior knowledge about content and substance of what is to be learned • Knowledge about learners and learning and a desire to help students develop, improve and do better • Skills in selecting and creating assessment tasks •Knowledge of criteria and standards appropriate to assessment tasks • Evaluative skills and expertise in the analysis and use of assessment information • Expertise in giving appropriate and targeted feedback.
Green & Mantz (2002); Sheppard (2000)	Assessment literate teachers must be able to design and administer more than summative end-of-unit tests and exams if they are to realize improvements in schools
McMillan (2000)	Teachers with a solid background in assessment are well positioned to integrate assessment with instruction so that they utilize appropriate forms of teaching.

<div align="right">To be continued</div>

Continued

Researchers/ organizations	Definitions
In a useful online publication from SERVE at the University of North Carolina (2004), they recommend that assessment-literate teachers know:	• How to define clear learning goals, which are the basis of developing or choosing ways to assess student learning • How to make use of a variety of assessment methods to gather evidence of student learning • How to analyze achievement data (both qualitative and quantitative) and make good inferences from the data gathered • How to provide appropriate feedback to students • How to make appropriate instructional modifications to help students improve • How to involve students in the assessment process (e.g., self and peer assessment) and effectively communicate results. • How to engineer an effective classroom assessment environment that boosts student motivation to learn.
Melter (2004)	Recognize sound assessment, evaluation, and communication practices Understand which assessment methods to use to gather dependable information and student achievement Communicate assessment results effectively, whether using report card grades, test scores, portfolios, or conferences Can use assessment to maximize student motivation and learning by involving student as full partners in assessment, record keeping, and communication
Center for School Improvement and Policy Studies, Boise State University (2007)	Assessment literate educators recognize sound assessment, evaluation, and communication practices; they Understand which assessment methods to use to gather dependable information and student achievement. Communicate assessment results effectively, whether using report card grades, test scores, portfolios, or conferences. Can use assessment to maximize student motivation and learning by involving student as full partners in assessment, record keeping, and communication.

In the TESOL/NCATE standards for ESOL teacher education, in the assessment domain, teachers are expected to understand the issues of assessment for ESL and language proficiency assessment for ESL (including how to develop assessments and use them to inform instruction).

In short, those who are assessment literate understand what assessment methods to use in order to gather dependable information about student achievement, communicate assessment results effectively, and understand how to use assessment to maximize student motivation and learning.

Five essential competencies of sound classroom assessment practice are described as follows by Stiggins:

a) Why assess? Assessment procedures and results serve clear and appropriate purposes.

b) Assess what? Assessments reflect clear and valued learning targets.

c) Assess how? Learning targets are translated into assessments that yield accurate results.

d) Communicate how? Assessment results are managed well and communicated effectively.

e) Involve students how? Students are involved in their own assessment (Stiggins, 2006, p.18; Stiggins et al., 2004, p.27).

3. ASSESSMENT LITERACY SURVEY

The issue of teacher training in classroom assessment has long been considered important in general education circles. Many researchers show that assessment training in teacher education is important to improve teachers' assessment literacy (Plake & Impara, 1993; Schafer, 1993; Stiggins, 2001). As Stiggins (2002, p.762) put it, "Few teachers are prepared to face the challenges of classroom assessment because they have not been given

the opportunity to learn to do so." To address the problem of inadequate assessment training for teachers, the American Federation of Teachers (AFT), National Council on Measurement in Education (NCME), and National Education Association (NEA) collaborated to develop the "Standards for Teacher Competence in Education Assessment of Students" or STCEAS to guide pre-service teachers' and educators' learning and course assessment. This standard remains an important authority in the field of teacher assessment literacy.

The Standards define assessment as "the process of obtaining information that is used to make educational decisions about students, to give feedback to the student about his or her progress, strengths, and weaknesses, to judge instructional effectiveness and curricular adequacy, and to inform policy" (AFT et al., 1990). They provide criteria for teacher competence with respect to the various components of this definition of assessment. They consist of the following seven principles:

* **STANDARD 1**—Teachers should be skilled in choosing assessment methods appropriate for instructional decisions.

* **STANDARD 2**—Teachers should be skilled in developing assessment methods appropriate for instructional decisions.

* **STANDARD 3**—The teacher should be skilled in administering, scoring and interpreting- the results of both externally produced and teacher produced assessment methods.

* **STANDARD 4**—Teachers should be skilled in using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement.

* **STANDARD 5**—Teachers should be skilled in developing valid pupil grading procedures that use pupil assessments.

* STANDARD 6—Teachers should be skilled in communicating assessment results to students, parents, other lay audiences, and other educators.

* STANDARD 7—Teachers should be skilled in recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information.

All seven standards apply to teachers' development and use of classroom assessments of the instructional goals and objectives that form the basis for classroom instruction. Standards 3, 4, 6, 7 also apply to large scale assessment, including administering, interpreting, and communicating assessment results, using information for decision making, and recognizing unethical practices (Brookhart, 2001).

Numerous research studies have been conducted over the past 20 years that have addressed one or more of the seven Standards. However, Plake (1993) addressed all teacher competencies as specified by *the Standards* for in-service teachers. One other study (Campbell et al., 2002) attempted to apply *The Standards* to groups of undergraduate preservice teachers. And Merlter (2005) tried to compare pre-service and in-service teachers' assessment literacy and found out that there were not significant correlation between teaching experiences and levels of assessment literacy of teachers.

In 1991, a national study measured teachers' assessment literacy (Plake, 1993) by using *the Standards* as a blueprint for the development of a survey instrument. The instrument (the Teacher Assessment Literacy Questionnaire) consisted of 35 items (5 per standard). Items were developed as application-type questions that were realistic and meaningful to teachers' actual practices. The instrument went through extensive content validation and pilot testing. A representative sample from around country was selected, and a total of 98 districts in 45 states participated, with a total usable sample of 555 surveys (Plake, 1993). The KR-20 reliability for the entire test was equal to .54 (Plake, Impara, & Fager, 1993). Teachers answered an average of slightly more than 23 out of 35 items correctly. The teachers' highest performance occurred on Standard 3-Administering, Scoring, and Interpreting the Results of Assessments *(M*=3.96/5.00); the lowest performance occurred on Standard 6 -Communicating Assessment Results *(M*=2. 70/5.00). On 10 of the 35 items, 90% or more of teachers answered the item correctly. These items addressed selecting appropriate assessments, acceptable test taking behavior for standardized testing situations, explanation of the basis for the grade to a child's parent, and the recognition of unethical practices in standardized test administration. On 5 items, less than 30% answered correctly. Two of the five came from Standard 5-Developing Valid Grading Procedures. Only 13% answered correctly an item that focused on steps to reliability of a test score. The two remaining items with low performance addressed Standard 7- Recognizing Unethical or Illegal Practices.

A similar study, conducted by Campbell et al. (2002), attempted to use the same assessment literacy instrument with undergraduate preservice teachers. The renamed *Assessment Literacy Inventory* (ALI) was administered to 220 undergraduate students following a course in tests and measurement. The course included topics such as creating and critiquing various methods of assessment, discussing ethical considerations related to assessment, interpreting and communicating both classroom and standardized assessment results, and discussing and evaluating psychometric qualities (i.e., validity and reliability) of assessments. The data from the undergraduate preservice teachers exhibited a higher level of reliability (ox=.74) than their inservice counterparts in the Plake et al. study (Campbell et al., 2002). The preservice teachers (*M*=21) averaged two fewer questions answered correctly than did the inservice teachers (*M*=23). Six items (numbers 5, 7, 22, 28, 31, and 35) demonstrated poor item discrimination values (<.20). The inservice teachers in the Plake et al. (1993) study scored higher than the preservice teachers on all but Standard 1 (Choosing Appropriate Assessment Methods). The preservice teachers scored highest on Standard 1, whereas the inservice teachers scored highest on Standard 3. Both groups of teachers scored lowest on Standard 6 (Communicating Assessment Results).

During the fall of 2002, Melter (2005) surveyed both preservice and inservice teachers with respect to their assessment literacy. The group of preservice teachers was comprised of 67 undergraduate students (science, and social studies) at a midwestern university. At the time of data collection, they were enrolled in methods courses (i.e., the semester preceding student teaching) and had just completed a specific course in classroom assessment. The group of inservice teachers consisted of 101 teachers, representing nearly every district and school in a three-county area surrounding the same institution. The schools were selected based on convenience due to their geographic location. All secondary grade levels and content areas were represented in the final sample.

Both groups were surveyed using the *Classroom Assessment Literacy Inventory,* or CALI, which consisted of the same 35 content-based items (five per standard) used in the *Teacher Assessment Literacy Questionnaire* (Plake, 1993) with a limited amount of rewording (e.g., changing some names of fictitious teachers, changing word choice to improve clarity, etc.), as well as seven demographic items. The items were grouped by Standard; Table 1 shows the alignment of items with their respective Standard. The 35 items presented the respondents with assessment-related scenarios, followed by a question with a specific correct answer. All items used a multiple-choice format with four options, one being the correct response. The original instrument has been shown to have reasonable reliability with both inservice teachers, rKR-20=.54 (Plake et al., 1993), and preservice teachers, ot=.74 (Campbell et al., 2002).

Table 2
Alignment of the Standards With Respective CALI Items

Standard	Item numbers
1. Choosing Appropriate Assessment Methods	# 1, 2, 3, 4, 5
2. Developing Appropriate Assessment Methods	# 6, 7, 8, 9, 10
3. Administering, Scoring and Interpreting the Results of Assessments	# 11, 12, 13,14,15
4. Using Assessment Results to Make Decisions	# 16,17, 18,19, 20
5. Developing Valid Grading Procedures	# 21, 22, 23, 24, 25
6. Communicating Assessment Results	# 26, 27, 28, 29, 30
7. Recognizing Unethical or Illegal Practices	# 31, 32, 33, 34, 35

Note. The Classroom Assessment Literacy Inventory (CALI) can be viewed. Retrieved from www.bgsu .edu/colleges/edhd/LPS/EDFI/ faculty/mertler/CALI. html

Interestingly, the architects of the ALI suggest that pending some additional research, the survey could also be used as a diagnostic instrument geared toward the identification and remediation of classroom assessment misconceptions or weakness of in-service teachers (Volante & Fazio, 2007).

In mainland China, theoretical and empirical studies on teachers' assessment literacy are also carried out in the general educational circles and focused more on primary and secondary teachers. A recent doctoral study of Chinese primary and secondary teachers' assessment literacy and its development (Zheng, 2009) conducted a teachers' assessment literacy survey administering 954 questionnaires to primary and secondary teachers working in Zhejiang Province, China. The questionnaire used in the study consisted of three sections, addressing background information of the instructors' experience (Part I), teachers' attitude toward assessment (Part II), and measurement of teachers' levels of assessment literacy (Part III). The first two sections were drafted by the researcher and the last one was adapted from Plake and Impara (1993). After the quantative data was professionally processed, he found out that the level of assessment literacy of those teachers in Zhejiang Province is quite low, even unacceptable. Based on the survey results and literature review, the researcher proposed a structural model of teachers' assessment literacy. The researcher also discussed the development of teachers assessment literacy should depend on the collaborative efforts of the government, school authorities and individual teachers. Although his investigation on assessment literacy and its development is very pioneering in China, besides teachers' self-report survey, it is a pity that he didn't employ other research methods to triangulate the survey data. Another limitation of the study is that it is not centered on one subject matter, such as English.

4. THE COMPONENTS OF ASSESSMENT LITERACY TRAINING

Based on the above discussion on assessment literacy, the second research focus of this study is to the training components of assessment literacy. Clearly, language instructors needs some measure of assessment training to gain higher level of assessment literacy if they are engaged in selecting, administering, interpreting, and sharing results of large-scale tests produced by professional testing organizations, or in developing, scoring, interpreting, and improving classroom-based assessment (Taylor, 2009). Then, what should they acquire to become proficient in the language of assessment?

Davies (2008) reviews several seminal textbooks used to teach language testing between 1960 and the present day, showing how they exemplify changing trends and perceptions in training focus. He noted that a steady curricular expansion from basic technical testing expertise (the how-to), to knowledge about language and measurement (the what), and, most recently, to understand of key principles with regard, for example, to validity and ethics. Davies characterized the present paradigm as that of Skills + Knowledge + Principles:

> Skills provide the training in the necessary and appropriate methodology, including item-writing, statistics, test analysis, and increasingly software programmes for test delivery, analysis and reportage. Knowledge offers relevant background in measurement and language description as well as in context setting. Principles concern the proper use of language tests, their fairness and impact, including questions of ethics and professionalism. (Davies, 2008, p.335)

4.1 Brindley's Theoretical Framework for Assessment Professional Development

Language assessors, particularly teachers, are expected to engage in classroom assessment practices, report on learners' progress aligned with external criteria as well as prepare learners for external examinations (Inbar-Lourie, 2008). To comply with these demands, Brindley (2001) offers an outline for programs for professional development in language assessment, which focuses on the knowledge components required for conducting language assessment in an educational context. The outline is modular in that it acknowledges different assessment needs, some of which are regarded as core and some as optional, and provides a useful framework for considering the components of the language assessment literacy knowledge base for analyzing and discussing language assessment competences in relation to the

assessment knowledge dimensions mentioned above: The reasoning or rationale for assessment (the "why"), the description of the trait to be assessed (the "what"), and the assessment process (the "how").

This framework construes language assessment not as a collection of assessment tools and forms of analyses, a body of knowledge and research grounded in theory and epistemological beliefs, and connected to other bodies of knowledge in education, linguistics and applied linguistics. The core competences will reflect current views about the social role of assessment in general and language assessment in particular, contemporary views about the nature of language knowledge, and give due emphases to both classroom and external assessment practices. The focus and intensity would vary and depend on the target audience, but an introduction to the core components will be obtained by all participants, including discussion of some of the unsolved controversies and tensions in the field (Inbar-Lourie, 2008).

These theoretical views of assessment training provide a starting point for understanding language assessment literacy. However, empirical findings about teacher assessment knowledge (assessment courses) can extend our understanding of assessment literacy.

4.2 Survey on Language and Testing Course Outside of Mainland China

Clearly the general purpose of language testing and assessment courses partly aims to train teachers' language assessment literacy. However, language assessment courses do not come in one shape or size. And not much is known about the nature and contents of language assessment courses, as little research has been targeted specifically at the objectives and contents of the courses.

In the mid-1990s Bailey and Brown (1995) reported on a survey of language testing courses in the context of language teacher training programs around the world. They examined various course characteristics, including topics covered and core textbooks used. A decade later, when they replicated their study, they found "the presence of a stable knowledge base that is evolving and expanding rather than shifting radically" (Bailey & Brown, 2008, p.371). They also found evidence of other types of advanced language testing courses being provided to prospective language teachers in some graduate programs. It remains unclear, however, how far the content of such courses is likely to equip teachers to manage the complex social and educational issues that are implicated in assessment today.

While the Bailey and Brown study surveyed a substantial number of respondents about their language testing courses, researchers in two more recent studies report directly and in more detail on their own language assessment course. The methodology used in both cases is narrative analysis but from different perspectives: while one study (O'Loughlin, 2006) explores the course from the point of view of two of the course participants (using their written contributions on an online forum), the second research reflects on the experience of transforming the course from the instructor's perspective (Kleinsasser, 2005). The descriptions of both courses allow insight into the course objectives and contents as well as the manner in which the courses were conducted. Considering the O'Loughlin (2006) study is more relevant to the current study, more detailed information is offered here. The course included both practical components and discussion of conceptual themes, such as social issues in language testing. The researcher reports that two of the students, whose narratives form the research focus of the study, both attained the course objectives. However, differences emerged as to their willingness and capacity to embrace new ideas in the area of language assessment. These differences are attributed to personal background and professional experience and context, emphasizing the need to consider the learners' diverse cultural background and experiences when planning and conducting the course (O'Loughlin, 2006). Following this line of thoughts, the current study on Chinese in-service College English instructor will contribute to the literature from a Chinese cultural perspective.

4.3 Survey on Language Testing and Assessment Courses Within Mainland China

Following Bailey and Brown's suit, by using survey and follow-up interview by emails and/ or telephone, Jin (2010) investigated the design and instruction of language testing and assessment courses at the tertiary level in mainland China with a view to improving the quality of such courses and to better preparing potential language teachers for their future careers. 86 instructors of such courses across the country participated in the survey commenting on the current situation in terms of the instructor background, teaching contents, teaching methodology, teaching materials and student perceptions of the courses provided as part of the curriculum of relevant programs training future foreign language teachers. An encouraging finding of this study is that language testing and assessment courses in tertiary institutions in China did adequately cover the essential theoretical and practical aspects of language testing. However, the analysis of teaching focuses revealed that recent developments in language testing and assessment, the social dimension of language testing for example, were not fully incorporated into the teaching content. Topics on assessment (e.g., alternative assessment, formative and summative assessment) did not have efficient coverage. In-service teacher training regarding language assessment is not discussed at all. She only mentioned about the concept of the core competences of language assessment literacy (Inbar-Lourie, 2008), but no further suggestions were given on how to improve teachers' language assessment literacy. Based on her

survey, this study will try to generate a Chinese College English instructors' language assessment literacy training model to direct Chinese pre-service and in-service English teacher training in assessment.

At the end of the paper, Jin (2010) pointed out that

improving the quality of courses and providing foreign language teachers with "the core competences" in assessment, therefore, clearly represent a step forward in raising the professional standards of practitioners in foreign language education, which will, in due course, benefit millions of foreign language learners in China. (Inbar-Lourie, 2008, p.396, 570)

The present study is a step towards expanding our understanding of the professional development of tertiary level English teachers in China.

4.4 Empirical Studies on Preservice and Inservice Assessment Literacy Training

Assessment training at both the preservice and inservice levels is crucial. Researchers have also examined the needs of preservice and inservice teachers in developing assessment literacy. In fact, as far as teacher preparation in assessment is concerned in EFL context, teachers in Hong Kong report that they received little or no training in assessment (Falvey & Cheng, 1995). Shohamy (1998) and Ferman (1998) found that EFL teachers in Isreal felt they lacked the knowledge and training required to practice assessment procedures. Another study done with tertiary-level English language teachers in the United Arab by Emirates and Kuwait, Troudi, Coombe and Al-Hamly (2009) found that teachers often felt marginalized in the area of assessment because of their perceived lack of knowledge about the subject. Similar views are also reflected in the recent study (Kiomrs, Abdolmehdi, Rashidi, & Naser, 2011) that the Iranian EFL teachers of secondary schools have a very poor knowledge base in language assessment.

This research base, although limited, demonstrates the need to understand language teacher assessment literacy more deeply so that we can design teacher preparation and professional development program to meet their needs.

Volante and Fazio (2007) explored more generally the assessment literacy development of primary/junior teacher candidates (grades1-6; ages 5-10) in a teacher education program. Based on results from an assessment literacy questionnaire administered throughout the four-year program, they reported that candidates maintained relatively low levels of self-efficacy in each year of the program. Candidates in this study emphasized the need for practical knowledge about assessment (i.e. assessment practices, tool development) and they overwhelmingly supported the notion of a specific course in classroom assessment and evaluation. The study also found that within their current model of assessment programming, candidates relied heavily on the mentorship of their associate teacher in aiding the development of their assessment literacy. Melter (2009) examined the effectiveness of a two-week classroom assessment workshop for in-service teachers. The workshop was based on the *Standards for Teacher Competence in the Educational Assessment of Students* and focused on discussion, practice, and practical application through performance assessment tasks. The training was shown to be highly effective for the teachers. DeLuca and Klinger (2010) examined assessment education at one pre-service teacher education program in Ontario, Canada. Through a questionnaire administered to 288 teacher candidates, this study identifies teacher candidates perceived confidence levels in assessment practice, theory, and philosophy. These teacher candidates also provided their view toward assessment topics which are important to include in a pre-service educational assessment course.

Teacher's dual role as both teacher and assessor of curriculum attainment, in addition to being (in the case of language teacher) a "facilitator of language development" (Rea-Dickins, 2007, p193) entails a special kind of teachers' assessment literacy. In the Chinese culture, examinations are valued, and decisions that are made on the basis of examination results are generally accepted almost without question, as fair (Bachman, 2010). Despite the crucial role of assessment in shaping the quality of teaching, there is evidence that teachers universally suffer from poor assessment literacy in classroom assessment (Volante & Faziao, 2007). Several reasons have been suggested which affects the enhancement of teachers' assessment literacy. A commonly-held belief is that an individual knows how to teach a language, he or she knows how to assess the product and the process of language learning as well (Spolsky, 1978). Such common mistaken beliefs contribute negatively to further neglect of teachers' knowledge base in language assessment. Additionally, traditional delivery approaches to teaching assessment courses in pre-service and in-service programs have also resulted in teachers' alienation from assessment issues (Inbar-Lourie, 2008).

At the same time, more training alone is insufficient to meet the needs of training in language assessment literacy. It is important that such training includes the necessary content for language instructors to apply what they have learned in the classroom and understand the available resources to supplement their formal training when they enter the classroom. Melter (1999) concluded that this potentially implies that teachers tend to develop assessments skills on the job, as opposed to structure environments such as courses or workshops. Therefore, it is very important to know more about the in-service teachers' assessment practice and their assessment literacy level to serve as initiative for training program design.

Perhaps the most serious issue, though, is that of assessment literacy and the field's lack of knowledge as to the extent of this professional knowledge about assessment among language instructors. Swender, Abbott, Vicars, and Malone (2006) have revealed that major gaps

exist in the applications (or perceived applications) of foreign language tests by language instructors. The first step is to determine what language instructors need to know about assessment in order to perform their jobs and secondly, to determine how to provide such training.

Therefore, it is worthwhile to explore the current state of in-service College English teachers' assessment literacy and their classroom assessment practice. So, by the way of survey, interviews, classroom observation, this research seeks to investigate several interrelated topics surrounding Chinese College English instructors' assessment literacy and their classroom-based assessment practices, including what teachers should know, what they do know, what they are trained to know, and what might be done to improve the quality of training.

CONCLUSION

Assessment is the bridge that links the curriculum and drives the instruction (Oleksak, 2007), as curriculum changes to reflect many and varied goals, the form of assessment must also change. While English classrooms and textbooks have made great strides in the last two decades, changes in assessment practices are in an urgent need. As illustrated by these theoretical and empirical studies have contributed a great deal to the conceptualization of teachers' assessment literacy in general and language teachers' assessment literacy in particular. However, there also exist some limitations. First, as shown by the review, the majority of the research on teachers' assessment literacy done in the general educational circle; and less research in language assessment community. Moreover, most of the assessment literacy research also has focused on primary and secondary teachers other than tertiary level teachers and more on pre-service teacher than in-service teachers. Second, few of empirical studies on language teachers' assessment literacy have deployed a triangulation of instruments and sources to gather data. Third, as indicated by the review, assessment literacy is a hierarchical and relativistic concept, due to the paucity of relevant research and literature, whether the language instructors are proficient users of classroom assessment or not have been unclear. So have been their training needs, therefore, more empirical studies in this field are urgently needed.

REFERENCES

American Federation of Teachers, National Council on Measurement in Education, National Education Association. (1990). Standards for teacher competence in educational assessment of students, *Educational Measurement: Issues and Practice, 9*(4), 30-32.

Bachman, L. F. (2000) Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17,* 1.

Bailey, K. M., & Brown, J. D. (1995). Language testing courses. What are they? In A. Cumming & R. Berwick (Eds.), *Validation in language testing* (pp.236-256). Clevedon, UK: Multilingual Matters.

Bailey, K. M., & Brown, J. D. (2008). Language testing courses: What are they in 2007? *Language Testing, 25*(3), 349-383.

Black, P., & William, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*(2), 139-148.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy& Practice, 5*(1), 17-18.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). Working inside the black box: Assessment for learning in the classroom. (Cover story). *Phi Delta Kappan, 86*(1), 9-21.

Borg, W. R., Worthen, B. R., & Valcarce, R. W. (1993). Teachers' perceptions of the importance of educational measurement. *Journal of Experimental Education, 55,* 9-14.

Boyles, P. (2005). Assessment literacy. In M. Rosenbusch, (Ed.), *National assessment summit papers* (pp.11-15). Iowa State University, Ames, Iowa.

Brindley, G. (2001). Outcomes-based assessment in practice: Some examples and emerging insights. *Language Testing, 18*(4), 393-408.

Brindley, G. (1991). Assessing achievement in a learner-centered curriculum. In J. C. Alderson & North, B. (Eds.), *Language Testing in the 1990s* (pp.153-166). London: Modern English Publications.

Brindley, G. (1997). Assessment and the language teacher: Trends and transitions. *Language Teacher Online, 21*(9). Retrieved from http://www.jalt_publications.org/tlt/files/97/sep/brindley.html

Brookhart, S. M. (2001) *The standards and classroom assessment research*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Dallas, TX, March (ERIC Document Reproduction Service No. 451189).

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

Buhagiar, M. (2006). The classroom assessment cycle within the alternative assessment paradigm: Exploring the role of the teacher. *Journal of Maltese Education Research*, *4*(2), 17-16.

Chang, C.-W. (2005). Oral language assessment: Teachers' practices and beliefs in Taiwan collegiate EFL classrooms with special reference to Nightingale University. UK: Unpublished doctoral dissertation, University of Exeter.

Cheng, L. (1999). Changing assessment: Washback on teacher perspectives and actions. *Teaching and Teacher Education, 15*(3)

Cheng, L. (2001) An investigation of ESL/EFL teachers' classroom assessment practices. *Language Testing Update, 29,* 53-83.

Cheng, L., Rogers, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods and procedures. *Language Testing, 21*(3), 360-389.

Cheng, L., & Qi, L. (2006). Description and examination of the national matriculation English test in China. *Language Assessment Quarterly: An International Journal, 3,* 53-70.

Cheng, L, Rogers, T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classroom. *Assessment and Evaluation in Higher Education, 33*(1), 9-32.

Cheng, L. (2008). The key to success: English language testing in China. *Language Testing, 25*(1), 15-37.

Cheng, L., & Curtis, A. (2010). E*nglish language assessment and the Chinese learner.* New York and London: Routledge Taylor & Francis Group.

Chinese Ministry of Education. (2004, 2007). *College English curriculum requirements.* Beijing, China: Foreign Language Teaching and Research Press.

Cizek, G. J., Fitzgerald, S. M., & Rachor, R. E. (1996). Teachers' assessment practices: Preparation, isolation, and the kitchen sink. *Educational Assessment, 3,*159-179.

Cobb, C. T. Joyner, J. M., & William, J. L. (1999). *Assessment literacy for educators: Linking instruction and assessment* (pp.3-11). ERS Spectrum, Spring.

Coombe, C., Al-Hamly, M., & Troudi, S. (2009). Foreign and second language teacher assessment literacy: Issues, challenges and recommendations. *Research Notes, 38*(3), 14-18.

Cowan, J. (1998). *On becoming an innovative university teacher.* Buckingham: RHE and Open University

Cui, Y. H. (2008). Teachers should learn to assess prior to instruction. *Curriculum in Elementary and Secondary Education, 11,* 55-58.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327-347.

DeLuca, C., & Klinger, A. D. (2010), Assessment development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy& Practice, 17*(4), 419-438.

Dietel, R., Herman, J. L., & Knuth, R. A. (1991). *What does research say about assessment?* NCREL, Oak Brook.

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research, 38*(1), 47-65.

Hargreaves, E. (2005). Assessment for learning? Thinking outside the black box. *Cambridge Journal of Education, 35*(2), 213-224.

Harlen, W. (2005). Teachers' summative practices and assessment for learning: Tensions and synergies. *The Curriculum Journal, 16*(2), 207-223.

He, A. E. (1996). *Teaching English at China's tertiary universities: Policy, practice and future prospective.* Melbourne, Australia: Monash University.

Heritage, M. (2007). Formative assessment: What do teacher need to know and do? *Phi Delta Kappan, 89*(2), 140-145.

Hu, Z. L. (2004). Individualization, Collaboration, Modularization, and Hypertextualization in College English Reform. *Chinese Foreign Languages,1*(1), 49-75.

Hu, Z. L. (2005). China's English Education Reform: Trends and Issues. *Asia TEFL, 2*(5), 1-4.

Impara, J., Plake, B., & Fager, J. (1993). Teachers' assessment background and attitudes toward testing. *Theory Into Practice, 32*(2), 113-117.

Jin, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing, 27*(4), 555-584.

Kiomrs, R., Rashidi, A., & Naser, R. (2011). On the interaction of test washback and teacher assessment literacy: The case of Iranian EFL secondary school teachers. *English Language Teaching, 4*(1), 156-161.

Kleinsasser, R. C. (2005). Transforming a postgraduate level assessment course: A second language teacher educator's narrative. *Prospect, 20,* 77-102.

Kunnan, A. J. (2004). Regarding language assessment. *Language Assessment Quarterly, 1*(1), 1.

Language Testing. (2008). A special issue of 25th anniversary. *Language Testing, 2*(3).

Liu, R. Q., & Dai, M. C. (2004). On the reform of college English teaching in China. *CELEA Journal, 27*(4), 3-8.

Mavrommatis, Y. (1997). Understanding assessment in the classroom: Phases of the assessment process—the assessment episode. *Assessment in Education: Principles, Policy & Practice, 4*(3), 381-400.

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practice, 22*(4), 34-43.

Mertler, C. A. (1999) Assessing student performance: A descriptive study of the classroom assessment practices of Ohio teachers. *Education, 120*(2), 285-296.

Mertler, C. A., & Campbell, C. (2004). *Assessing those who assess: Development of an instrument to measure teachers' assessment literacy*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus, Ohio.

Mertler, C. A. (2005). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(2), 76-92.

Mertler, C. A., & Campbell, C. (2005). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the Assessment Literacy Inventory.* Paper presented at the annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.

Mertler, C. A. (2009). Teachers' assessment knowledge and their perceptions of the impact of classroom professional development. *Improving School, 12*(2), 101-113.

Mousavi, S. A. (2002). *Encyclopedic dictionary of language testing* (3rd ed.). Taipei: Tung Hua Book Company.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385-402

Oleksak, R. (2007). Building a framework for national assessment literacy. *Foreign Language Annals, 40*(4), 565.

O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate student on-line subject forum. *University of Sydney Papers in TESOL, 1,* 71-85.

O'Sullivan, R. S., & Johnson, R. L. (1993). *Using performance assessments to measure teachers' competence in classroom assessment.* Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA, April.

O'Sullivan, R. G., & Chalnick, M. K. (1991). Measurement related course work requirements for teacher certification and recertification. *Educational Measurement: Issues and Practice, 10*(1) 17-19, 23.

Plake, B. S. (1993) Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher, 6*(2), 21-7.

Plake, B., Impara, J., & Fager, J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice, 12*(4), 10-12.

Patton, M. Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications.

Popham, W. J. (2004). All about accountability: Why assessment illiteracy is professional suicide. *Educational Leadership, 6*(1), 82-83.

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing*, *21*(3), 249-258.

Quilter, S. M., & Gallini, J. K. (2000). Teachers' assessment literacy and attitudes. *The Teacher Educator*, *36*(2), 115-131.

Nashville, TN.

Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and Teacher Education, 27*, 472-482.

Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel: Recommendations and reality. *Journal of Teacher Education, 38*(3), 57-63.

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice, 10*(1), 3-6.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4-14.

Shohamy, E. (1998). Inside the "black box" of classroom language tests, *Studia Anglica Posnaniensia, XXXIII,* 343-52.

Shohamy, E., & Hornberger, N. (Eds.). (2008). Encyclopedia of language and education (2nd ed.): Vol.7. *Language Testing and Assessment*. New York: Springer Science+ Business Media

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4-14.

Spolsky, B. (1978). Introduction: linguists and language testers. In B. Spolsky (Ed.), *Advances in language testing series 2: Approaches to language testing* (pp.v-x). Arlington, VA: Center for Applied Linguistics.

Stiggins, R. J., & Bridgeford, N. J. (1985). The ecology of classroom assessment. *Journal of Educational Measurement, 22*(4), 271-286

Stiggins, R. J., Conklin, N. F., & Bridgeford, N. J. (1986). Classroom assessment: A key to effective education. *Educational Measurement: Issues and Practice, 5,* 5-17.

Stiggins, R. J. (1991a). Assessment literacy. *Phi Delta Kappan, 72*(7), 534 -540.

Stiggins, R. J. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice, 10*(1), 20-32

Stiggins, R, J, & Conklin, N. F. (1992). *In teachers' hands: Investigating the practice of classroom assessment.* Albany, NY: State University of New York Press.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-246.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23-17.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan, 83*(10), 758-765.

Stiggins, R. J. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan, 86*(2), 86-92.

Stoynoff, S. (2009). A survey of developments in ESOL testing. In C. Coombe, P. Davidson, & D. Lloyd (Eds.), *Fundamentals of language assessment: A practical guide for teachers.* Dubai: TESOL Arabia Publications.

Swender, E., Malone, M., Vicars, R., & Abbott, M. (2006). *ACTFL assessment for performance and proficiency of languages: Progress update*. Paper presented at 2006 ACTFL Conference.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29,* 21-26.

TESOL/NCATE Program Standards. (2003) Retrieved from http://clas.uncc.edu/linguistics/Internal%20documents/NCATEP12Standards.pdf Alexandra, VA: TESOL.

Troudi, S., Coombe, C., & Al-Hamly, M. (2009). EFL teachers' views of English language assessment in higher education in the United Arab Emirates and Kuwait. *TESOL Quarterly, 43*(3), 546-555.

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749-770.

Wang, S. F. (2008). Review on assessment literacy in China. *Contemporary Education Science, 12,* 3-6.

Ward, J. G. (1980). Teachers and testing: A survey of knowledge and attitudes. In L. M. Rudner (Ed.), *Testing in our schools* (pp.15-24). Washington, DC: National Institute of Education.

Wise, A. E. (Ed.). (1996). Quality teaching for the 21st Century. *Phi Delta Kappan*, *78*(Special Issue), 190-224.

Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly, 43*(3), 49-513.

Xu, Y. T., & Liu, Y. C. (2008). Narrative study on college English teachers' knowledge of formative evaluation. *Foreign Language Learning, Theory and Practice, 3,* 61-68.

Zhen, D. H. (2009). *Study on development of teachers' assessment literacy* (Doctoral dissertation). East China Normal University.

Zhou, Y. (2010). Teachers as the key factor for English teaching in EFLT. *Foreign Language Teaching and Research, 4,* 294-297.