

Exploring Novice EFL Teachers' Classroom Assessment Literacy Development: A Three-Year Longitudinal Study

Hao Xu¹

Published online: 5 June 2017
© De La Salle University 2017

Abstract This three-year longitudinal study examines the development of four Chinese novice EFL teachers' classroom assessment literacy, which refers to teachers' knowledge of and abilities to conduct classroom assessment and is reflected in their planned and improvised assessment practices. The dataset includes 26 classroom observations, 48 interviews, and 171 participants' journal entries. Data analysis following the paradigmatic analytic procedures reveals three developmental stages for novice teachers' classroom assessment literacy: the first stage features the acquisition of practical techniques for planned classroom assessment; the second stage allows novice teachers to gain a growing awareness to connect planned classroom assessment to teaching objectives, but the usefulness of assessment results in teaching adjustments was only confined to an improved plan of a future teaching situation; in the third stage, novice teachers are more inclined to conduct improvised formative assessment in teaching which shows immediate assessment-driven teaching improvements in response to dynamic classroom interactions.

Keywords Assessment literacy · Classroom assessment · Novice teacher · EFL teacher · Teacher development

Introduction

Classroom assessment provides teachers with crucial information regarding the accomplishment of teaching objectives and the difficulties students experience in their learning process (Graham 2005; Phye 1997). Therefore, it is regarded as an indispensable component of teachers' teaching competence, which can be conceptualised as *classroom assessment literacy*. Assessment literacy generally refers to teachers' knowledge of and abilities to conduct assessment (Coombe et al. 2012; Inbar-Lourie 2013; Lam 2014; Popham 2009, 2011; Stiggins 1991), and accordingly, classroom assessment literacy can be regarded as the use of assessment literacy in classroom assessment, which is a most common type of assessment for teachers.

Classroom assessment literacy, however, is not simply assessment literacy enacted in classroom teaching as a specific domain. It is fundamentally concerned about the contingency between assessment practice, learning/teaching goal and learner performance (Black 2009; Black and William 2009; Johns and Gerard 1967; Shepard 2005; William and Thompson 2007). Therefore, classroom assessment literacy can be defined as teachers' knowledge of assessment in general and of the contingent relationship between assessment, teaching, and learning, as well as abilities to conduct assessment in the classroom to optimise such contingency. Because classroom scenarios feature huge dynamicity, such contingent relationships may also change rapidly and frequently (Rea-Dickins 2001). This vastly makes classroom assessment both unplanned and non-plannable. Teachers often improvise in their assessment practices, besides what they have planned beforehand (Erickson 2010). In light of this, it is also important to examine more closely how *planned* and *improvised* assessment practices interact, and how they change over

✉ Hao Xu
xuhaokent@bfsu.edu.cn; xuhaokent@139.com

¹ National Research Centre for Foreign Language Education, Beijing Foreign Studies University, Room 655, Zong He Lou, West Campus, 19 Xi San Huan Bei Road, Beijing 100089, People's Republic of China

time in terms of proportion and effectiveness in the development of teachers' classroom assessment literacy.

As to assessment literacy in itself, previous studies basically focused on (1) teacher' knowledge of and abilities to use assessment methods or techniques (Black and William 1998), and (2) knowledge of and abilities to use the assessment outcome to support and guide practice (Cheng et al. 2004; Inbar-Lourie 2008; Jeong 2013; Stiggins 1988; Stiggins and Conklin 1992). Therefore, assessment literacy concerns both a dimension of *methods/techniques* and a dimension of *utilisability*. Previous research examined the various ways teachers employ assessment methods/techniques such as designing, selecting, scoring, evaluating, and revising (see Stiggins 1988 for a comprehensive review). More recent research shows an increasing interest in how assessment results are used to inform pedagogical practices, including formulating teaching objectives, selecting contents and activities, employing strategies to deliver the contents, judging whether the objectives have been achieved, and identifying improvement needs (Clark-Gareca 2016; Hailaya et al. 2014; König et al. 2014). This line of research on the utilisability dimension shows that assessment practices and the teaching adjustments based on them constitute an iterative system, enabling incremental improvements in classroom teaching (Rea-Dickins 2007; Xu and Liu 2009). Xu and Brown (2016) systematically review this notable development of assessment literacy research and propose a reconceptualisation of assessment literacy as Teacher Assessment Literacy in Practice (TALiP), which further acknowledges that the more essential part of teacher assessment literacy does not lie in the methods/techniques that teachers are able to employ, but in how the results produced by these methods/techniques are used. In short, existing literature has provided a useful framework for analysing teacher assessment literacy that includes both dimensions of *methods/techniques* and *utilisability*. However, very few studies have investigated how teachers' assessment literacy develops as a linking system between assessment and teaching. It looks even more interesting to explore when and how the initial awareness of and practices toward such linkage between assessment practices and teaching adjustments emerge.

Based on the shared understandings manifested in existing literature on assessment literacy in general and that on assessment literacy in various disciplinary subjects including English as a Foreign Language (EFL), the current study aims to delineate how beginning teachers' classroom assessment literacy develops by tracking the changes of both of their planned and improvised assessment practices. In particular, the study intends to reveal how the two types of assessment practices and teaching adjustments interact and how such interaction changes over time. Novice teachers, compared with more experienced teachers, tend

to experience more fluctuations in professional development and demonstrate more conspicuous changes. Therefore, investigation into the situations of novice teachers may provide more insights into the dynamics and complexities of their development. With four EFL teachers as research informants, the study aims to answer the following research questions:

- (1) What assessment practices did the novice EFL teachers undertake in their beginning years?
- (2) How did their assessment practices change over time?

Methodology

Research Design

The current study adopted a qualitative methodology to address the research questions. As to data collection, this study drew on a variety of data sources, i.e. observations of the participants' classroom teaching, individual interviews, and journal protocols written by the participants regarding their reflections on the classroom assessment they conducted. Observation provided data on the participants' actual classroom assessment behaviours, while interview and journal provided data on the participants' cognitions of classroom assessment. The three data sources also provided data triangulation to inspect and validate each other. Data were analysed by the paradigmatic analytic procedures to 'produce taxonomies and categories out of the common elements across the database' (Polkinghorne 1995, p. 5).

Participants

The four female participants, Zhao, Qian, Sun, and Li (all pseudonyms), began to participate in the study as first-year EFL teachers in a junior high school, a senior high school, an elementary school, and a junior high school, respectively, in Beijing. Before that, they had just completed their undergraduate studies in a 4-year B.A. programme for EFL teacher education at a normal university in Beijing, where I got in contact with them. As a teaching fellow at that university, I was asked to supervise the four participants during a 6-week teaching practicum at a junior high school in the last year of their undergraduate studies. However, I did not assume a strong institutional role, because I was only asked to pay regular visits to the school and to check the progress of the teaching practicum, and a professor from the university was responsible for mentoring them in pedagogy. During the teaching practicum, I learnt that they all planned to teach after graduation, so I invited them to be the participants of this novice teacher study, which was a

way of convenient sampling. They approved, and signed written consents which clarified the participants' rights.

Data Collection

Data were collected in Chinese, and translated into English as presented in this paper. In the participants' first 3 years of teaching, I carried out 26 detailed observations on the participants' classroom teaching (6 from Zhao, 5 from Qian, 8 Sun, and 7 from Li). Among them, 16 observations were invited sessions (0 from Zhao, 3 from Qian, 8 from Sun, and 5 from Li) in which the participants hoped to seek advice from me on their teaching. In these observation sessions, I took field notes recording the behaviours of the participants and their students as well as my thoughts throughout the observation.

I also conducted two individual interviews with each participant in every semester at regular intervals, which totalled to 48 interviews. These interviews lasted 40 min–2 h each, and were conducted in places recommended by the participants. Various topics that were related to the participants' experience of and reflection on life as novice teachers were covered in the interviews, but the topic of classroom assessment persisted in all of the interviews. To obtain data regarding classroom assessment, I asked open questions which varied from interview to interview as I tried to contextualise them. Below are a few examples of the questions:

- What did you do to assess your students' learning in the classroom? What was your thinking when you did that?
- How well did your students learn recently? How did you know that?
- Did you do anything if you found out that your students learnt well or had problems in learning?

All of the interview data were transcribed and carefully proofread before analysis.

In addition to being observed in classroom teaching and regularly interviewed, the participants were encouraged to write journals recounting their feelings about and reflections on their teaching and life as novice EFL teachers. Among the 171 journal entries I obtained, which ranged from 75 to 891 Chinese characters, 39 entries contained content related to classroom assessment and were added to the dataset of this study.

Data Analysis

The vast amount of data were analysed in four stages in which paradigmatic analytic procedures were used 'to produce taxonomies and categories out of the common elements across the database' (Polkinghorne 1995, p. 5). First, I read through all of the data in order to get a general

idea of what each participant had done for classroom assessment, why they had done that, and how they had thought of their assessment, with no intention of categorising them. In the mean time, I tried to identify whether there were issues in the data that required me to contact the participants for clarification or confirmation. After the data were clarified or confirmed, I began synthesising the data regarding each participant into a case report in order to have a global understanding of each of the participants. Then I examined each case report to identify possible answers to the research questions. In another reading of the case reports, I compared the answers to the research questions from each of the four participants and tried to identify possible patterns. Repeated readings of the data were also done when I was uncertain about the interpretation of a data extract.

As to specific analysis of participants' assessment practices, I first identified from the observation data an assessment event in which the participant obviously employed certain assessment methods/techniques. Then I reviewed the data in the context to determine the purpose and impact of this assessment practice. I also tried to locate information regarding this assessment event for confirmation or/and clarification in the interview transcripts and journal entries. In sum, the analysis of assessment practices mainly relied on the observation data with interviews and journals serving as a crosscheck.

Data analysis showed particular features of assessment practices that are exclusive to and reoccur within certain periods in the beginning years of the participants' teaching. I therefore tried to identify three such periods as possible developmental stages of their classroom assessment literacy, which will be showcased and delineated in the following section.

Findings and Discussion

This section will first focus on the three developmental stages that emerged from data analysis, describing the novice teachers' classroom assessment practices, and how they impacted on teaching adjustments. Then a discussion will follow.

Building the Tool Box

The first year for these novice teachers could be identified as the first stage of classroom assessment literacy development, because this year featured the acquisition of specific practical techniques for classroom assessment. Below are two extracts from my observations on Qian's and Sun's classes:

At the end of the lesson, Qian gave each student a piece of paper, and told them to translate some phrases from Chinese to English. She read these phrases one by one in Chinese, and the students put down the English translation on the paper. [...] These phrases had just been taught at the beginning of the lesson. (Qian; Observation, Year 1)

Sun was teaching new vocabulary items. [...] She first read aloud a new word. Then she asked each pupil to utter the new word, and gave them a sticker if they did it correctly. [...] She did this with each new word, and gave out many stickers. (Sun; Observation, Year 1)

As shown in the extracts above, both Qian and Sun organised activities for classroom assessment purposes, although the materials they used were simple and the interaction such assessment required was brief.

When I asked them in following interviews how these ideas for classroom assessment occurred to them, they told me that most of these techniques were learnt from their more experienced colleagues:

The head of the EFL section told me that many teachers in the school used this method, because it can help the teachers immediately know how well the students have mastered the new words. [The researcher then asked why knowing this was important.] Anyway, as a teacher, you should know the progress of the students. [...] This is part of the teacher's responsibility, I guess. (Qian; Interview, Year 1)

When I just started to teach, I observed another teacher's class. She gave out the stickers whenever the pupils got things right. Later, I found no elementary school teacher was *not* doing this. [...] This is effective, because the pupils are happy to receive the stickers. [The researcher further asked how this would affect teaching.] The most important thing about this is that the pupils are kept motivated. (Sun; Interview, Year 1)

It can be seen that these practical techniques for classroom assessment, which are easy to learn and use with fixed procedures, were generally learnt from the experienced teachers. Interestingly, the novice teachers in this stage did not seem to understand the fundamental utilisability of classroom assessment, as they only considered 'know[ing] the progress of the students' important as 'part of the teacher's responsibility,' and considered such classroom assessment beneficial to pupils' learning, for instance, the sustaining of motivation. The following extracts on their conceptualisations of classroom assessment further

confirmed that the novice teachers in this stage only had a vague idea about the utilisability of classroom assessment:

It's about assessing how well the teacher teaches in a particular lesson. (Zhao; Interview, Year 1)

It's something like a [...] for example, a quiz to check students' progress, say, at the end of the class. (Qian; Interview, Year 1)

It's about testing, I guess, as it's a kind of assessing. (Sun; Interview, Year 1)

Teachers need to try to know how well their students learn. (Li; Interview, Year 1)

Further evidence of this lack of assessment awareness was also found in later interviews that were conducted in the second year of their teaching:

I've come to a clearer understanding of what classroom assessment is truly for. [...] It's about improving teaching, not a pure act of assessment as a teaching step. (Zhao; Interview, Year 2)

They've been emphasising a lot about assessment, assessment, and assessment, and it took me quite a while to see it's not a formality item in the teaching manual. (Li; Interview, Year 2)

This may indicate that in the first year, they simply regarded classroom assessment as a 'teaching step' or a 'formality item.'

To sum up, in this first stage, the novice teachers began to build the tool box but did not fully understand the pedagogical utilisability of the tools in it. Based on classroom observation, I also found that in this stage, novice teachers' assessment practices were planned practices. They planned not only specific methods/techniques to be adopted in classroom teaching, but also the time points in a lesson to adopt them.

Combining the Tool and the Purpose

The second stage for the development of their classroom assessment literacy seemed to last as long as two years, i.e. throughout the second and third years of their teaching. In this stage, their previous vague ideas about the utilisability of classroom assessment were gradually replaced by a growing awareness to connect classroom assessment to teaching objectives. Below are four extracts on their conceptualisations of classroom assessment in this stage:

It's about assessing how much and how well teaching objectives have been accomplished. (Zhao; Journal, Year 2)

It's about knowing what has been learnt and what has not, to be more accurate, what in your plan has been learnt or has not. (Qian; Interview, Year 3)

It's like responding to a checklist which contains all you need to do for a class. (Sun; Journal, Year 2)

It's about monitoring the learning progress against the roadmap the teacher has intended. (Li; Interview, Year 2)

The novice teachers used different wordings such as 'plan,' 'all you need to do,' and 'roadmap,' which, however, all referred to the 'teaching objectives' in general. The extracts also show that their assessment practices in this stage were still planned ones as in the first stage, although they were already connected with learning/teaching objectives.

In tandem with their conceptual shifts in the utilisability of classroom assessment, changes in their actual classroom assessment practices were also manifest. The extracts below show my observation of Zhao's class and her explanations of the classroom assessment she conducted in a following interview:

In this lesson, Zhao taught a reading passage. She set three learning goals for her students: (1) acquiring seven new words from the reading context; (2) understanding the general idea and details; (3) guessing three new words in the reading context (but the students didn't need to learn to use them). [...] She designed various activities/tasks not only for achieving these goals but also for checking if these goals have been met. For instance, after the seven new words were presented and taught, she asked the students to do a piece of compound dictation in which the seven new words were among the ten words/phrases to be filled in. (Zhao; Observation, Year 2)

I've been increasingly aware of the importance of checking how well the teaching objectives have been accomplished. You can't only spend time on the teaching itself without devoting some time, say just a small proportion of your class time, to assessing how rewarding your time spent on teaching is. [...] This is one of the few principles I've been most impressed about since I started to teach. [...] I've been repeatedly reminded of this in various teacher training programmes. (The researcher further asked what she would do if she found out some teaching objectives were not achieved.) In this case, I would reflect and reconsider if the teaching objectives are not properly set. [...] They may be too high for the students to reach. (Zhao; Interview, Year 2)

These extracts show that Zhao's classroom assessment practices concurred with her beliefs about classroom

assessment. She designed the compound dictation materials and implemented a compound dictation activity to assess if the students were able to produce the seven newly learnt words, which was one of the three teaching objectives for the lesson. The materials she used in this stage also seemed to be more complex than those in the previous stage. However, this stage is similar to the previous one in that the interaction the classroom assessment required was still as brief as a quiz, despite the more complexity involved in the content of the quiz. We can also see that her understanding of the utilisability of classroom assessment did not go beyond the teaching objective issue to teaching improvement—she only mentioned her reflection on the appropriateness of the teaching objectives rather than how she would draw on assessment results for teaching improvement.

With regard to how assessment results can specifically inform teaching adjustments, Zhao wrote in one of her journal entries:

If a problem occurs, the teacher must reflect on it and get things right in the next lesson. [...] If nothing can be done with this year's students, the teacher will have to be very careful the next time she teaches the same lesson. (Zhao; Journal, Year 2)

This means that, in Zhao's belief, assessment results could only guide future teaching rather than bringing about immediate teaching modifications. Qian also revealed a similar belief in her journal:

Learning to teach well takes a couple of 'errors and trials.' When something goes wrong, you will see that, and you will need to face it and fix it next time. [...] The teacher does not need to cry over the spilt milk. Learning a lesson from the lesson is what really makes a difference. (Qian; Journal, Year 2)

Interestingly, Qian mentioned the 'spilt milk.' She used this metaphor to argue that the problems were irreversible on the scene, and that remedial actions were only possible in future teaching.

To sum up, the second stage allowed novice teachers to gain a growing awareness to connect classroom assessment to teaching objectives. Their classroom assessment in this stage featured the use of more complex materials, although the interaction such classroom assessment required was still brief. In this stage, novice teachers began to combine the tool box and the purpose of classroom assessment, and this combination was often planned prior to classroom teaching. They also began to use assessment results to evaluate the realisation of teaching objectives, but they seemed to be confined to a stereotype that assessment results could only inform future teaching rather than immediate teaching adjustments.

Toward the Formative Assessment

The third stage seemed to begin in the middle of or toward the end of the third year, when the novice teachers were more conscious of the utilisability of classroom assessment regarding immediate teaching improvement. They used the results of classroom assessment to reflect on the problems in their teaching and instantly seek improvement. This is primarily reflected in their updated conceptualisations of classroom assessment:

It's about knowing where your students are and, more importantly, where *you* are as the teacher, and you'll know if you [the students and the teacher] are going in the same direction. [...] You'll have to stop and make sure you are. So, just stop now and then, check and look back. If anything goes wrong, get it right before it gets worse. (Zhao; Interview, Year 3)

The teacher needs to monitor the students' learning progress often and at important junctures so that the teacher will know if adjustment is needed in the following step. (Li; Journal, Year 3)

These extracts show that the novice teachers became more concerned about 'stop[ping] now and then,' and making 'adjustment' if 'needed,' which indicates enhanced understanding of classroom assessment as a means of online teaching diagnosis and improvement. Correspondingly, the 'stop[ping]' and the 'monitor[ing]' did not seem to be planned or plannable. The novice teachers needed to improvise in assessing how learner performance and their teaching goal were contingent. As seen in the extracts above, this improvisation may have been initiated by the teachers' noticing of possible learning difficulties or teaching problems (e.g. feeling that the teacher and students are not 'going in the same direction' or that 'anything goes wrong').

Given the enhanced knowledge of the utilisability of classroom assessment, Zhao and Li became more inclined to adopt, among various types of classroom assessment, formative assessment in teaching (Leung 2004). Li's extract below shows how formative classroom assessment was conducted, which also demonstrates that assessment practices in this stage feature improvisation:

Li's class featured much interaction with her students for checking the progress and for inquiring into difficulties students might have encountered. The following were a few turns of her conversation with one of the students:

'S (Student): Pandas were less and less because of natural disasters.

L (Li): How did you know that?

S: I got this from the book.

L: Could you tell me where?

S: In the second paragraph. The third line.

L: OK. Can you read that line?

S: The number of pandas decreases, and this does not have as much to do with natural disasters as human activities. Ah, yes, sorry, not because of natural disasters.

L: Right, you should pay closer attention.'

After this, Li gave a few more examples of 'not have as much to do with ... as ...,' and she also asked two students to make sentences using this structure. (Li; Observation, Year 3)

These extended conversation turns between Li and the student clearly show her concern with the causes of her student's mistake. Her concern was improvisational and even spontaneous. When she knew it was due to the student's misunderstanding of a structure (i.e. 'not have as much to do with ... as ...'), she paused, and turned to focusing on this structure. The way she inquired into the student's problem was apparently a kind of classroom assessment; the way she used the results of classroom assessment in her following teaching was obviously formative, as she explained to me in a later interview that '[she] guessed that it might be a difficult point, and [that she] needed to help the students focus on this a little bit.' (Li; Interview, Year 3) Besides, we can also see from the above extract that the interaction included in this assessment scenario was no longer brief and procedural as in previous stages. It was authentic and dynamic. As to how she developed the ability of formative classroom assessment, Li told me that 'it [was] basically based on [her] personal learning and reflection, such as reading books of theories and professional journals and always trying to see if what [was] happening in the classroom [was] related to the theories.' (Li; Interview, Year 3)

To sum up, in the third stage in the middle of or toward the end of the third year, the novice teachers might develop an enhanced knowledge of the utilisability of classroom assessment regarding teaching improvement, and thus become more inclined to conduct improvised formative classroom assessment, which shows immediate assessment-driven teaching improvements in response to dynamic classroom interactions.

Discussion

So far I have tried to describe the three developmental stages for novice teachers' classroom assessment literacy, which have emerged from data analysis. In the first stage, the novice teachers tended to acquire practical, easy-to-learn techniques with fixed procedures for planned classroom assessment without much awareness of the

utilisability of classroom assessment regarding teaching improvement. In the second stage, the novice teachers seemed to gain a growing awareness to connect their planned classroom assessment to teaching objectives, and the classroom assessment they conducted featured the use of more complex materials, although the interaction in such classroom assessment was still brief. More importantly, novice teachers' understanding of the usefulness of assessment results in teaching adjustments was only confined to an improved plan of a future teaching situation. In the third stage, the novice teachers developed an enhanced knowledge of the utilisability of classroom assessment regarding teaching improvement, and thus became more inclined to conduct improvised formative classroom assessment when noticing possible learning difficulties or teaching problems, which shows immediate assessment-driven teaching improvements in response to dynamic classroom interactions.

With regard to novice teachers' development in classroom assessment techniques, the complexity of materials used for classroom assessment seems to increase over time, and interaction included in classroom assessment also becomes increasingly enriched and dynamic. This may indicate that novice teachers tend to begin with simple and procedural techniques, and gradually master the more complex and flexible ones. This also corroborates previous research findings regarding teachers' use of assessment methods/techniques (e.g. Black and Wiliam 1998).

As to the development of novice teachers' understanding of the utilisability of classroom assessment, an awareness of the connectivity between classroom assessment and other important aspects of teaching seems to play a pivotal role. As novice teachers' classroom assessment literacy develops, such awareness may be first reflected in the connection between classroom assessment and teaching objectives, which calls for teaching improvements in a future teaching situation. This awareness may then develop and be reflected later in the connection between classroom assessment and immediate teaching improvements, allowing more formative assessment to be undertaken. In other words, the way novice teachers understand and use classroom assessment results in relation to teaching improvements may experience an evolution from a 'next time' mindset to a 'right now' ideology. The significance of this finding has moved beyond building up taxonomies of assessment methods/techniques (Hailaya et al. 2014; Stiggins 1988) or testing the effectiveness of using classroom assessment results in making teaching adjustments (Coombe et al. 2012; Phye 1997; Stiggins 1991). This has shed some light upon a possible pattern that defines a very kernel aspect of novice teachers' classroom assessment literacy development, and should open up new possibilities

of specific intervention for teacher educators (Popham 2011).

It is also interesting to notice that different classroom assessment knowledge which is developed in different stages seems to be acquired through different mediations. The simple, easy-to-use techniques were mostly learnt from peers in daily communication and school-based lesson observation activities; the acquisition of knowledge regarding connecting classroom assessment to teaching objectives seemed to be mediated by in-service teacher training; as to connecting classroom assessment to immediate teaching improvement. However, personal learning and reflection seemed to play a major meditational role, which, of course, requires a high level of autonomy on the teacher's part (Dierking and Fox 2013).

The findings of the current study also suggest that improvisation in classroom assessment may be a crucial indication of classroom assessment literacy development, or a more difficult point of attainment, as they were only found in two of the four participants and only in the third year of their beginning career. Teachers' ability to improvise in classroom assessment can be seen as an outgrowth of and further enhancement beyond their previous mastery of assessment methods/techniques as well as the emerging awareness to plan assessment in connection with the teaching goal. Improvised assessment seems to be a kind of flexible implementation of teachers' knowledge of assessment methods/techniques and knowledge of its utilisability in a more responsive way to the constantly changing classroom dynamics.

Conclusions

This three-year longitudinal study examines the development of four Chinese novice EFL teachers' classroom assessment literacy with a specific focus on how they used assessment results to make teaching improvements. Data analysis following the paradigmatic analytic procedures reveals a three-stage developmental route: planned assessment for *no* improvement, planned assessment for *future* improvement, and improvised assessment for *immediate* improvement.

Based on the findings of the study, a few implications for teacher education and development may be discerned. Because pre-service teacher education did not seem to play an evident role in the development of novice teachers' classroom assessment literacy, more efforts are needed for pre-service teacher education to better equip future teachers with the basics of classroom assessment. Pre-service teacher education programmes may need to incorporate a more explicit introduction to frequently used classroom assessment methods/techniques and try to enhance pre-

service teachers' awareness of the contingency between assessment, teaching, and learning. Importantly, the enhancement of the contingency awareness should go together with the learning of assessment methods/techniques so that the developmental route may be shortened. Second, it may be a more effective way to promote novice teachers' classroom assessment literacy by combining the assessment methods/techniques and the awareness of utilisability in various school-based and non-school-based teacher development activities. In this study, the natural order of classroom assessment literacy development is from assessment for no improvement to future improvement and then to immediate improvement, but it seems quite possible that this developmental route can be shortened if novice teachers are shown in earlier stages of development how immediate improvement can be achieved through assessment practices.

Acknowledgements This work was supported by a Grant from the National Social Science Fund of P. R. China [number 13CYY028]. I am tremendously grateful to the editors and anonymous reviewers for their valuable comments and suggestions. I also feel grateful to Professor Charlene Polio for her suggestions, and to the discussants who shared their invaluable opinions on an early version of the paper presented at the Global Curriculum & Instruction Network Conference 2014 which was held in June 2014 in Hong Kong.

References

- Black, P. (2009). Formative assessment issues across the curriculum: The theory and the practice. *TESOL Quarterly*, *43*, 519–524.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, *80*, 139–148.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, *21*, 5–31.
- Cheng, L., Roger, T., & Hu, H. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods and procedures. *Language Testing*, *21*, 360–389.
- Clark-Gareca, B. (2016). Classroom assessment and English language learners: Teachers' accommodations implementation on routine math and science tests. *Teaching and Teacher Education*, *54*, 139–148.
- Coombe, C., Troudi, S., & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges, and recommendations. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyoff (Eds.), *The Cambridge guide to second language assessment* (pp. 20–29). Cambridge: Cambridge University Press.
- Dierking, R. C., & Fox, R. F. (2013). 'Changing the way I teach': Building teacher knowledge, confidence, and autonomy. *Journal of Teacher Education*, *64*, 129–144.
- Erickson, F. (2010). Taking advantage of structure to improvise in instruction: Examples from elementary school classroom. In S. Sawyer (Ed.), *Structure and improvisation in creative teaching* (pp. 113–132). Cambridge: Cambridge University Press.
- Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through pre-service teacher education. *Teaching and Teacher Education*, *21*, 607–621.
- 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*, 297–317.
- Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*, 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*, 345–362.
- Johns, K. E., & Gerard, H. B. (1967). *Foundations of social psychology*. New York: Wiley.
- König, J., Blömeke, S., Klein, P., Suhl, U., Busse, A., & Kaiser, G. (2014). Is teachers' general pedagogical knowledge a premise for noticing and interpreting classroom situations?: A video-based assessment approach. *Teaching and Teacher Education*, *38*, 76–88.
- Lam, R. (2014). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, *1*–29.
- Leung, C. (2004). Developing formative teacher assessment: Knowledge, practice, and change. *Language Assessment Quarterly*, *1*, 19–41.
- Phye, G. D. (1997). Classroom assessment: A multidimensional perspective. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, achievement, and adjustment* (pp. 33–51). California: Academic Press.
- Polkinghorne, D. E. (1995). Narrative configuration in qualitative analysis. In J. A. Hatch & R. Wisniewski (Eds.), *Life history and narrative* (pp. 5–23). London: Falmer.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, *48*, 4–11.
- Popham, W. J. (2011). Assessment literacy overlooked: A teacher educator's confession. *The Teacher Educator*, *46*, 265–273.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, *18*, 429–462.
- Rea-Dickins, P. (2007). Learning or measuring? Exploring teacher decision-making in planning for classroom-based assessment. In S. Fotos & N. Nassaji (Eds.), *Form-focused instruction and teacher education* (pp. 195–211). Oxford: Oxford University Press.
- Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, *63*, 66–70.
- Stiggins, R. J. (1988). Revitalizing classroom assessment: The highest instructional priority. *Phi Delta Kappan*, *69*, 363–368.
- Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, *72*, 534–539.
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Albany, NY: State University of New York Press.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? In C. A. Dwyer (Ed.), *In the future of assessment: Shaping teaching and learning* (pp. 53–82). New York: Erlbaum.
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, *58*, 149–162.
- Xu, Y., & Liu, Y. (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, *43*, 493–513.