TEACHER ASSESSMENT LITERACY IN PRACTICE: A RECONCEPTUALIZATION

Yueting Xu

Guangdong University of Foreign Studies, Guangzhou, China

Gavin T. L. Brown

The University of Auckland, New Zealand

Correspondence concerning this paper should be addressed to to Ms. Yueting Xu, School of English and Education, Guangdong University of Foreign Studies, Higher Education Mega Centre, Guangzhou, People's Republic of China 510006 or by email to traceyxu@gdufs.edu.cn

Abstract:

This paper aims to reconceptualize teacher assessment literacy (AL) by connecting two fields of research: educational assessment and teacher education. It begins with a scoping review of AL studies. By synthesizing and analyzing 100 studies on teacher AL, a new conceptual framework of teacher assessment literacy in practice (TALiP) is proposed. This framework is illustrated by a discussion of the various components of teacher AL and their interrelationships. This paper concludes with the theoretical contributions of the framework, a working definition of TALiP, and implications for policy and practice of assessment education.

Keywords: teacher assessment literacy, reconceptualization, teacher education

1. Introduction

Assessment literacy (AL), traditionally defined as a basic understanding of educational assessment and related skills to apply such knowledge to various measures of student achievement (Stiggins, 1991a), is increasingly being recognized as an integral part of teacher professionalism (Abell & Siegel, 2011; Brookhart, 2002; Engelsen & Smith, 2014; Schafer, 1993; Stiggins, 1995). Such growing interest in AL is due partly to the central role of assessment in student learning (Black & Wiliam, 1998), and to strong evidence that teachers are key agents in educational assessment (Leung, 2014). To help students attain higher levels of academic achievement, teachers need to develop appropriate types and levels of AL (Stiggins, 1995). Despite the compelling arguments for AL (Brookhart, 2011), many teachers are often involved in assessment-related decision-making without sufficient background or training in assessment (DeLuca, 2012; Lam, 2015; Popham, 2009; Schafer & Lizzitz, 1987). As a result, "assessment illiteracy abounds" (Stiggins, 2010, p. 233).

To address this problem, evidence has been gathered concerning the knowledge and skills that teachers need to be considered assessment literate, their training requirements, efficacy in assessment, as well as contextualized understanding of AL (e.g., DeLuca & Klinger, 2010; Gottheiner & Siegel, 2012; Plake, Impara, & Fager, 1993). In addition, models of teacher AL have been proposed with different foci, such as different stakeholders' perspectives of assessment education (DeLuca, 2012), theoretical knowledge of classroom-based assessment (Abell & Siegel, 2011), and socio-cultural perspectives of 'assessment literacies' within multiple discourses (Willis, Adie & Klenowski, 2013). As these prior models do not integrate with the principles of pre- and in-service teacher education, a new framework of assessment literacy education is needed to create an overall trajectory of professional development in AL and to encompass all phases of teacher education and development. To address this need, the present paper, based on synthesis and analysis of prior studies, develops a framework that inter-relates the knowledge and skill components of teacher AL with known characteristics of effective pre- and in-service teacher education, as well as the details of sociocultural contexts (e.g., policy, cultural values, social norms) in which assessment is implemented.

To that end, this paper is organized into three parts: the first part reviews AL research for identifying emerging themes and implications for teacher education; the second part describes our conceptual framework of Teacher Assessment Literacy in Practice (TALiP); the third part summarizes the theoretical contribution of the framework, proposes a working definition of TALiP, and highlights its implications for policy and practice of teacher assessment education. This study intends to map out key conceptual issues surrounding teacher AL to guide future empirical endeavors and interventions.

2. Review of AL Studies

The goal of reviewing the literature on current research into AL was to understand

what worked and did not for developing teacher AL, and to use that knowledge to draw clear connections between various emphases in previous studies.

2.1 Search method and results

To comprehensively identify relevant studies on teacher AL, we used a criteria-based scoping review (Arksey & O'Malley, 2005). While definitive definitions of scoping reviews do not yet exist (Pham, et al., 2014), scoping reviews, as opposed to systematic reviews, tend to address broader issues emerging from relevant studies in order to create an overview of a diverse body of work, regardless of methodological approaches. The time span was set from 1985 to 2015 since research prior to 1985 has been reviewed by Campbell (2013). The search was restricted to peer-reviewed journal articles, chapters, and academic books published in English. Unpublished theses and dissertations, conference papers, and textbooks or resource books were not eligible for inclusion. Articles not available in full-text were also excluded, as was research on the assessment literacy of other stakeholders (e.g., students, testers). Hence, this scoping review may not be exhaustive, but was deemed sufficient for an overview of the field.

During the first stage, searches were performed in GoogleScholar, Academic Search Premier, and ERIC with a combination of the search terms: 'teacher,' 'assessment literacy,' 'assessment competence,' 'professional development in assessment,' and 'assessment expertise.' A search of these terms identified 254 journal articles, book chapters, and books. Titles and abstracts of these publications were read to select relevant studies. To be included in our initial search, a study had to reference teachers as the main targets or participants of the research and to discuss assessment literacy as a main theme.

In the second stage, citations within relevant studies were located to identify possible new sources. These studies were subsequently evaluated in the same manner with regards to main targets and themes. To ensure the scoping was comprehensive, key journals in educational assessment and teacher education were inspected (i.e., *Assessment in Education: Principles, Policy & Practice, Assessment & Evaluation in Higher Education, Educational Assessment, Educational Measurement: Issues and Practice, Educational Assessment, Evaluation and Accountability, Teaching and Teacher Education, Journal of Teacher Education,* and *Teachers and Teaching: Theory and Practice*). Relevant articles were identified using the same selection process and citations within relevant papers.

A total of 100 publications were identified, which consisted of 71 empirical and 29 non-empirical studies. These selected studies were inductively analyzed using thematic coding procedures (Patton, 2002). In the first cycle of coding (Miles, Huberman, & Saldaña, 2014), studies were read to elicit key phrases that represented their research foci. We either utilized the keywords provided by the authors or created new ones ourselves. This round of coding produced 173 codes, as many studies had more than one code. The second cycle of coding reduced the initial codes into 10 sub-themes, which are shown in *italics* in Table 1. These 10 sub-themes were further

aggregated into three main themes: (1) knowledge and skills within AL, (2) assessment education and its relationships with various mediating factors, and (3) contextual considerations of AL, potentially representing a progression in understanding AL. Table 1 lists the studies under each theme and sub-theme, orders them chronologically by date of publication, and indicates type of research (empirical or non-empirical). In addition, implications for teacher education in each study are included. These implications are then summarized, making clear the connections between teacher education and educational assessment and, more importantly, highlighting the aspects incorporated into a reconceptualization of teacher AL.

2.2 Knowledge and Skills within AL

The first line of research into AL began with attempts to describe the appropriate content and standards of knowledge and skills needed by teachers in order to be considered assessment literate. As Table 1 shows, the primary goal seems to have been the establishment of a 'knowledge base' for AL. The *Standards for Teacher Competence in Educational Assessment of Students* (hereafter the *Standards*) (AFT, NCME, & NEA, 1990) was a seminal outcome of this work. The *Standards* prescribe seven competency areas in which teachers should be skilled. The seven areas are: 1) choosing assessment methods appropriate to instructional decisions; 2) developing assessment methods appropriate for instructional decisions; 3) administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods; 4) using assessment results when making decisions about individual students, planning teaching, developing curriculum, and school improvement; 5) developing valid pupil grading procedures; 6) communicating assessment results to various stakeholders; and 7) recognizing unethical, illegal, and inappropriate assessment methods and uses of assessment information.

Based on the *Standards*, subsequent studies have discussed the AL knowledge base for different subject areas (e.g., language, science), assessment purposes (e.g., learning or accountability purposes), and stakeholders (e.g., teachers, students, administrators) (Abell & Siegel, 2011; Inbar-Lourie, 2008; Taylor, 2013). Recent studies have proposed updated lists of the knowledge base, calling for inclusion of capabilities based upon recent developments in assessment policy and practice (i.e., formative assessment and accountability contexts) (Brookhart, 2011; JCSEE, 2015; Stiggins, 2010). These standards set goals for assessment education and lay a solid foundation for empirical studies on AL (see DeLuca, LaPointe-McEwan, & Luhanga, 2015).

Table 1 Synthesized AL research

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
Theme 1: Knowledge and skills within AL			
Knowledge Base			
1	(AFT, NCME, & NEA, 1990)	NE	Standards to be incorporated into future teacher training and certification programs
2	(Schafer, 1993)	NE	Involving classroom assessment specialists in schools to supervise teachers' AL development; improving both teachers' knowledge and motivation in assessment with a reward system
3-7	(Stiggins, 1991a, 1991b, 1995, 1999, 2010)	NE	More on-going assessment training; right time to promote AL; careful curriculum design for AL; tending to teacher emotion; a need for faculties of education self-assessing whether they are producing assessment-literate teachers; joint efforts from various stakeholders
8-9	(Inbar-Lourie, 2008, 2013)	NE	A knowledge base for language assessment literacy (LAL) and more assessment training for language teachers
10	(Popham, 2009)	NE	Assessment training for both accountability and learning purposes
11	(Brookhart, 2011)	NE	An updated list of teacher assessment competencies to guide teacher educators, teachers, and educational specialists
12	(Abell & Siegel, 2011)	E	Incorporating constructive view of learning and principles of assessment into the AL knowledge base for science teachers; providing opportunities for teachers to make connections between assessment theories and practice
13-14	(Taylor, 2009, 2013)	NE	Going beyond offering opportunities for professional development; more classroom-based assessment competencies for teachers
15	(Lian, Yew, & Meng, 2014)	NE	Sustained efforts for AL development to be incorporated as teachers' daily practice
16	(DeLuca, LaPointe-McEwan,	NE	A need for considering contemporary assessment standards, local policies and priorities, and assessment education and support for teachers when constructing AL measures; a need for

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
	& Luhanga, 2015)		differentiating AL measures for different teacher populations (i.e. pre- & in-service)
17	(JCSEE, 2015)	NE	Classroom assessment principles to be incorporated into teacher AL
Measuring teacher AL			
18	(Impara, Divine, Bruce, Liverman, & Gay, 1991)	E	More training of assessment results interpretations
19	(Popham, 1991)	E	Concerns for ethics in teachers' test-preparation practice
20	(Plake, Impara, & Fager, 1993)	E	Understanding teacher assessment needs and preferences
21	(Impara & Plake, 1995)	E	AL development of school counselors and administrators as resources for teacher AL
22	(Mertler, 2004)	E	More on-going 'on-the-job' assessment training
23	(Maclellan, 2004)	E	A need for teacher education paying more attention to enabling pre-service teachers to develop greater cognitive complexity in thinking about assessment
24	(Alkharusi, Kazem & Al-Musawai, 2011)	E	Continuous assessment training for in-service teachers and diagnosis for pre-service teachers' conceptions of assessment
25	(Kiomrs, Abdolmehdi, & Rashidi, 2011)	E	More down-to-earth approaches focusing more on assessment practice than on theoretical principles
AL measurement validation			
26	(Alkharusi, 2011)	E	Nil
27	(Gotch & French,	NE	Nil

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
28	2014) (Hailaya, Alagumalai, & Ben, 2014)	E	Nil
Theme 2: Assessment education and its relationships with various mediating factors			
Assessment courses			
29	(Schafer, 1991)	NE	A need for basic assessment training and joint support from principals, curriculum specialists, and teacher educators
30	(Airasian, 1991)	E	A call for educational assessment specialists to understand informal classroom assessment activities and to include relevant concrete training materials to prepare teachers for these activities
31	(O'Sullivan & Chalnick, 1991)	E	A need for teacher AL as integral part of teacher accreditation and certification, and for assessment training to be organized consistently with current continuing education credit structure
32	(Brookhart, 1999)	NE	Classroom assessment courses to be taught through a mixture of direct instruction in the concepts and application scenarios for classroom practice, simulation, and discussion
33-34	(Bailey & J. D. Brown, 1995; J. D. Brown & Bailey, 2008)	E	Informing language testing instructors of the content to be covered
35	(Jin, 2010)	E	A need for a network of teacher-testers for an exchange of experiences in the design and instruction of assessment
36	(DeLuca, Klinger, Searle, & Shulha, 2010)	E	A need for pre-service teacher education to develop student teachers' self-directed learning about assessment
37	(Popham, 2011)	NE	A need for a regular assessment course based on legitimate assessment textbooks

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
38	(DeLuca, Chavez, Bellara, & Cao, 2013)	E	Continued inquiry into the pedagogical conditions for effective assessment education; helping pre-service teachers develop foundational skills (esp. critical reflection) to negotiate and navigate their learning about assessment, and creating opportunities for them to connect theory with practice
39	(DeLuca & Bellara, 2013)	E	More alignment of pre-service assessment curriculum to professional standards and delineating AL as a core teaching construct
40	(Jeong, 2013)	E	Attention to the backgrounds of language assessment course instructors
Assessment training programs & resources			
41	(Schafer & Lizzitz, 1987)	E	Including measurement competencies as part of teacher certification standards
42	(Wise, Lukin, & Roos, 1991)	E	A call for measurement as a compulsory course in pre-service curriculum
43	(Lukin, Bandalos, Eckhout, & Mickelson, 2004)	E	Building a learning community for teachers involved in AL training
44	(Graham, 2005)	E	Attending to pre-service teachers' prior beliefs of assessment and providing qualified mentor teachers who model good assessment practices
45	(Campbell & Collins, 2007)	E	Core assessment content to be included in assessment textbooks and coursework irrespective of disciplines
46	(Sato, Wei, & Darling-Hammond, 2008)	E	National certification as a fruitful way for promoting AL
47	(Mertler, 2009)	E	Applying in-depth and intensive approach to AL development
48	(Schneider & Randel, 2010)	NE	A need for administrative support, individualized learning goals and content, sufficient contact hours, collaboration, and coherent training system

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
49	(Fan, Wang, & Wang, 2011)	E	Using e-learning environment to enhance teacher AL
50	(K. Smith, 2011)	E	Entailing teachers' internal need for change
51	(Koh, 2011)	E	More localized & sustained professional development; design of assessment tasks and rubrics as sites for effective teacher learning
52	(Leahy & Wiliam, 2012)	NE	Subject-specific training with modules allowing for teachers' free choice, flexibility, small steps, and supportive accountability environment
53	(DeLuca, 2012)	NE	Using the proposed framework of assessment education to strengthen the delivery of assessment courses by incorporating perspectives of teacher education policy makers, teacher educators and teacher candidates
54	(Greenberg & Walsh, 2012)	NE	A need for assessment course syllabi across teacher training programs in U.S. to reflect the practical and current demands of school teachers and administrators
55	(Hill, Ell, Grudnoff, & Limbrick, 2014)	E	A need for both a dedicated assessment course and embedded assessment training across courses and practicum in pre-service teacher education, as well as for further support to pre-service teachers who need to use assessment to enable student learning in schools
56	(DeLuca, Klinger, Pyper, & Woods, 2015)	E	Ongoing support and opportunities for teachers to share and explore their learning
57	(Lam, 2015)	E	A need for AL training to meet teachers' localized needs

The relationship among assessment training, teacher conceptions of assessment and AL

58	(Quilter & Gallini, 2000)	E	A call for assessment training to begin with addressing teachers' perceptions & attitudes about assessment
59	(G. T. L. Brown, 2008a)	E	A need for considering teacher conceptions of assessment

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
60	(Hill, Cowie, Gilmore, & Smith, 2010)	NE	A need for investigating pre-service teachers' conceptions and their interactions with assessment learning experiences
61	(L. F. Smith, Hill, Cowie, & Gilmore, 2014)	E	Greater liaison between universities and school to attend to teachers' beliefs and their evolving process
62	(Gunn & Gilmore, 2014)	E	The important role of pre-service assessment education in orientating student teachers' conceptions of assessment towards professional expectations
63	(Hill, Gunn, Cowie, Smith, & Gilmore, 2014)	E	A need for teacher educators to take account of the student teachers' prior assessment knowledge and beliefs about assessment into their programs; greater liaison between universities and schools to provide more consistency for pre-service teachers to move between the two contexts
64	(L. F. Smith & Galvin, 2014)	E	A call for an integrated approach to assessment curriculum with repetition of materials, connections between theory and practice, incorporation of political landscapes and sufficient time for student teachers to try out new ideas
65	(DeLuca & Lam, 2014)	E	Assessment education curriculum attending to teachers' understanding of assessment in inclusive education
66	(Levy-Vered & Alhija, 2015)	E	A need for increasing AL training and understanding teacher conceptions of assessment

Teacher assessment training needs and self-reported efficacy

67	(Gullickson, 1993)	E	More alignment between teachers' assessment training needs and assessment specialists/professors' prioritized areas
68	(Brookhart, 2002)	NE	Classroom assessment to be prioritized over large-scale assessment in the knowledge base; teacher educators and assessment specialists listening to teachers' assessment needs
69	(Zhang &	E	A need for assessment training programs to be tailored to suit the differential needs of teachers

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
	Burry-Stock, 2003)		working in different content areas and grade levels
70	(Volante & Fazio, 2007)	E	Joint efforts from government, faculties of education, school district leaders, administrators, and associate teachers to promote AL
71	(DeLuca & Klinger, 2010)	E	A need for multiple perspectives for assessment education, including those of associate teachers, beginning teachers and teacher education faculty
72	(Leighton, Gokiert, Cor, & Hefferman, 2010)	E	Scrutinizing teacher conceptions of assessment and extending AL to include competencies for cognitive diagnostic assessment
73	(Fulcher, 2012)	E	Informing design of assessment curriculum and online resources
74	(Ell, Hill, & Grudnoff, 2012)	E	Informing teacher educators that that the prior knowledge of assessment that student teachers bring to pre-service teacher education can be both a resource and a challenge
75	(DeLuca, Chavez, & Cao, 2013)	E	The responsibility of pre-service teacher education to expand teacher conception of assessment and to improve AL confidence
76	(Howley, Howley, Henning, Gillam, & Weade, 2013)	E	Teachers as active contributors to AL knowledge base and teachers' learning about assessment through reflection and participation
77	(Vogt & Tsagari, 2014)	E	Contextualized AL training to meet the needs of local teachers
78	(Rorsyth, Cullen, Ringan & Stubbs, 2015)	E	A shared language and norm to help teachers to feel trusted to make appropriate assessment decisions

Theme 3: Contextual considerations of AL

Macro and micro contexts

| 79 | (Xu & Liu, 2009) | E | Joint-efforts from teachers, teacher educators and administrators to promote PD in assessment |

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
80	(Forsberg & Wermke, 2012)	E	Increasing teacher autonomy in their learning of assessment; teaching experiences and cooperation with colleagues as two most important sources for teachers' assessment knowledge
81	(Gu, 2014)	E	More AL training for teachers and joint-ownership of assessment
Teacher's identity as assessor			
82	(Adie, 2012)	E	Online moderation as a revenue for developing 'assessor' identity
83	(Scarino, 2013)	E	Considering teachers' preconceptions about assessment and interpretive process of their practices and identity as assessors
84	(Cowie, Cooper, & Ussher, 2014)	E	Tensions and contradictions between assessment coursework and practicum experiences as generative sites for productive changes of pre-service teachers' conceptions of assessment yet requiring careful orchestration by teacher educators
Understanding and developing AL in practice			
85	(Lomax, 1996)	E	Hands-on experiences for pre-service training and inclusion of school teachers as course partners
86	(Borko, Mayfield, Marion, Flexer, & Cumbo, 1997)	E	More explicit attention to both beliefs and practices of assessment, and a need for situating in-service teacher assessment training within teachers' daily classroom practice
87	(McMillan, 2003)	E	A need for modifying the nature of assessment principles and providing a more relevant set of guidelines that will help teachers' classroom assessment practices
88	(Buck, & Trauth-Nare, 2009)	E	A need for providing in-service teachers with opportunities to contemplate the efficacy of classroom assessment practices
89	(Frey & Fisher, 2009)	E	Using teacher-created, standard-aligned formative assessment as sources for professional development
90	(Wyatt-Smith, Klenowski, &	E	Developing a common language among teachers to talk about assessment

Themes & subthemes	Authors & Date	Empirical or non-empirical	Implications for Teacher Education (if any)
91	Gunn, 2010) (Siegel & Wissehr, 2011)	E	Connection between theory and practice in pre-service curriculum and more attention to teacher conceptions of assessment
92	(Allal, 2011)	E	Social moderation as an avenue for AL development
93	(Gottheiner & Siegel, 2012)	E	Providing teachers with useful assessment resources for well-informed decisions
94	(Lyon, 2013)	E	Translating theoretical progression of assessment expertise into tools and instruction to support teacher preparation and learning
95	(Campbell, 2013)	NE	Understanding how contextually unique classroom environments may present barriers to AL and how informational resources potentially negate the positive effects of assessment training
96	(Leung, 2014)	NE	Sensitizing teachers to specific affordances and purposes of classroom assessment and promoting teacher awareness of their conceptions and practices
97	(Willis, Adie, & Klenowski, 2013)	NE	Enabling teachers to engage in critical enquiry of their assessment practice
98	(Engelsen & Smith, 2014)	E	Teachers, students and school leaders 'speaking the same language' when engaging in dialogues about assessment
99	(Fleer, 2015)	E	Regular dialogues among teachers about their assessment practices
100	(Herman, Osmundson, Dai, Ringstaff, & Timms, 2015)	E	Inclusion of content and pedagogical knowledge into teacher AL

Using the 1990 *Standards* as a blueprint, instruments were developed to investigate teacher AL levels, such as *Teacher Assessment Literacy Questionnaire* (TALQ) (Plake, et al., 1993) and *Classroom Assessment Literacy Inventory* (CALI) (Mertler, 2004). These objective tests were intended to measure what teachers knew about the prescribed competencies and to identify strengths and weaknesses in AL (e.g., Alkharusi, Kazem & Al-Musawai, 2011). Although the strong and weak areas found in these studies varied with different samples, the consensus was that teacher assessment knowledge was generally inadequate relative to standards and expectations. A similar conclusion was found in Maclellan's (2004) study which examined teacher knowledge of assessment by analyzing their written scripts. Although these studies are a useful entry point for empirical research on AL, the results need to be interpreted and generalized with caution due to the limited evidence for the psychometric properties of the instruments (Gotch & French, 2014).

Several concerns can be identified from this approach. First, the 1990 *Standards* and more recently updated and revised guidelines (e.g., Brookhart, 2011; JCSEE, 2015) set reasonable goals for teachers and teacher educators to work towards, and yet the existence of such goals alone does not inherently advance teacher AL. Second, any knowledge base consists predominantly of 'core' knowledge dimensions that are applicable to all subject areas and contexts; however, as in most educational practices, assessment is implemented in subject-specific and context-dependent circumstances (Abell & Siegel, 2011; Taylor, 2013). Third, the knowledge base is not static; for example, the 1990 *Standards* have been replaced by newer benchmarks (e.g., JCSEE, 2015) based on latest assessment research findings. Fourth, while measuring teachers' mastery of the principles drawn from the knowledge base can give an estimate of teacher AL knowledge, transfer to the practical realm is not guaranteed. Consequently, research into AL now focuses on the effectiveness of assessment education and its mediating factors.

2.3 Assessment education and its relationships with various mediating factors

Subsequent studies are concerned with assessment education for both pre- and in-service teachers as well as its relationships with various mediating factors, such as teachers' training needs, conceptions of assessment, and efficacy. The underpinning assumption is that teacher AL would improve if pre- and in-service programs prepare teachers sufficiently and, if not, there must be some mediating factors that inhibit its effectiveness. In what follows, sub-themes under this heading are reviewed consecutively, and distinctive features of pre- and in-service assessment education are discussed. Notwithstanding these features, integrating assessment education practices across the spectrum of teacher employment and seeing them as falling along the same developmental trajectory would be beneficial for an overall understanding of teacher AL.

As a formal way of improving teacher AL, the quality of assessment courses was an initial concern (see Greenberg & Walsh, 2012). Thus, studies were conducted into course content and its inclusion rationale (Brookhart, 1999; Popham, 2011; Schafer,

1991), evaluations of course characteristic factors (e.g., instructors, content, students, and alignment with professional standards) (J. D. Brown & Bailey, 2008; DeLuca & Bellara, 2013; Jeong, 2013; Jin, 2010), and pedagogies that facilitate learning about assessment (DeLuca, Chavez, Bellara, et al., 2013). Despite criticisms of assessment courses for being theory laden and disconnected from teachers' assessment practices, these studies all suggest essential components for quality assessment courses, including carefully-tailored content aligned with professional standards, well-trained instructors who can make connections between theory and practice, and useful pedagogies that engage student teachers in critical reflection of assessment. Sadly, many pre-service teacher education programs still cannot benefit from these findings because they only offer a one-semester assessment course that provides a general introduction to assessment (Greenberg & Walsh, 2012), or else do not have such a course at all (Popham, 2011).

The second focus was on the effectiveness of assessment education in both pre- and in-service programs. For pre-service teachers to become assessment literate, a number of conditions need to be met: 1) assessment education taking various forms and integrating different stakeholders' perspectives (DeLuca, 2012; Hill, Ell, Grudnoff, & Limbrick, 2014; Mertler, 2009); 2) AL becoming part of teacher accreditation and certification (Sato, Wei, & Darling-Hammond, 2008; Schafer & LIzzitz, 1987); 3) mentors attending to student teachers' prior beliefs on assessment (Graham, 2005); and 4) training content being localized, subject-area specific that allow for teachers' free choice (Lam, 2015; Leahy & Wiliam, 2012). In-service teachers who have little time and limited opportunities for formal training may need to learn from on-line learning resources (Fan, Wang, & Wang, 2011), seek support from within the workplace (Lukin, Bandalos, Eckhout, & Mickelson, 2004), and utilize daily classroom practices as sites for their AL development by implementing assessment for learning (AfL) (K. Smith, 2011), instructional rounds (DeLuca, Klinger, Pyper, & Woods, 2015), and design of assessment tasks and rubrics (Koh, 2011). Notwithstanding these different conditions, consensus is that assessment training needs to be sustainable. Given that many assessment courses in pre-service programs are one-offs, a critical question then arises: what other factors contribute to the (in)effectiveness of assessment training?

The answer to this question can be found in those studies that explored the relationship among assessment training, conceptions of assessment and teacher AL. The results are inconclusive: some studies found that the amount and type of training in assessment was independent of the beliefs teachers had about the purposes of assessment (G. T. L. Brown, 2008a), suggesting that perceptions might resist training; while others reported a positive relation between assessment training and teacher AL (Levy-Vered & Alhija, 2015; Quilter & Gallini, 2000). Nevertheless, they seem to agree that teachers have their own deeply-rooted conceptions of assessment that arise from their experiences with school assessment (i.e., assessment evaluates students rather than is used to improve teaching and learning) (e.g., Gunn & Gilmore, 2014; L. F. Smith, Hill, Cowie & Gilmore, 2014), which constitutes a significant barrier to developing pre-service teacher AL (Quilter & Gallini, 2000). Clearly, assessment

education needs to encompass both technical knowledge of assessment and more consciousness-arousing components that prompt teachers to reexamine their conceptions (DeLuca & Lam, 2014; Hill, Cowie, Gilmore, & Smith, 2010).

Further studies into teacher training needs and their assessment efficacy have produced several intriguing findings. First, there was a mismatch in preferred content emphases between teachers and their professors, with the former desiring more topics about formative and summative uses of tests and non-test evaluation, while the latter preferred test statistics and analysis (Gullickson, 1993). Second, although teachers feel an urgent need for assessment training, there was a tendency for them to express a need for 'everything' rather than being specific about their needs (DeLuca & Klinger, 2010). Third, teachers by and large lacked confidence in their assessment practices, and were in dire need of assessment training, resources, and opportunities (Volante & Fazio, 2007). Fourth, teachers reported that they acquire assessment knowledge in the field through reflection and collaboration about classroom experiences (Howley, Howley, Henning, Gillarn, & Weade, 2013; Wise, Lukin, & Roos, 1991).

These studies present obvious challenges for AL. Teacher perceptions of their assessment training needs may not be what they actually want. An agreed basis is needed to ensure alignment between what experts (e.g., professors in colleges of education) and teachers consider to be important knowledge. Since pre-service assessment education is either non-existent or weak and, when shown to be effective, its impact is limited by the amount of time available to develop AL (e.g., Airasian, 1991), more evidence is needed about the contribution of 'on-the-job' learning can make towards AL. These challenges suggest that a contextual understanding of AL may provide a more useful approach to teacher assessment education.

2.4 Contextual considerations of AL

Although direct investigations into how contextual factors influence teacher AL are generally lacking, inferences can be made concerning how factors at different levels exert an influence on AL in different ways. At the national level, policies influence AL by shaping professional development activities and teacher autonomy (Forsberg & Wermke, 2012) and by mandating curriculum standards, textbook use, and large-scale tests (Gu, 2014). At the institutional level, structural conditions like power relations in the workplace shape teachers' personal practical knowledge of assessment (Xu & Liu, 2009). At the personal level, teachers' awareness of – and actions to construct – their identity as assessors also matter (Adie, 2013; Cowie, Cooper, & Ussher, 2014; Scarino, 2013). Although these mediating factors are addressed separately in each study, how they interrelate with each other and what other factors might matter remain unknown. Particularly when these factors are put together, it is unclear how AL is developed and enacted in day-to-day classroom practice.

Some studies have taken a more integrated approach to AL, placing emphasis on teachers as professionals whose conceptions and practice of assessment are situated in specific contexts rather than on the mastery of assessment principles (Frey & Fisher, 2009; Lomax, 1996; Wyatt-Smith, Klenowski, & Gunn, 2010). Informed by

sociocultural views of learning, AL is conceptualized as "a dynamic context dependent social practice" (Willis et al., 2013, p. 242). This conceptualization of assessment literacies in the plural form is understood as contextualized and culturally responsive practices that teachers negotiate with their students in learning communities. These insights provide a starting point for us to rethink the nature of AL with perspectives from different learning theories and views of assessment.

With this contextualized perspective, teacher AL becomes a joint enterprise requiring input and support from many stakeholders, such as students, school administrators, policy makers, and even the general public (Allal, 2011; Engelsen & Smith, 2014; Fleer, 2015), although teachers remain the agents of assessment education. Indeed, only when teachers become the main drivers of their own AL development will they make full use of classroom-based assessment as sites for learning (Gottheiner & Siegel, 2012), engage in regular dialogues and collaboration with colleagues (Wyatt-Smith, et al., 2010), and participate in assessment activities within communities of practice (Willis, et al., 2013).

These studies of contextual understanding of AL suggest that, although teachers are expected to acquire sufficient AL knowledge as promulgated in various standards, the reality is that teachers who are placed in complex contexts have to make professional decisions about assessment in response to various factors that may facilitate or inhibit their practices. Their decision-making is thus a process by which they balance the demands of external factors with their own beliefs and values (McMillan, 2003). Solutions to improve teacher AL, therefore, are by no means universal, but rather contextual. Situating teachers' assessment decision-making and action-taking processes among the different dynamics at play in specific contexts appears essential to understanding and developing solutions to improve teacher AL.

2.5 Implications for teacher education

Table 1 shows that most of the reviewed AL studies have meaningful implications for teacher education. Four pertinent issues are derived. First, teachers need a solid AL knowledge base and teacher education programs need to include assessment courses as part of their curriculum and certification (Sato, et al., 2008; Schafer & Lizzitz, 1987). Further, the content of the assessment knowledge base needs to keep abreast of research and policy innovations (Fan, et al., 2011; Popham, 2009).

Second, teacher AL training needs to become long-term, sustainable, individualized, and 'on-the-job' (Graham, 2005; Koh, 2011). Assessment education needs to be long enough to engage teachers in complex and deep learning about assessment, which will possibly lead to changes in their conceptions and practices regarding assessment. In addition, assessment education ought to make strong connections between the knowledge base and the context of practice (Lyon, 2013). AL development needs to be situated within the requirements of different educational contexts and, thus, will have different priorities at different times and places (Vogt & Tsagari, 2014). AL training also needs to be placed within a supportive system with joint efforts from various stakeholders (Frey & Fisher, 2009; Impara & Plake, 1995).

Third, there is a pressing need for better understanding of teachers as individuals and professionals (Hill, et al., 2010) because the effectiveness of assessment training might be offset by teachers' conceptions, emotions, needs, and prior experiences about assessment (G. T. L. Brown, 2008a; DeLuca & Bellara, 2013; L. F. Smith, et al., 2014). Careful diagnoses of these factors and deliberate efforts to guide teachers toward assessment knowledge and skills are essential (DeLuca, Chavez, & Cao, 2013; Siegel & Wissehr, 2011).

Fourth, teacher AL development is not merely an accumulation of assessment knowledge, but rather the development of a sophisticated, contextually-appropriate set of inter-related competencies. Teacher professionalization in assessment, like any other area of development, involves many factors at play interactively and simultaneously. Teacher autonomy, identity as assessor, and critical inquiry of their assessment practices may enable agency (Cowie, et al., 2014; Scarino, 2013; Willis, et al., 2013). Teachers need to be involved in learning communities where they have a common language to share, negotiate, and make decisions about their assessment practices (Lukin, et al., 2004; Wyatt-Smith, et al., 2010).

These four issues suggest that teacher AL is dependent on a combination of cognitive traits, affective and belief systems, and socio-cultural and institutional influences, all of which are central to teacher education. They pinpoint the complexity of AL and entail a need for reconceptualizing teacher AL with insights from prior studies and new perspectives from teacher education.

3. Reconceptualizing AL: Teacher Assessment Literacy in Practice

Building upon the reviewed studies and relevant literature in teacher education, a conceptual framework of TALiP is presented (Figure 1). Before discussing the specific components of the framework, a few clarifications are in order. First, prior AL models have informed our framework in many ways. For example, our framework has gained insights from Willis, et al.'s (2013) model that AL is a social practice situated in multiple discourses and DeLuca's (2012) framework that assessment principles need to be connected and extended to classroom practice. Second, there seems to remain a need for a new conceptual framework that bridges the distinctive concerns of teacher education and educational assessment, integrates pre- and in-service teacher education as a whole, and delineates key components and influences in preparing teacher to become assessment literate. Third, the content and structure of this framework are based upon both the prior studies reviewed earlier and our accumulated experiences as a teacher educator (first author) and an educational assessment specialist (second author). Fourth, research work contributing to this framework goes beyond the AL studies reviewed earlier by including key literature from teacher education.

This framework consists of six components. From the bottom to the top respectively, they are the knowledge base, teacher conceptions of assessment, institutional and socio-cultural contexts, TALiP the core concept of the framework, teacher learning, and teacher identity (re)construction as assessors. In the following

sections, the role of each component in the conceptual framework is outlined and their reciprocal relationships are further examined.
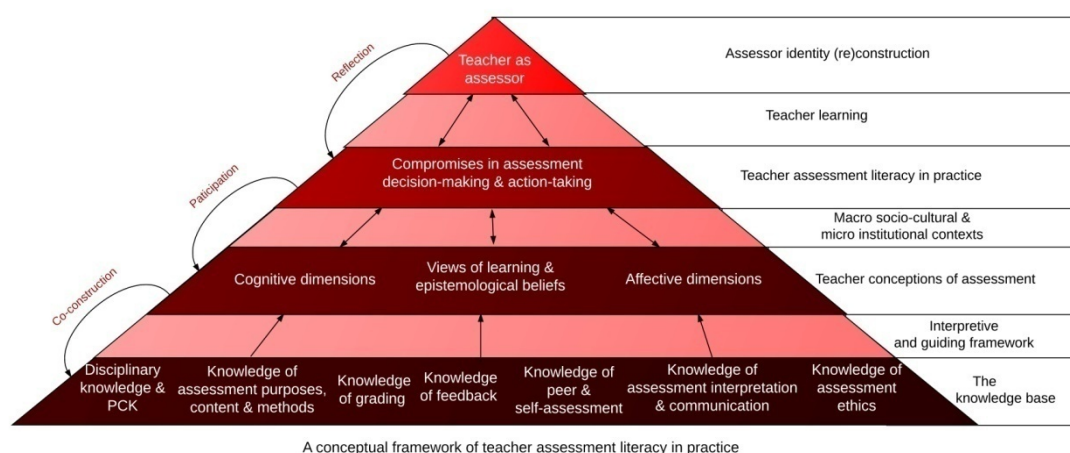

A conceptual framework of teacher assessment literacy in practice

Figure 1 A conceptual framework of teacher assessment literacy in practice

3.1 The knowledge base as a necessary but insufficient condition for TALiP
At the bottom of this pyramid is the knowledge base, which is the basis of all other components, since clarity of the knowledge that both pre- and in-service teachers need for effective assessment practice is essential (Maclellan, 2004). Without the knowledge base, there would be no standards or criteria by which the appropriateness of assessment practice could be evaluated, potentially causing failed outcomes for teachers and students (Fulcher, 2012).

However, the knowledge base is a necessary, but insufficient, condition for TALiP. It is a necessary condition because education systems need a core body of formal, systematic, and codified principles concerning good assessment practice. These can be reflected in standards or other codified documents, representing shared knowledge useful for teacher licensure, accreditation, and design of classroom assessment practices. Yet, the knowledge base is insufficient because these principles only serve as decontextualized guidelines and are not ready-made solutions to problems that arise within complex and diverse classroom assessment scenarios.

Based on the *Standards* (AFT, NCME, & NEA, 1990) and recent updates (e.g., Brookhart, 2011; JCSEE, 2015), as well as some major contemporary assessment textbooks for teachers (Airasian & Russell, 2008; G. T. L. Brown, Irving, & Keegan, 2014; McMillan, 2001; Nitko & Brookhart, 2011), an appropriate assessment knowledge base is proposed.

3.1.1 Disciplinary knowledge and pedagogical content knowledge (PCK). Since educational assessment is about measuring the curriculum content taught in schools/universities, knowledge of disciplines and how to teach that content cannot be excluded from the assessment knowledge base (Abell & Siegel, 2011; Brookhart, 2011; Shulman, 1987).

3.1.2 Knowledge of assessment purposes, content, and methods. Teachers need to

know why they assess (i.e., formative, summative), how different assessment methods can be related to the learning goals and specific content being learned, and what a variety of relevant assessment strategies are (AFT, NCME, & NEA, 1990; Brookhart, 2011; JCSEE, 2015).

3.1.3 Knowledge of grading. Depending on the grading system within a school or jurisdiction, teachers need to know rationales for grading, different methods for creating grades (i.e., norm-, criterion-, or ipsative-referenced), the content of grades (i.e., academic achievement or affective performances), principles for using criteria or rubrics, as well as different ways for assigning or weighing components into grades (Nitko & Brookhart, 2011). Related to this is an understanding of scoring techniques for objectively-marked testing and principles of consistency and moderation for judgment-based assessments.

3.1.4 Knowledge of feedback. Teachers need to know the purposes and principles of feedback, different types of feedback (e.g., descriptive, evaluative, or supportive) with their respective functions, as well as the strengths and weaknesses of various kinds of feedback (e.g., task, process, metacognitive, and self-oriented) in facilitating student learning (Hattie & Timperley, 2007; JCSEE, 2015).

3.1.5 Knowledge of assessment interpretation and communication. Teachers need to know ways of interpreting evidence generated from assessment, ensuring that their interpretations have sufficient warrant (Brookhart, 2011). They also need to know ways of communicating assessment results to stakeholders such as students, parents, managers/administrators, and the general public in such a way that inappropriate interpretations (e.g., ignoring margin of error or reliability indices) are discouraged.

3.1.6 Knowledge of student involvement in assessment. Teachers need to understand the benefits of involving students in assessment (e.g., self- and peer assessment and transparent learning objectives or standards). Teachers also need to know strategies of using self- and peer assessment in different assessment tasks, as well as those of training students to effectively participate in assessment (JCSEE, 2015).

3.1.7 Knowledge of assessment ethics. Teachers need to understand legal and ethical responsibilities concerning the use, storage, and dissemination of assessment results (Tierney, 2013). Furthermore, teachers are required to know how to work towards equity, non-discrimination, inclusion, and social justice (AFT, NCME, & NEA, 1990; Brookhart, 2011).

These domains represent a general body of key theoretical principles of assessment independent of specific contexts. A basic mastery of this body of knowledge serves as a threshold; teachers who have crossed it would have the opportunity to engage in assessment at a deeper level. It should be noted that this knowledge base is dynamic and evolving, depending in part on the context in which assessment is deployed.

3.2 Teacher conception of assessment as an interpretive and guiding framework

Educating teachers with the content of the knowledge base, whether by lecture,

workshop, or seminar, that knowledge is filtered and interpreted by teacher conceptions of assessment (Barnes, Fives, & Dacey, 2015; Fives & Buehl, 2012). Teacher beliefs are an interpretive and guiding framework by which they mediate their uptake of theoretical knowledge and its implementation. Conceptions of assessment denote the belief systems that teachers have about the nature and purposes of assessment, and that encompass their cognitive and affective responses.

Teacher conceptions of assessment have both cognitive and affective aspects, both of which are framed by broader views of teaching and learning and their epistemological beliefs (G. T. L. Brown, 2008b). The cognitive dimension denotes what teachers believe is true and false about assessment. How teachers respond to information in the knowledge base depends in part on the degree to which new knowledge is consistent with, or dissonant from, their current conceptions. In sum, teachers tend to adopt new knowledge, ideas, and strategies of assessment that are congruent with their conceptions of assessment, while rejecting those that are not.

The affective dimension denotes emotional inclinations that teachers have about various aspects and uses of assessment. Arising from various assessment experiences in their careers, teachers have strong and weak, positive and negative emotions about assessment (Crossman, 2007). This means that, consistent with Green (1971), some aspects of teacher conceptions of assessment are relatively stable and deeply-held systems that are resistant to change, while other aspects, perhaps less deeply experienced, are easier to change. The emotional dimension of conceptions may make conceptual change difficult, leading to less effective learning about assessment and reduced effectiveness in implementing new assessment policies.

In addition, teacher conceptions are both collective due to the social nature of education and individualized. They are influenced by policy priority in specific socio-cultural and institutional contexts (G. T. L. Brown, 2008b; L. F. Smith, et al., 2014) and subject to an individual's personal and educational experiences (Hill, et al., 2010). Furthermore, as Ajzen (2005) makes clear, it is difficult for an individual to implement changes in assessment practice in a work context in which the social norm is resistant to a changed practice. Therefore, professional development efforts that ignore the emotional component and existence of implicit conceptions by relying purely on rational persuasion are likely to be less successful (G. T. L. Brown, 2008a; DeLuca, Chaves, & Cao, 2013). To improve teacher AL inevitably involves a long process of attending to, and possibly changing, teachers' existing conceptions of assessment. Although some studies have acknowledged the role that teacher conceptions play in shaping AL (G. T. L. Brown, 2008a; Levy-Vered & Alhija, 2015), few studies into teacher AL have included it as a legitimate dimension.

3.3 Micro- and macro-contexts as the boundaries for TALiP

Despite the decisive role of teacher conceptions of assessment, in-service teachers cannot do whatever they please in actual practice since they are employed within an immediate workplace community and larger social, political, and cultural contexts. These micro- and macro-contextual variables exert an influence on teachers'

assessment practices individually or in concert through policies, norms, rules, regulations, and conventions to "create a culture of certainty and compliance that is not easily challenged by teachers" (Scarino, 2013, p. 312). These boundaries take multiple forms, as small as pre-specified criteria or as large as national assessment policies. These contextual variables set boundaries for teachers' assessment practices in terms of what they should do and what they should refrain from doing (Gu, 2014). These boundaries are also socio-culturally specific, and, thus, teachers working in different contexts are expected to work towards different goals and outcomes.

Likewise, pre-service teachers' learning about assessment is oftentimes constrained by their institutional contexts and local and national policies. For countries where there exist national legislation and professional standards about assessment, these directives play a large part in prescribing assessment education curriculum (DeLuca & Bellara, 2013). In addition, the structures of assessment education (e.g., explicit, integrated, and blended) and their actual implementation (e.g., consistent or discrete content in these courses) (DeLuca & Klinger, 2010) may facilitate or constrain AL development. For countries where such national standards or assessment courses are not available, social norms and dominant discourses in the workplace instead play a prominent role in setting these boundaries. Pre-service teachers' learning about assessment through 'apprenticeship of observation' (Lortie, 1975) may largely depend on the quality of assessment experiences modeled by the university professors or associate teachers in schools (Graham, 2005).

Apart from official regulations, in-service teachers' assessment practices are oftentimes under the influence of needs and interests of various stakeholders, including the media, school administrators, parents, students, and colleagues (Wyatt-Smith, et al., 2010; Xu & Liu, 2009). For example, in face of students with diverse learning needs, some teachers may adjust their feedback (e.g., cover truth of poor performance with excessive praise [Pajares & Graham, 1998]), while others may change their instruction to drill students for high-consequence, externally-mandated school-accountability tests (Nichols & Harris, in press). Unlike the community rules and educational policies that set 'hard' boundaries for teacher assessment practice, the needs and interests of these stakeholders seem to set 'soft' boundaries that are more flexible, and teachers may have more autonomy over whether or not to accommodate these needs into their practices.

The tighter the boundaries, the less space there is for professional autonomy. Tensions arise for teachers when they have less autonomy (Fleer, 2015; Forsberg & Wermke, 2012) and can arise because of incongruence between their conceptions and the boundaries imposed upon them within their context. Hence, teachers are forced to make compromises to exercise assessment literacy in practice.

3.4 Teacher assessment literacy in practice as compromises made among tensions

As summarized by McMillan (2003), teachers' assessment decision making is a process by which teachers balance the demands of external factors and constraints with their own beliefs and values. Good assessment practices therefore involve the art

of compromise (Carless, 2011). Building upon these insights, we argue that AL is better understood as teacher assessment literacy in practice (TALiP), which consists of various compromises that teachers make to reconcile tensions. As TALiP is constantly negotiated between teachers' conceptions of assessment and the macro socio-cultural, micro institutional contexts and expected knowledge base, it reflects a temporary equilibrium reached among tensions. Such equilibrium may be disrupted by changes in any of the various factors. For example, when the institutional context changes through local policy adjustment (e.g., promulgation of school accountability testing), new tensions for TALiP arise. Accordingly, teachers may need to make new compromises to reach a balancing point among the tensions generated by context, knowledge, and conceptions.

These compromises may take many forms. They can be internal to assessment itself, such as compromises made between multiple purposes that a single assessment task is intended to serve (Pellegrino, Chudowsky, & Glaser, 2001). They can be internal to the teacher, such as tensions between beliefs in using assessment for improved teaching and learning and a new policy using assessment for school accountability and accreditation purposes. Therefore, teachers need to make compromises in their decision-making and action-taking about whether to treat assessment as a quality assurance mechanism or learning-oriented tool. Compromises can also be required because of contextual factors external to assessment (e.g., class size, teaching schedule, etc.) that impact on the kinds of assessment practices that are feasible (Carless, 2015). For example, despite believing that performance assessment is more beneficial to student learning than paper-and-pencil tests, a teacher may have to compromise such beliefs when confronted with a class of 60+ students, requiring new strategies of assessment compatible with personal conceptions, while complying with the constraints of the context (i.e., large-class setting).

Obviously TALiP is neither a static concept nor an idealized body of knowledge and skills. Rather, it is a dynamic, complex entity combining teachers' assessment knowledge, their conceptions of assessment, and their responses to the external contexts embedded with actual constraints and affordances in the environment. TALiP reflects the true realities concerning how teacher AL is enacted in the classroom. The advancement of one's TALiP is not straightforward; it needs to be driven by principles of teacher learning.

3.5 Teacher learning as the impetus for advancing TALiP

The process of becoming assessment literate is fundamentally a transformative, consciousness-evoking one. However, teachers may be content to have conceptions and practices of assessment that are entirely consistent with external contexts without casting doubt on their own practices. They would then end up repeating traditional practices that are inconsistent with research evidence about effective practice or even current policy expectations. The mission then is how to challenge teacher (mis)conceptions of assessment so that legitimacy of tradition is questioned and improved TALiP is achieved.

Teacher learning research (Cochran-Smith & Lytle, 1999) is a major source of input for how teacher AL is construed. Teacher learning is "both a process of active individual construction and a process of enculturation into the ... practices of wider society" (Cobb, 1994, p. 13). Teacher learning is the impetus for TALiP for two reasons. First, the dynamic, interactive, and contingent nature of classroom-based assessment changes the environment in which AL is operationalized. The effectiveness of classroom assessment largely depends on various dynamics at work within and outside the classroom, such as forms of assessment tasks, cognitive and emotional preparedness of students, the teacher's familiarity with his/her students and course materials, characteristics of the course, and the physical environment (Carless, 2015). If teachers do not seek for a better understanding of assessment during their day-to-day practices, it is unlikely that they will be able to make better compromises among tensions arising from these dynamics, nor will student learning be facilitated. Second, teachers have their own concerns and queries arising from their classroom practices (Brookhart, 2002), which may lead to reflection and changes. Teacher learning can thus take place in the classroom and other workplace communities if teachers are aware of the reciprocal interactions among many processes, mechanisms, and actions arising from assessment activities (Koh, 2011).

To be specific, there are two main ways for teacher learning to occur: (a) reflective practice (Schön, 1983) and (b) participation in community activities (Westheimer, 2008). Although reflective practice is not a novel idea, research evidence consistently shows that it has the power to help teachers understand the links between what they do and how they might improve their effectiveness (Cornish & Jenkins, 2012; DeLuca, 2012). Without reflection, teaching may be prompted by impulse, intuition, or routine. Reflection requires more than just thinking; it is a progression from spontaneous interpretation of an experience to an intellectualizing of it (Dewey, 1933). Reflective practice in assessment is a recursive process that is not only retrospective in examining what was done and how, but also prospective in planning for new actions or changes (Schön, 1983). Such reflective practice is critical to advancing TALiP because it provides opportunities for "assumption hunting" (Brookfield, 1995, p.21) concerning pre-conceptions of assessment, leading to "cognitive conflict" (Cobb, Wood, & Yackel, 1990, p.127). These cognitive conflicts may open up a "change-provoking disequilibrium" (Woolfolk, Hoy, & Davis, 2009, p.645) in which the teacher may realize the need to unlearn and relearn about assessment.

Collectively, teachers need to participate in assessment-related activities (e.g., moderation, rubric development, assessment task design) within real or virtual communities (Adie, 2013; Lukin, et al., 2004). Participation in community activities engages teachers in professional conversations about their assessment practices, offering opportunities to understand alternative thinking and practice of assessment, and allowing them to defend their own conceptions and negotiate their ideas with colleagues (Fleer, 2015; Leahy & Wiliam, 2012). Such engagement may lead teachers to make subsequent changes in their day-to-day assessment practices and new compromises in the interest of student learning. In this sense, teacher learning can be

considered as the impetus for effecting changes in assessment and advancing TALiP.

3.6 Teacher as 'assessor' identity (re)construction as the ultimate goal

An interesting challenge in educational assessment is the position of the teacher not only as the instructor but also as an assessor of learning. Traditionally, decisions about student qualifications or certification were made by an external body, meaning that the teacher was only the instructor. Thus the teacher's traditional role involved drilling and repeatedly preparing students for success on an externally produced examination. In this sense, the teacher's professional identity was free of obligations, duties, or authority associated with evaluation. However, under formative assessment policies, teachers assess students to make pedagogical decisions, and these decisions are sometimes used as part of formal certification processes. Under such circumstances, the teacher's role and identity now includes 'assessor.' Understanding this role and successfully integrating it into the teacher's pedagogical function requires a new way of understanding what it means to be a teacher.

Sitting atop the TALiP pyramid is the teachers' identity (re)construction as assessors, which represents the ultimate goal of TALiP. There are three reasons for this positioning. First, teacher identity (re)construction exists throughout their professional life. Pre-service teachers need to undergo a shift in identity as they progress through teacher education programs and assume teaching responsibility in school contexts full of complexity. They need to transform their identities from being students who are assessed into teachers who plan, implement and evaluate students using assessment tools. In-service teachers need to constantly negotiate their roles as assessors within interactions with others (Cowie, et al., 2014; Scarino, 2013), especially when innovations are being implemented.

Second, teacher identity (re)construction is a natural by-product of teacher learning. For example, through reflection and participation, teachers can become more in tune with their sense of self and gain a deeper understanding of how this self fits in with larger contexts involving others (Beauchamp & Thomas, 2009; Beijaard, Meijer, & Verloop, 2004). Thus, learning of assessment, either individually or collectively, leads teachers to make sense of their work in assessment and themselves as assessors. With deeper engagement in assessment, teachers come to understand the value of assessment and how their educational decisions in assessment can make a positive difference to student learning. The awareness of being assessors who are responsible for both accountability and learning purposes of assessment may empower teachers with greater autonomy and resources, assist them in (re)claiming their ownership of assessment practices, and open up more opportunities in which their AL can be recognized by the community (Wyatt-Smith, et al., 2010).

Third, the dynamic and evolving nature of TALiP suggests that teacher identity as an assessor is by no means static. It is a complex and dynamic entity that brings various cognitive, affective, and socio-cultural dimensions into play (Beauchamp & Thomas, 2009). As changes are introduced to curriculum, pedagogy or administrative procedures, assessment itself is impacted, and identity construction as assessors

connects the centrality of teacher as agents of assessment with ubiquitous influences from contexts.

Identity construction can play a powerful role in provoking teachers' thinking, promoting dialogues and shared meaning, and shaping their professional judgments (Adie, 2013; Scarino, 2013; Wyatt-Smith, et al., 2010). If a teacher simply regards him or herself as a teacher, teaching practice may cease when the teaching content is delivered, coached, or scaffolded. In contrast, if the teacher is conscious of also being an assessor, he/she ought to mindfully plan and implement assessment activities to measure whether what was taught has been learnt, and to give quality feedback that helps students make progress. In addition, if teachers are placed within a community where they need to negotiate and rework their identities as assessors, their traditional decision-making processes about student learning may be challenged. Instead of treating poor assessment results as a function of student effort or ability, with an assessor identity teachers may reflect on their practices, reconsider alternative ways of practice, relearn and unlearn assessment knowledge, trial new assessment approaches or strategies, scrutinize their conceptions of assessment, and eventually calibrate or change their assessment practices to facilitate student learning. In other words, if teachers start to construct their identities as assessors, it is likely that they will change their practices and conceptions of assessment accordingly.

It should be noted, however, that there is no 'ideal' assessor identity. Teachers who have a clearer sense of their own identity as assessors are likely to be better authors of their own assessment practices, more likely to engage in self-reflection, and in a better position to integrate others' perspectives into their own values. In this way, their identity (re)construction as assessors empowers teachers with enhanced agency to make more justified compromises in their assessment practices, which may lead to improved TALiP.

3.7 The relationships among TALiP components

The components within the framework of TALiP are interrelated as depicted by the multi-directional arrows (Figure 1). The upward arrows from the knowledge base represent varying training opportunities in which principles are disseminated to teachers. This knowledge goes through the interpretive framework of teacher conceptions of assessment; usually, only knowledge compatible with teachers' current conceptions is accepted and used by teachers. Meanwhile, teachers' conceptions of assessment, being mainly built upon prior experiences, are both receptive to, and conflicting with, influences from the 'boundaries' set by micro- and macro-contexts. Teachers have to make constant compromises to strike a temporary equilibrium among these tensions within a particular context. Such compromises are not necessarily in the best interest of student learning or teachers' own work because of the constraints of boundaries placed upon autonomy and professionalism.

Such equilibrium, however, will be disrupted by every new wave of tensions. Yet the advancement of the TALiP can only be possible through teacher learning. Both teachers' reflective practice and participation in community-based assessment

activities will generate affordances for teachers to (re)construct their identities as assessors. The double-headed arrows between teacher learning and their identity construction indicates a reciprocal relationship, in which learning and identity construction generate opportunities for and exert influence upon each other.

Teachers' identity (re)construction facilitates changes in how compromises are made, and results in new insights and conceptions of assessment. With such new insights, teachers may gain more autonomy to exercise their assessment expertise and use their time-tested practice or findings to contribute to the AL knowledge base itself. In this sense, the knowledge base can be dynamic and evolving. Old principles will be replaced by new ones derived from implementing AL in practice, and once-correct theories or time-tested assessment practices may be proven wrong by conflicting evidence.

The TALiP framework is cyclical in nature. Changes occurring in one component are contingent upon changes in another. Thus, the framework has multi-directional flows. On one hand, research-based theories and principles flow from the bottom (i.e., the knowledge base) to the top (i.e., teachers' identity construction as assessors); on the other hand, context-tested assessment practices and research-based evidence flow from the top to the bottom to renew and reformulate the knowledge base. In this way, not only is a sharable knowledge base and the centrality of teachers recognized, but the complex processes by which TALiP is enacted are also identified.

4. Conclusions

After reviewing AL studies from the past three decades, the present paper proposes a conceptual framework of TALiP that bridges the distance between research on educational assessment and teacher education, each with its own distinctive aims and limited scope, and yet sharing the same concern for helping teachers acquire AL and implement high quality assessment. This framework contributes to the theorization of AL by moving the field beyond a focus on the knowledge base to consideration of a situated, dynamic, and evolving system in which teachers constantly make compromises among competing tensions, suggesting that the improvement of teacher AL is a systematic enterprise that depends on forming a virtuous circle of TALiP with joint efforts from appropriate stakeholders.

As an initial attempt of delineating different layers and stages of AL development, we propose a working definition of TALiP that breaks down the dichotomy of being either assessment 'literate' or 'illiterate,' and places it instead on a continuum with different levels of mastery, contingent upon the context in which assessment is conducted.

TALiP consists of three levels of mastery. First is a basic mastery of educational assessment knowledge, which includes the fundamental principles of the 'what', 'why', and 'how,' without which teachers cannot engage with assessment at a deeper level. Second is an internalized set of understanding and skills of the interconnectedness of assessment, teaching, and learning. Unlike the 'should-do' kind of knowledge indicated by the first level, this is a more personal perception of how assessment should be, formed among the tensions between theoretical knowledge and

teachers' own conceptions of assessment. Third is a self-directed awareness of assessment processes and one's own identity as an assessor. Such awareness allows teachers to accommodate and translate assessment policies and principles into their classroom realities and institutional contexts while driving them to reflect on their assessment practices and to gain new insights.

The presumed complexity of the framework does not mean that regularity and generalized understandings of TALiP are impossible. It is true that TALiP depends on how teachers make compromises, which varies by teacher and context. It is also true that cases of teachers who have become reflective and resourceful assessors in one context may be quite different from those who have done so in another. However, there are possible generalizations about how TALiP advances that would be applicable to all teachers across contexts. According to the proposed framework, assessment literate teachers are those who constantly reflect on their assessment practice, participate in professional activities concerning assessment in communities, engage in professional conversations about assessment, self-interrogate their conceptions of assessment, and seek for resources to gain a renewed understanding of assessment and their own roles as assessors. Ultimately, we need more empirical studies across contexts to support this framework.

This framework can be used as an operationalized model for researching teacher AL. Each component within TALiP can be used as a point of entry for conducting AL research. On the level of the knowledge base, future studies could be conducted to: 1) explore the relationship of 'common-core' and 'discipline-specific' assessment knowledge with a view towards establishing aspects of AL that are unique to various disciplines or subjects; 2) compare the professional standards across contexts or along the historical timeline (e.g., Deluca, LaPointe-McEwan, et al., 2015); or 3) interrogate how contextualized this knowledge base needs to be given the different contexts for assessment (e.g. exam-oriented culture in East Asia vs. assessment for learning culture in many Western countries). On the teacher conception level, we concur with Hill, Ell, et al. (2014) that longitudinal studies which explore how teachers' conceptions of assessment influence their AL development, ideally from school student status to pre- and on into in-service, are much needed to better understand the complex process of becoming assessment literate. In particular, evidence of how assessment training can change teacher conceptions into a more learning-facilitating direction would be welcome (e.g., DeLuca, Chavez, Bellara, et al., 2013). In addition, inquiries into the way various levels of context facilitate or inhibit teacher AL development are wanted. Both successful and unsuccessful cases about in-service teachers learning about assessment in a workplace community would be most beneficial to facilitating on-the-job assessment education. Last but not least, future research needs to address the issues of teacher learning and teacher identity construction as an assessor, given that these areas have long been neglected in teacher AL studies (e.g., Cowie, et al., 2014). Research assuming an ethnographic stance to observe teachers' classroom practices and listen to different voices of stakeholders would be particularly helpful to unravel the complexities of these issues. In sum, this model provides researchers with clues to follow in terms of where to start and what possible positioning their chosen

topic is within the larger landscapes of teacher education and assessment research.

The framework has a number of implications for both policy and practice of teacher assessment education. For policy development, the different layers of influence and their interrelationships within the framework may be helpful in terms of resource allocation. For example, given that the knowledge base is a necessary, but insufficient, condition for teacher AL, policy makers need to first devise standards that reflect both contemporary development of assessment research and culturally valued assessment heritage. These standards may promote teacher AL by acting as thresholds in teacher licensure and accreditation. For those countries where such standards are already existent, policy makers may go to greater lengths along the flow of the framework, preferably collaborating with school administrators, to provide follow-up interventions, such as diagnosing specific problems in pre- or in-service assessment education in specific contexts, and working out supplementary mechanism to achieve better implementation of professional standards or policies.

For pre-service teacher education, there are implications for the content and structures of assessment education. For instance, given the dynamic nature of the AL knowledge base, teacher educators need to constantly update their assessment curriculum with insights from latest assessment research findings and time-tested practices. Also, given the importance of teacher conceptions and contexts in shaping teacher AL, they may need to include in the curriculum scenarios that are both cognitively challenging and emotionally appealing (e.g., how do you assess a hard-working student whose work is poor quality and a lazy student whose work is high quality) so that student teachers will reflect upon their own conceptions and practices of assessment. They may also need to increase internship opportunities for student teachers to experience how to connect and make compromises between theory and practice.

For in-service teacher education, the framework reminds school administrators of both the cognitive and emotional support that teachers need. As most of the conventional assessment training is cognitively oriented, TALiP suggests that school administrators show sufficient attention, care, and patience to understand their teachers' conceptions. In addition, they may need to create conditions for facilitating individual or collective teacher learning of assessment and identity construction as assessors. Such conditions may have a wide range of forms, such as building a lounge where teachers can sit comfortably together to share their assessment practices, setting incentives for encouraging teachers to keep a reflective journal of assessment practice, or building up a cyber-space for teachers to 'complain' about the enacting assessment policies.

Last but not least, this framework provides teachers with both general predictions of an array of challenges that may confront them in advancing their AL and workable solutions to cope with such challenges. It also reminds teachers that these challenges are oftentimes not linear, but occur simultaneously. It will assure them that feeling overwhelmed and incompetent is a natural part of the growing process of becoming assessment literate. This may, to a great extent, alleviate teacher anxiety and enhance resilience. Using the framework as a checklist, teachers may know where they are,

where they need to be, and how best to proceed on their developmental trajectories.

References

Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do. In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The Professional knowledge base of science teaching* (Vol. 12, pp. 205-221). Dordrecht: Springer.

Adie, L. (2013). The development of teacher assessment identity through participation in online moderation. *Assessment in Education: Principles, Policy & Practice, 20*(1), 91-106.

Airasian, P. W. (1991). Perspectives on measurement instruction. *Educational Measurement: Issues and Practice, 10*(1), 13–16, 26.

Airasian, P. W., & Russell, M. K. (2008). *Classroom assessment* (6th ed.). New York: McGraw-Hill.

Ajzen, I. (2005). *Attitudes, personality and behavior* (2nd ed.). New York: Open University Press.

Alkharusi, H. (2011). Psychometric properties of the teacher assessment literacy questionnaire for preservice teachers in Oman. *Procedia - Social and Behavioral Sciences, 29*, 1614-1624.

Alkharusi, H., Kazem, A. M., & Al-Musawai, A. (2011). Knowledge, skills, and attitudes of preservice and inservice teachers in educational measurement. *Asia-Pacific Journal of Teacher Education, 39*(2), 113-123.

Allal, L. (2013). Teachers' professional judgement in assessment: A cognitive act and a socially situated practice. *Assessment in Education: Principles, Policy & Practice, 20*(1), 20-34.

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

Arksey, H., & O'Malley, L. (2005) Scoping studies: Towards a methodological framework. *International Journal of Social Methodology, 8*(1), 19-32.

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). Cleveton, UK: Multilingual Matters.

Barnes, N., Fives, H., & Dacey, C. M. (2015). Teachers' beliefs about assessment. In H. Fives & M. G. Gill (Eds.), *International handbook of research on teacher beliefs* (pp. 284-300). New York: Routledge.

Beauchamp, C., & Thomas, L. (2009). Understanding teacher identity: An overview of issues in the literature and implications for teacher education. *Cambridge Journal of Education, 39*(2), 175-189.

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(2), 107-128.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in*

Education, 5(1), 7-74.

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

Brookfield, S. D. (1995). *Becoming a critically reflective teacher.* San Francisco, CA: Jossey-Bass.

Brookhart, S. M. (1999). Teaching about communicating assessment results and grading. *Educational Measurement: Issues and Practice, 18*(1), 5–14.

Brookhart, S. M. (2002). What will teachers know about assessment, and how will that improve instruction. In R. W. Lizzitz & W. D. Schafer (Eds.), *Assessment in educational reform: Both means and ends* (pp. 2-17). Boston, MA: Allyn & Bacon.

Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*(1), 3-12.

Brown, G. T. L. (2008a). Assessment literacy training and teachers' conceptions of assessment. In C. Rubie-Davies & C. Rawlinson (Eds.), *Challenging thinking about teaching and learning* (pp. 285-302). New York: Nova Science.

Brown, G. T. L. (2008b). *Conceptions of assessment: Understanding what assessment means to teachers and students.* New York: Nova Science.

Brown, G. T. L., Irving, E., & Keegan, P. (2014). *An introduction to educational assessment, measurement, and evaluation: Improving the quality of teacher-based assessment.* New Zealand: Dunmore Publishing Limited.

Brown, J. D., & Bailey, K. M. (2008). Language testing courses: What are they in 2007? *Language Testing, 25*(3), 349-384.

Buck, G. A., & Trauth-Nare, A. E. (2009). Preparing teachers to make the formative assessment process integral to science teaching and learning. *Journal of Science Teacher Education, 20*(5), 475–494.

Campbell, C. (2013). Research on teacher comptence in classroom assessment. In J. H. McMillan (Ed.), *Sage handbook of research on classroom assessment* (pp. 71-84). Los Angeles: SAGE.

Campbell, C., & Collins, V. L. (2007). Identifying essential topics in general and special education introductory assessment textbooks. *Educational Measurement: Issues and Practice, 26*(1), 9–18.

Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. London: Routledge.

Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. London: Routledge.

Cobb, P. (1994). Where is the mind? Constructivist and sociocultural perspectives on mathematical development. *Educational Researcher, 23*(7), 13-20.

Cobb, P., Wood, T., & Yackel, E. (1990). Classrooms as learning environments for teachers and researchers. In R. B. Davis, C. A. Mayer & N. Noddings (Eds.), *Constructivist views on the teaching and learning of mathematics* (pp. 125-146). Reston, VA: National Council of Teachers of Mathematics.

Cochran-Smith, M., & Lytle, S. L. (1999). Relationships of knowledge and practice: Teacher learning in communities. *Review of Research in Education, 24*(1), 249-305.

Cornish, L., & Jenkins, K. A. (2012). Encouraging teacher development through embedding reflective practice in assessment. *Asia-Pacific Journal of Teacher Education, 40*(2), 159-170.

Cowie, B., Cooper, B., & Ussher, B. (2014). Developing an identity as a teacher-assessor: Three student teacher case studies. *Assessment Matters, 7*, 64–89.

Crossman, J. (2007). The role of relationships and emotions in student perceptions of learning and assessment. *Higher Education Research and Development, 26*(3), 313-327.

DeLuca, C. (2012). Preparing teachers for the age of accountability: Toward a framework for assessment education. *Action in Teacher Education, 34*(5/6), 576-591.

DeLuca, C., & Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education curriculum. *Journal of Teacher Education, 64*(4), 356-372.

DeLuca, C., Chavez, T., Bellara, A., & Cao, C. (2013). Pedagogies for preservice assessment education: Supporting teacher candidates' assessment literacy development. *The Teacher Educator, 48*(2), 128-142.

DeLuca, C., Chavez, T., & Cao, C. (2013). Establishing a foundation for valid teacher judgement on student learning: The role of pre-service assessment education. *Assessment in Education: Principles, Policy & Practice, 20*(1), 107-126.

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

DeLuca, C., Klinger, D., Pyper, J., & Woods, J. (2015). Instructional rounds as a professional learning model for systemic implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice, 22*(1), 122-139.

DeLuca, C., Klinger, D., Searle, M., & Shulha, L. (2010). Developing a curriculum for assessment education. *Assessment Matters, 2*, 133-155.

DeLuca, C., & Lam, C. Y. (2014). Preparing teachers for assessment within diverse classrooms: An analysis of teacher candidates' conceptualizations. *Teacher Education Quarterly, 41*(3), 3-24.

DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*. Advance online publication. doi: 10.1007/s11092-015-9233-6.

Dewey, J. (1933). *How we think*. Buffalo, NY: Prometheus Books.

Ell, F., Hill, M. F., & Grudnoff, L. (2012). Finding out more about teacher candidates' prior knowledge: implications for teacher educators. *Asia-Pacific Journal of Teacher Education, 40*(1), 55-65.

Engelsen, K. S., & Smith, K. (2014). Assessment literacy. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *The enabling power of assessment: Designing assessment for quality learning* (pp. 140-162). New York: Springer.

Fan, Y.-C., Wang, T.-H., & Wang, K.-H. (2011). A Web-based model for developing assessment literacy of secondary in-service teachers. *Computers & Education, 57*(2), 1727-1740.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham & T. Urdan (Eds.), *APA educational psychology handbook: Individual differences and cultural and contextual factors* (Vol. 2, pp. 471-499). Washington, DC: American Psychological Association.

Fleer, M. (2015). Developing an assessment pedagogy: The tensions and struggles in re-theorising assessment from a cultural-historical perspective. *Assessment in Education: Principles, Policy & Practice, 22*(2), 224-246.

Forsberg, E., & Wermke, W. (2012). Knowledge sources and autonomy: German and Swedish teachers' continuing professional development of assessment knowledge. *Professional Development in Education, 38*(5), 741-758.

Frey, N., & Fisher, D. (2009). Using common formative assessments as a source of professional development in an urban American elementary school. *Teaching and Teacher Education, 25*(5), 674-680.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113-132.

Gotch, C. M., & French, B. F. (2014). A systematic review of assessment literacy measures. *Educational Measurement: Issues and Practice, 33*(2), 14-18.

Gottheiner, D. M., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education, 23*(5), 531-557.

Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through preservice teacher education. *Teaching and Teacher Education, 21*(6), 607-621.

Green, T. F. (1971). *The activities of teaching*. New York: McGraw-Hill.

Greenberg, J., & Walsh, K. (2012). *What teacher preparation programs teach about K-12 assessment: A review*. Washington, DC: National Council on Teacher Quality.

Gu, P. Y. (2014). The unbearable lightness of the curriculum: What drives the assessment practices of a teacher of English as a Foreign Language in a Chinese secondary school. *Assessment in Education: Principles, Policy & Practice, 21*(3), 286-305.

Gullickson, A. R. (1993). Matching measurement instruction to classroom-based evaluation: Perceived discrepancies, needs, and challenges. In S. L. Wise (Ed.), *Teacher training in measurement and assessment skills* (pp. 1-25). Lincoln, Neb: Buro Institute of Mental Measurement, University of

Nebraska-Lincoln.

Gunn, A. C., & Gilmore, A. (2014). Early childhood initial teacher education students' learning about assessment, *Assessment Matters, 7*, 24–38.

Hailaya, W., Alagumalai, S., & Ben, F. (2014). Examining the utility of Assessment Literacy Inventory and its portability to education systems in the Asia Pacific region. *Australian Journal of Education, 58*(3), 297-317.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Herman, J., Osmundson, E., Dai, Y., Ringstaff, C., & Timms, M. (2015). Investigating the dynamics of formative assessment: Relationships between teacher knowledge, assessment practice and learning. *Assessment in Education: Principles, Policy & Practice, 22*(3), 344-367.

Hill, M. F., Cowie, B., Gilmore, A., & Smith, L. F. (2010). Preparing assessment-capable teachers: What should pre-service teachers know and be able to do? *Assessment Matters, 2*, 44-64.

Hill, M. F., Ell, F., Grudnoff, L., & Limbrick, L. (2014). Practise what you preach: Initial teacher education students learning about assessment. *Assessment Matters, 7*, 90–112.

Hill, M. F., Gunn, A., Cowie, B., Smith, L. F., & Gilmore, A. (2014). Preparing primary and early childhood initial teacher education students to use assessment in teaching. *Assessment Matters, 7*, 4–23.

Howley, M. D., Howley, A., Henning, J. E., Gillam, M. B., & Weade, G. (2013). Intersecting domains of assessment knowledge: School typologies based on interviews with secondary teachers. *Educational Assessment, 18*(1), 26-48.

Impara, J. C., Divine, K. P., Bruce, F. A., Liverman, M. R., & Gay, A. (1991). Does interpretive test score information help teachers? *Educational Measurement: Issues and Practice, 10*(4), 16-18.

Impara, J. C., & Plake, B. S. (1995). Comparing counselors', school administrators', and teachers' knowledge in student assessment. *Measurement & Evaluation in Counseling & Development, 28*(2), 78-87.

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

Inbar-Lourie, O. (2013). Language assessment literacy. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 2923-2931). Oxford: Blackwell.

Jeong, H. (2013). Defining assessment literacy: Is it different for language testers and non-language testers? *Language Testing, 30*(3), 345-362.

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.

Joint Committee on Standards for Education Evaluation (JCSEE). (2015). Classroom assessment standards for PreK-12 teachers. [Kindle version]. Retrieved from https://www.amazon.com/Classroom-Assessment-Standards-PreK-12-Teachers-ebook/dp/B00V6C9RVO

Kiomrs, R., Abdolmehdi, R., & Rashidi, N. (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.

Koh, K. H. (2011). Improving teachers' assessment literacy through professioal development. *Teaching Education, 22*(3), 255-276.

Lam, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing, 32*(2), 169-197.

Leahy, S., & Wiliam, D. (2012). From teachers to schools: Scaling up professional development for formative assessment. In J. Gardner (Ed.), *Assessment and learning* (pp. 49-72). London: SAGE.

Leighton, J. P., Gokiert, R. J., Cor, M. K., & Heffernan, C. (2010). Teacher beliefs about the cognitive diagnostic information of classroom- versus large-scale tests: Implications for assesment literacy. *Assessment in Education: Principles, Policy & Practice, 17*(1), 7-21.

Leung, C. (2014). Classroom-based assessment issues for language teacher education. In A. J. Kunnan (Ed.), *The companion to language assessment* (Vol. 3, pp. 1510-1519). Chichester, UK: Wiley Blackwell.

Levy-Vered, A., & Alhija, F. N.-A. (2015). Modelling beginning teachers' assessment literacy: The contribution of training, self-efficacy, and conceptions of assessment. *Educational Research and Evaluation, 21*(5-6), 378-406.

Lian, L. H., Yew, W. T., & Meng, C. C. (2014). Enhancing Malaysian teachers' assessment literacy. *International Education Studies, 7*(10), 74-81.

Lomax, R. G. (1996). On becoming assessment literate: An initial look at pre-service teachers' beliefs and practices. *The Teacher Educator, 31*(4), 292-303.

Lortie, D. (1975). *Schoolteacher: A sociological study*. London: University of Chicago Press.

Lukin, L. E., Bandalos, D. L., Eckhout, T. J., & Mickelson, K. (2004). Facilitating the development of assessment literacy. *Educational Measurement: Issues and Practice, 23*(2), 26-32.

Lyon, E. G. (2013). Conceptualizing and exemplifying science teachers' assessment expertise. *International Journal of Science Education, 35*(7), 1208-1229.

Maclellan, E. (2004). Initial knowledge states about assessment: Novice teachers' conceptualizations. *Teaching and Teacher Education, 20*(5), 523–535.

McMillan, J. H. (2001). *Essential assessment concepts for teachers and administrators*. Thousand Oaks, California: Corwin Press, Inc.

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. (2004). Secondary teachers' assessment literacy: Does classroom experience make a difference? *American Secondary Education, 33*(1), 49-64.

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

Nichols, S. L., & Harris, L. R. (in press). Accountability assessment's effects on teachers and schools. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of social and human conditions in assessment* (pp.40-56). New York: Routledge.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Thousand Oaks, CA: SAGE.

Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). Boston: Pearson Education, Inc.

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.

Pajares, M. F., & Graham, L. (1998). Formalist thinking and language arts instruction: Teachers' and students' beliefs about truth and caring in the teaching conversation. *Teaching and Teacher Education, 14*(8), 855-870.

Patton, M. Q. (2002). *Qualitative research & evaluation methods*. Thousand Oaks, CA: SAGE.

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

Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A., & McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods, 5*(4), 371-385.

Plake, B. S., Impara, J. C., & Fager, J. J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice, 12*(4), 10-12, 39.

Popham, W. J. (1991). Appropriateness of teachers' test-preparation practices. *Educational Measurement: Issues and Practice, 10*(4), 12-15.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11.

Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator, 46*(4), 265-273.

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

Rorsyth, R., Cullen, R., Ringan, N., & Stubbs, M. (2015). Supporting the development of assessment literacy of staff through institutional process change. *London Review of Education, 13*(3), 34-41.

Sato, M., Wei, C. R., & Darling-Hammond, L. (2008). Improving teachers' assessment practices through professional development: The case of National Board Certification. *American Educational Research Journal, 45*(3), 669-700.

Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing, 30*(3), 309-327.

Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice, 10*(1), 2-6,12.

Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into Practice, 32*(2),

118-126.

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

Schneider, M. C., & Randel, B. (2010). Research on characteristics of effective professional development programs for enhancing educators' skills in formative assessment. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 251-276). New York: Routledge.

Schön, D. (1983). *The reflective practitioner: How professional think in action*. New York: Basic Books.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Preservice teachers' assessment literacy. *Journal of Science Teacher Education, 22*(4), 371-391.

Smith, K. (2011). Professional development of teachers—A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation, 37*(1), 55-61.

Smith, L. F., & Galvin, R. (2014). Toward assessment readiness: An embedded approach in primary initial teacher education. *Assessment Matters, 7*, 39–63.

Smith, L. F., Hill, M. F., Cowie, B., & Gilmore, A. (2014). Preparing teachers to use the enabling power of assesment. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 418-445). Dordrecht: Springer.

Stiggins, R. J. (1991a). Assessment literacy. *Phi Delta Kappan, 72*, 534-539.

Stiggins, R. J. (1991b). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice, 10*(1), 7-12.

Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-245.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23-27.

Stiggins, R. J. (2010). Essential formative assessment competencies for teachers and school leaders. In H. L. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 233-250). New York, NY: Taylor & Francis.

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

Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing, 30*(3), 403-412.

Tierney, R. D. (2013). Fairness in classroom assessment. In J. H. McMillan (Ed.), *SAGE handbook of research on classroom assessment* (pp. 125-144). Thousand Oaks, CA: SAGE.

Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of an European study. *Language Assessment Quarterly, 11*(4), 374-402.

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.

Westheimer, J. (2008). Learning among colleagues: Teacher community and the shared enterprise of education. In M. Cochran-Smith, S. Feiman-Nemser & D. J. McIntyre (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (pp. 756-783). New York, NY: Routledge.

Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualizing teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher, 40*(2), 241-256.

Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education, 42*(1), 37-42.

Woolfolk, H. A., Hoy, W. K., & Davis, H. A. (2009). Teachers' self-efficacy beliefs. In K. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 627-655). Mahwah, NJ: Lawrence Erlbaum.

Wyatt-Smith, C., Klenowski, V., & Gunn, S. (2010). The centrality of teachers' judgement practice in assessment: A study of standards in moderation. *Assessment in Education: Principles, Policy & Practice, 17*(1), 59-75.

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), 493-513.

Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education, 16*(4), 323-342.