
Developing Classroom-Based Formative Assessment Literacy: An EFL Teacher's Journey

Jiayi LI & Peter Yongqi GU

Victoria University of Wellington, New Zealand

Abstract

Formative assessment has become an increasingly popular topic in educational reform since the publication of Black and Wiliam's (1998) highly influential review article. In China, formative assessment has been promoted by the Ministry of Education in its curriculum standards for over two decades. A lack of implementation in classrooms has been attributed to the lack of assessment literacy among classroom teachers. A 12-week professional development program was designed and implemented for a group of five secondary school EFL teachers in China aiming at developing their classroom-based formative assessment literacy. The program took the form of collaborative action research in which a researcher joined the teachers to form a community of practice for continuing professional development. This paper focuses on one teacher's growth in assessment literacy for classroom-based formative assessment. Classroom video recordings and interviews were transcribed and coded using NVivo 12. Baseline data indicated that the teacher lacked the knowledge and beliefs needed to conduct formative assessment. She did not have clear targets for teaching, learning and assessment; and she did not make any targets clear to her students. In addition, the teacher's assessment practices did not aim to improve student subject core competencies as stipulated in the senior secondary English language curriculum. By the end of the program, the teacher's knowledge, beliefs and practices in formative assessment were significantly enhanced. In particular, her goal setting for classroom teaching, learning and assessment was intentionally aligned with the development of subject core competencies. More cycles of formative assessment practices were found to help students close the learning gaps and achieve their learning goals. These findings lend encouraging support to the feasibility of teachers developing formative assessment literacy within a continuing professional development framework.

Keywords: *assessment literacy, formative assessment, teacher learning, continuing professional development, collaborative action research*

1. Introduction

Classroom assessment occupies an important position in teaching activities. A third to a half of a teacher's professional time is spent on involvement in assessment-related activities

(Stiggins, 1999). Classroom assessment is closely related to teaching and learning, directly affecting the quality of classroom teaching (Popham, 2017). Teacher assessment literacy determines the quality of assessment since classroom assessment decisions are mostly derived from teachers' judgments (Volante & Fazio, 2007). From the perspective of teacher professional development, teacher assessment knowledge is one of the elements of Pedagogical Content Knowledge (Shulman, 1986). Improving teacher assessment literacy can empower teachers (Popham, 2009), and should be an important goal of teacher education.

This study thus aimed to design and implement a continuing professional development program for a group of five EFL teachers at two secondary schools in Hebei, China. The focus of the program was the provision of "just-in-time, job-embedded assistance" (Guskey & Yoon, 2009) for the on-going and on-site development of teacher assessment literacy.

1.1 Assessment Literacy

Assessment-literate educators come to any assessment "knowing what they are assessing, why they are doing so, how best to assess the achievement of interest, how to generate sound samples of performance, what can go wrong, and how to prevent those problems before they occur" (Stiggins, 1995, p. 240). Classic conceptualizations of assessment literacy such as the Standards for Teacher Competence in Educational Assessment of Students by American Federation of Teachers et al. (1990) and Stiggins (1995) have been complemented by later reformulations such as Heritage (2007), Brookhart (2011), Willis et al. (2013), and DeLuca et al. (2016), taking into consideration of recent development in formative assessment. Heritage (2007) highlighted the significance of formative assessment and proposed the required knowledge and skills for teachers to use formative assessment successfully in the classroom. Brookhart (2011) proposed an updated list of assessment-related knowledge and skills for teachers to perform in a professional and competent manner.

Most recent models of assessment literacy have focused both on theoretical reformulations of the construct (Andrade et al., 2021; Xu & Brown, 2016) and on the teacher's ability for practical enactment in the classroom (Herppich et al., 2018). Xu and Brown's (2016) teacher assessment literacy in practice included (1) the knowledge base, (2) teacher conceptions of assessment, (3) institutional and socio-cultural contexts, (4) teacher assessment literacy in practice, (5) teacher learning, and (6) teacher identity (re)construction as assessors. Andrade et al. (2021) conceptualized classroom assessment as co-regulated learning. In this view, formative assessment in the classroom was conceptualized as socially and technologically mediated planning, monitoring, control, and reflection. Herppich et al. (2018) saw "teacher assessment competence" as a learnable ability that allows a teacher to translate knowledge into a judgment through decision-making processes.

Empirical research on language assessment literacy has examined the knowledge, skills, and principles needed for various stakeholders (Taylor, 2013). Assessment literacy for language teachers has been explored from either a knowledge-based perspective for teacher education or the assessment training needs of language teachers (Fulcher, 2012). Most of these conceptualizations of assessment literacy cover a whole range of capacities relevant to language testing and assessment. How these knowledge and abilities are enacted inside the

classroom, and how this “assessment literacy in action” can be improved have not received enough attention.

A number of researchers have attempted to help teachers develop their assessment literacy. Torrance and Pryor (1998) documented a longitudinal study among primary schools in the UK and examined classroom-based formative assessment through a “micro-sociological” perspective. Black et al. (2003) worked with a group of forty-eight teachers of English, mathematics and science in the UK and developed ideas about formative assessment together and helped the teachers to put the ideas into practice. McMunn et al. (2004) described a district level project in the US which spanned three phases over eight years. A recent study by Andersson and Palm (2017) reported the results of an assessment training program for a semester and worked with a group of 22 mathematics teachers in Sweden. The study found significant improvement in both assessment literacy of the teachers and the academic achievement of the students. Similar research for the enhancement of classroom assessment literacy for formative purposes among teachers of English as a foreign language is hard to find, although a large number of exploratory studies about assessment practices and assessment literacy is beginning to emerge (e.g., Gu, 2014; Hill & McNamara, 2012).

1.2 Formative Assessment

Formative assessment is the process from “where the learner is right now” to “where the learner is going” (Black & Wiliam, 2009, p. 8). Evidence suggests that formative assessment is an important part of classroom work, and the development of formative assessment can raise standards of learning achievement (Black & Wiliam, 1998). In recent years, governments around the world have become increasingly active in pursuing this goal through educational reform. The Chinese government has introduced a series of policy documents as part of the English curriculum reform (Gu, 2012). Most of these documents, especially the curriculum standards, explicitly require the implementation of formative assessment in the classrooms (MOE, 2020).

The defining features of formative assessment practices are visualized by Gu (2021) in spiraling cycles which include five elements: clarification of goals, elicitation of evidence, interpreting the evidence, providing feedback, and student/teacher take-up and action. These elements are both sequential and interactive. Each complete cycle serves the formative purpose of “moving student understanding or learning closer to the target” (Gu, 2021, p. 14). This framework was used to construct the framework for teacher learning in classroom-based formative assessment, as well as to guide the coding scheme and data analysis of the classroom assessment practices of the participating teachers in the study reported in this article.

In formative assessment, where teaching, learning and assessment become an integrated whole, targets of teaching and learning and criteria for success are equal to the targets of assessment and success criteria. For formative assessment, especially classroom-based formative assessment, the entire process of how to assess, how to interpret the results, what feedback to provide, and whether and how to take follow-up actions depends on teachers' understanding of the assessment objectives and success criteria. Essentially, teaching, learning and assessment goals are the starting point and ultimate destination of formative assessment,

and there is no way to talk about monitoring progress and assessing success without knowing where the learning goals are and what the criteria for success are. Therefore, one crucial step teachers and their students need to do is to identify the teaching, learning and assessment goals and then explore ways to get closer to the targets of learning after knowing the gap between learners' current level and target level of learning.

In classroom assessment, teaching, learning and assessment objectives come from curriculum standards and the associated theoretical conceptions of language competence. Clear targets play an important role in guiding English language teaching, learning and assessment. In addition, criteria for successful achievement of the objectives should be clearly identified and the teachers' understanding and interpretation of the criteria form an important part of their pedagogical content knowledge. In classroom-based formative assessment, teachers should first familiarize themselves with the curriculum standards, which set out what is to be taught and learnt. Second, teachers should look beyond the objectives of each lesson to grasp the theoretical conceptualization of language competence, to map the objectives of each lesson to the theoretical construct, and to make judgments based on various situations in classroom teaching, learning and assessment.

1.3 Teacher Continuing Professional Development

"Teachers need substantial knowledge to implement formative assessment effectively in classrooms. It is doubtful that the average teacher has that knowledge, so most teachers will need substantial time and support to develop it" (Bennett, 2011, p. 20). Research has shown that traditional teacher training is not effective and that the short workshop approach does not work for teacher professional development (Guskey & Yoon, 2009).

Continuing professional development (CPD) includes "all the activities in which teachers engage during the course of a career which are designed to enhance their work" (Day & Sachs, 2005, p. 3). Effective teacher professional development is "structured professional learning that results in changes to teacher knowledge and practices, and improvements in student learning outcomes" (Darling-Hammond et al., 2017, p. 2).

Over the decades, there have been different ways in conducting teacher CPD. Kennedy (2005) summarized the models into nine types (1) training, (2) award-bearing, (3) deficit, (4) cascade, (5) standards-based, (6) coaching/mentoring, (7) community of practice, (8) action research, and (9) transformative. These have often been interpreted as a transmission type vs a transformative type. In 2014, she (Kennedy, 2014) removed "transformative" as a type of CPD, because it is "more a combination of experiences and contextual factors rather than a model itself" (p. 693). She regrouped the remaining eight types into three categories of purposes, transmissive, malleable, and transformative, highlighting the continuum nature of her list. In addition to this dimension of purposes, which Sachs (2016) referred to as "functional development" vs "attitudinal development," Sachs identified another dimension "organizational/managerial" vs "occupational/democratic" professionalism. When these two dimensions are combined, she has four "types of CPD": controlled professionalism, compliant professionalism, collaborative professionalism, and activist professionalism. This study falls mostly under the fourth category, although certain features of collaborative professionalism

will also be present.

Timperley et al. (2007) synthesized a wide range of research and proposed a model for teacher CPD which they termed “teacher inquiry and knowledge-building cycle.” This model starts with the teacher’s analysis of student learning needs and teacher learning needs before any teacher development tasks are designed, implemented, and evaluated. The model sees CPD as co- and self-regulatory learning cycles, with emphasis on teacher agency, and combines the features of communities of practice/learning, collaborative action research.

Effectiveness of CPD has been a main line of research on teacher professional development. Darling-Hammond et al. (2017) reviewed 35 studies conducted over a span of three decades and summarized a similar list of criteria, that effective professional development: (1) is content focused, (2) incorporates active learning, (3) supports collaboration, (4) uses models of effective practice, (5) provides coaching and expert support, (6) offers feedback and reflection, and (7) is of sustained duration. Most research findings in these intervention studies surveyed by Darling-Hammond et al. support both improvement in teaching and in student learning.

The approach adopted in this study was collaborative work practices that aim for the empowerment of teacher agency. Since teachers participating in this study were volunteers, and that participation in this research was not related to any form of performance and accountability, this model was more “democratic” than “managerial” (Sachs, 2001, 2016). Timperley et al.’s (2007) “teacher inquiry and knowledge-building cycle” was followed as the CPD model, and Darling-Hammond et al.’s (2017) core features of professional development was used as design criteria for the professional development plan for this study.

2. This Study

This study aims to explore the extent to which a CPD program helps in the development of teacher classroom assessment literacy for formative purposes. The usefulness of the CPD program in improving the students’ English language achievement is also studied. A 12-week professional development program was designed and implemented. Five teachers (4 females and 1 male) teaching classes equivalent to Grade 10 in the 2020-2021 school year were invited to participate in this study. The five teachers and one researcher (the first author) formed a community of practice for professional development in classroom assessment literacy. This paper zooms in on one of these participating teacher’s formative assessment knowledge, beliefs, and practices before, during, and after the CPD program.

2.1 The CPD Program

A 12-week CPD program of two “teacher inquiry and knowledge-building cycles” aimed at enhancing teacher classroom-based formative assessment literacy was implemented. The CPD program consisted of four stages. Stage 1 was a three-week preparation stage when key readings and ideas were discussed and shared followed by goal setting for the CPD. Stages 2 and 3 involved two rounds of actions that lasted four weeks each. At the beginning of Stage 2, the community of practice designed a classroom-based formative assessment task as a group to be used in their own classes. Then each of the five teachers implemented the formative

assessment task that they designed together and video-recorded one lesson that focuses on this task. Self- and group-reflections were initiated at the end of this stage using the video-recorded lessons. In Stage 3, the teachers took these reflections into account and designed another formative assessment task to be implemented in their own classes. Again, one lesson per teacher was video recorded to go through the same process as in the previous stage. Stage 4 focused on a one-week program evaluation. The five teachers went through a guided reflection of the whole CPD program.

2.2 Cassie's Teaching Context

The teacher, Cassie (pseudonym), was a female teacher in her late thirties teaching at a good school in a city in north China. She obtained double BA degrees in English and Japanese language and literature and became an English language teacher upon graduation. At first, she did not know how to teach because she had not received teacher training. She became a good teacher by reflection and improvement, and by learning from other teachers and external experts. By the time she participated in this study, she had had over 13 years of teaching experience at two secondary schools in Hebei Province. However, Cassie had never taken any assessment courses or read any assessment books. She was highly motivated to improve her classroom assessment literacy, and took an active part in the study.

The textbooks were Chinese General Senior Secondary English Compulsory Textbooks (Volumes 1-3, 2019 edition) published by People's Education Press. The textbooks were designed and written based on the 2017 edition of the General Senior High School Curriculum Standards. Each volume contained five units, so the three volumes contained a total of 15 units. Hebei Province, where Cassie is located, started to use this latest version of textbooks in September 2020. By the time of the study, the textbooks had newly been adopted. As such, lesson preparation and teaching had to start from scratch. An older edition of the textbook that had been used before did not contain explicit learning objectives. The 2019 edition, the latest version, has an introduction that appears at the beginning of each unit which resembles learning objectives.

2.3 Data Collection and Data Analysis

To capture teachers' formative assessment practices embedded in classroom teaching, learning, and assessment, classroom video recording data were collected and analyzed, and patterns were generated from the analysis. A total of four lessons of 40 minutes each were video recorded, including two lessons in the baseline phase and two lessons in the professional development program. A Reading & Thinking lesson in the textbook, Volume 3 of the 2019 edition, was chosen as the module for two rounds of action research in the CPD program. One lesson in each of the two rounds of actions was video recorded for teacher self- and group-reflections after class.

Cassie was interviewed about her knowledge, beliefs and practices of formative assessment at the beginning and at the end of the teacher professional development program. The interviews were semi-structured and lasted about 30 minutes each. The pre-interview was conducted after the video recordings in the baseline phase. The post-interview was conducted at the end of the CPD program. In the post-interview, Cassie was also asked to describe the

ways and extent to which the CPD helped in changing her assessment knowledge and beliefs. Follow-up questions were asked in the interviews with further description of assessment knowledge, beliefs, as well as assessment practices. Cassie was interviewed in Chinese. The transcripts were translated into English and were back translated. Two students in Cassie's class were interviewed as well before and after the action research. Cassie's reflection journals, teaching plans, and other artefacts throughout the semester were collected to obtain complementary data for analysis.

The classroom video recording data and interviews were transcribed and coded using NVivo 12. The coding scheme was derived from both top-down and bottom-up processes. The five components of classroom assessment practices illustrated in Gu (2021) was the starting point that guided the top-down perspective of what should be involved in classroom-based formative assessment. The subject core competencies, which were the goals for senior high school teaching and learning required by the curriculum standards, served as the targets of teaching, learning, and assessment in planning and implementing formative assessment tasks. Exactly what the teacher did in class in each of the components was analyzed as bottom-up processes. After the coding was done, a tally of each code was used as an indicator of the teacher's classroom assessment practices in these lessons.

3. Findings

The teacher CPD program was found to have had a significant impact on the growth of Cassie's classroom assessment literacy for formative purposes. Cassie gained assessment knowledge after the stage of reading, sharing, and discussion in the participatory action research. Not only did she implement the classroom-based formative assessment task the group designed together, but she also changed her own assessment practices.

3.1 Targets of Teaching, Learning, and Assessment

3.1.1 Before CPD

What is assessed, how it is assessed, how the assessment result is interpreted, what feedback is provided to the students, whether and what follow-up action needs to be taken, all depend on teachers' understanding of targets and success criteria for achieving the targets. Cassie's long-term goals for the first year of senior secondary school were on "the development of learning methods and learning habits" (Cassie-TIPre5¹). Her mid-term goals were "to improve their (students') oral expression" and to do "sustained silent reading (SSR) before the end of the semester" (Cassie-TIPre5). Her short-term goals included the unit objectives covered in the textbook, such as "the past participle as adverbial modifier" (e.g., "*Satisfied* with what he did, the teacher praised him in class") for the week (Cassie-TIPre5).

It was worth noting that these were general and vague statements as teaching objectives which were not clear and specific. In addition, she set unreasonable goals. For example, "sustained silent reading" was one of the practices for improving reading ability, which cannot

¹ In this article, the following short forms are used to indicate the source of a piece of data. TIPre: teacher pre-interview; TIPost: teacher post-interview; SIPost: student post-interview; RJ: reflection journal; Cassie-TIPre5: the fifth question in the pre-interview with Cassie; S2-SIPost7: the seventh question in the post-interview with the second student; Cassie-RJ3: Cassie's reflection journal three.

be counted as a goal. It was unclear what goals she hoped to achieve through this practice. Moreover, these long-, mid-, and short-term goals lacked coherence. The short-term goals for a unit were not contributing to the achievement of mid-term goals (improving oral ability and conducting SSR inside the classroom). Furthermore, improving oral ability and conducting SSR (mid-term goals) might not lead to the fulfilment of the long-term goals (development of learning methods and learning habits). Cassie's three-stage goals lacked systematicity and feasibility, and the achievement of short- or mid-term goals did not necessarily lead to the fulfilment of long-term goals.

In the two base-line lessons which focused on presenting new materials, Cassie never made her objectives of teaching and learning clear to her class. On the other hand, in the interviews in which Cassie talked about her teaching in general, she revealed an unclear understanding of what learning targets were. Most of the time, she mistook the problems she identified in previous homework as learning targets for future lessons. For example, she found that in her "students' homework there was a common problem with long sentences and complex sentences" (Cassie-TIPre6). She then made it the goal for the next lesson.

Another way Cassie set her learning targets was whether something was closely related to the university entrance examination. No matter whether it was a test or an exercise, where the tasks normally simulated the university entrance examination, problems were immediately identified as targets for the next lesson. In fact, this was typical of this whole group of senior secondary school teachers.

These activity-based targets had something to do with the ultimate goal of English language teaching since they were all about improving students' English language competence. However, these goals did not have any specific connection with the curriculum targets and textbook targets. Rather than setting goals for each unit, each stage, each month or each year, Cassie made it her goal to solve the problems she was currently encountering. Moreover, she did not even think about what this lesson had to do with the overall teaching goals or with what they needed to achieve in the long run. This explains why Cassie's students denied in the pre-interview that their teacher made learning targets clear to them.

3.1.2 After CPD

In the first round of action, clarification of goals was emphasized in group sharing and discussion sessions of the community of practice. Cassie showed four learning targets at the beginning of the lesson. These were closely related to the goals of developing student subject core competencies in the curriculum standards covering all the four aspects of language ability, cultural awareness, thinking capacity, and learning ability. Then she began her class without knowing how well students understood the goals. After the first round of action, Cassie reflected on the video recorded lesson, and realized that "it was not clear enough for students to understand the targets of learning by merely seeing learning targets on slides" (Cassie-RJ3). In addition, in her self-reflection journal, Cassie realized that she was "too ambitious trying to achieve all four teaching objectives by using six classroom activities during the lesson" (Cassie-RJ3). She found herself and her students "too busy in that lesson" because they "had many tasks to do in a 40-minute lesson" (Cassie-RJ3). Cassie did not even have

much time to close the learning gap she identified in class, so she left that as homework for her students.

In the second round of action, Cassie set three teaching targets for the lesson, and designed one activity for warming up and three classroom reading, thinking, and discussion tasks based on the three targets. She provided explanations together with exemplars to further clarify the sub-goals of the targets of learning after providing learning targets on the slide. She also organized group discussions on the learning targets for this lesson so that different ideas could be exchanged and agreement could be reached. At the end of the lesson, she found that she had “achieved every target of teaching and the students were given enough opportunities to close the learning gap” (Cassie-RJ4).

In addition to Cassie's changes in targets of teaching, learning, and assessment, it is worth noting that the learning objectives were woven into each step of the formative assessment cycle, allowing the formative assessment cycle to be guided by the objectives.

In the post-interview, Cassie was asked about her short-term, mid-term, and long-term English language teaching goals for year one senior secondary students. This time, her goals became much more coherent and systematic. As for her short-term goals, she used the teaching objectives for developing student core competencies of a lesson as an example to illustrate how she set goals for each lesson. She became very proficient in setting goals based on the curriculum standards. She was also able to cover the four dimensions of subject core competencies as teaching objectives, and then design classroom tasks according to the objectives. For her mid-term goals, Cassie broke down the end-of-semester goals into several stages, focusing on achieving mid-term goals, preparing for long-term goals, and solving problems whenever students encountered them. For Cassie's long-term goals, she set the end-of-year goals according to the curriculum standards. These were the core competencies that senior secondary school students should develop by the end of each academic year. Cassie said that the goals for English language teaching, learning, and assessment in the curriculum standards became clearer to her:

What I have benefited the most from participating in the action research is that I'm starting to clarify learning objectives for my students before each lesson. I used to think that the objectives only applied to teaching new lessons, but now I find that they apply to all types of lessons (Cassie-TIPost9).

Cassie also talked about the importance of criteria for success:

It is particularly necessary to give the students success criteria before they complete the task, or more intuitively, to give an example. We provide such criteria and an example, but there is more than one answer. We accept all possible answers as long as they meet our criteria. This is another lesson that I've learned from the action research (Cassie-TIPost3).

3.2 Elicitation of Evidence

Overall frequency. The overall frequency of elicitation dropped slightly from 485 times in the baseline phase to 474 times after the professional development phase. More specifically, the elicitation foci (especially the focus on subject core competences specified in the curriculum) increased from 280 times to 302 times after teacher learning, while the elicitation

techniques (questioning and observation) dropped from 205 times in the baseline phase to 172 times in the action research.

Elicitation focus. Table 1 shows that Cassie's elicitation focus shifted from classroom function in the baseline phase to subject core competencies after CPD. The frequency of elicitation for classroom function dropped from 184 times in the baseline phase to 124 times after professional development, while her practice of eliciting subject core competencies increased to 178 times from 96 times, becoming the focus of elicitation practices after CPD.

Table 1. *Changes in Elicitation Focus*

Elicitation Focus	Before CPD	After CPD
Classroom function	184	124
Warming up	59	19
Classroom management	56	23
Requests or directions	50	40
Learning check	19	42
Subject core competencies	96	178
Language ability	61	66
Cultural awareness	0	23
Thinking capacity	4	25
Learning ability	31	64

Within the focus on classroom function, the frequency of eliciting for warming up, classroom management, and requests or directions dropped, while the frequency of learning checks increased after professional development. After CPD, Cassie asked fewer warming up questions, with the frequency dropping to 19 times from 59 times in the baseline phase. Questions about student background knowledge (e.g., *Where is San Francisco?*) dropped the most as these only served as warming up to the new lesson, which did not contribute to the achievement of learning targets. Often the background knowledge being discussed was related to the art, science, and geography surrounding the topic they were learning about. Also, Cassie asked fewer management questions (e.g., *Who would like to have a try?*) after professional development. She mentioned in the reflection journal that her “students became confident and motivated to share their opinions and ideas after group discussion”, so she “no longer needed to spend much time asking management questions as they were willing to answer the questions voluntarily” (Cassie-RJ3). Besides, Cassie mentioned in the post-interview that “organizing more classroom tasks and group discussions” after CPD may be one of the reasons that “led to the decrease in requests or directions” (Cassie-TIPost13). In addition, Cassie asked more questions for learning check, e.g., questioning about task completion, questioning for confirmation, and questioning for understanding, etc.

There was a dramatic increase in eliciting subject core competencies, rising from 96 times in the baseline phase to 178 times after professional development. Cassie did not have any teaching targets in the baseline phase. After professional development, she intentionally set objectives to develop students' subject core competencies. The elicitation of learning ability

increased the most, as several tasks were designed for developing student learning strategies in the two lessons during the action research, with the frequency increased from 31 times before CPD to 64 times after CPD. The elicitation of language ability increased from 61 times in the baseline phase to 66 times after CPD. The frequency of eliciting students' cultural awareness increased to 23 times after teacher learning from 0 in the baseline phase. Cassie even designed a few tasks mainly about cultural diversity in the first round of action. She also paid more attention to the development of student thinking capacity, with the frequency of elicitation increasing from 4 times in the baseline phase to 25 times after professional development.

Elicitation technique. Table 2 shows that Cassie used a combination of observation and questioning as elicitation techniques before and after professional development. After CPD, she used more classroom tasks and less questioning to elicit students' learning or understanding.

Table 2. *Changes in Elicitation Techniques*

Elicitation Technique	Before CPD	After CPD
Classroom tasks	40	49
Listening	16	0
Reading	11	38
Speaking	13	8
Viewing	0	3
Questioning	165	123
Constructed-response questions	43	74
Open-ended questioning	7	14
Selected-response questions	43	5
Rising intonation	49	23
Yes or no questions	23	7

Using classroom tasks as an elicitation technique increased from 40 times before CPD to 49 times after CPD. Pair discussion and group discussion increased after teacher learning. Students became more confident to share their opinions after exchanging ideas and broadening thinking within groups or pairs. Cassie also initiated more reading and thinking tasks in the action research.

The frequency of questioning dropped after professional development, although whole-class questioning and individual student questioning remained dominant elicitation techniques. What matters is not the frequency of elicitation but the amount of learning evidence elicited. When Cassie watched her classroom video recordings for the baseline phase, she was a bit shocked. She “did not mean to ask so many questions” (Cassie-RJ2). She “simply felt awkward” for her class to remain silent and therefore “kept asking questions” (Cassie-RJ2). After reflection, Cassie gave her students more wait time to process her questions and even initiated classroom discussions in the action research phase.

The type of questioning with the largest increase was constructed-response questions, rising to 74 times after CPD from 43 times before CPD. Also, Cassie asked more open-ended questions (from 7 times to 14 times) in order to encourage critical thinking, logical thinking,

and creative thinking. She avoided asking yes or no questions (from 23 times to 7 times) and selected-response questions (from 43 times to 5 times) after professional development. In addition, using a rising intonation to wait for student answers normally does not elicit anything more than a word. This type of questions dropped from 49 times to 23 times.

3.3 Interpretation of Evidence

Overall frequency. Cassie's interpretation practices increased after professional development, from 277 times before CPD to 368 times after CPD. More specifically, interpretation focus increased from 136 times to 192 times, and her interpretation against success criteria increased from 141 to 176 times.

Table 3. *Changes in Interpreting Evidence*

Interpreting Evidence	Before CPD	After CPD
Interpretation focus	136	192
Classroom function	56	39
Warming up	48	11
Classroom management	3	1
Requests or directions	0	3
Learning check	5	24
Subject core competencies	80	153
Language ability	54	61
Cultural awareness	0	15
Thinking capacity	6	22
Learning ability	20	55
Success criteria	141	176
Information focused	126	131
Language focused	14	24
Student interpretation	0	11
Task focused	1	10

Interpretation focus. Cassie's interpretation for classroom function decreased and the interpretation for subject core competencies increased after professional development (Table 3). Interpretation for classroom functions dropped from 56 times in the baseline phase to 39 times after professional development in line with the decrease of elicitation for classroom functions after teacher learning. Specifically, warming up, classroom management, and requests or directions dropped, while learning checks increased after teacher learning. On the contrary, interpretation for subject core competencies increased dramatically from 80 times to 153 times. Cassie became much more deliberate in interpreting learning evidence against the targets of teaching, learning, and assessment.

Success criteria. Table 3 shows that information-focused criteria remained the dominant criteria for interpretation in the baseline phase and professional development phase, while more language-focused criteria, student interpretation, and task-focused criteria were used after professional development. Within the information-focused criteria, the dominant criteria

in the baseline phase were the correctness of information because Cassie elicited evidence of learning using mainly selected-response questioning. What she did for interpretation was to map student answers to the answer key. After CPD, however, her criteria focusing on information was changed to providing information and relevancy, indicating that her interpretation started to focus on more aspects of student ability.

Cassie made more use of language-focused criteria after CPD (from 14 times to 24 times), mainly focusing on language accuracy and pronunciation. Moreover, the students were involved in self-assessment and peer-assessment (increasing from 0 to 11 times after teacher learning). For most of the time in class before CPD, the students listened to their teacher and occasionally answered questions. After the CPD, student interpretation was introduced in the class, the students were actively involved in peer-assessment and self-assessment. One student mentioned this in the interview at the end, that student interpretation could make him more concentrated in class. Another student expressed her interest of being involved in self- and peer-assessment in class. Moreover, Cassie also started to interpret the completion and difficulty of the tasks (increased from 1 to 10 times) after the CPD, monitoring student progress in moving towards the learning objectives.

3.4 Providing Feedback

Overall frequency. Cassie provided more feedback to her students after professional development, with the frequency of feedback practice increasing from 248 times in the baseline phase to 351 times after CPD. More specifically, the feedback foci and the use of feedback techniques increased, with the frequency increasing from 91 to 169 times and from 157 to 182 times respectively.

Feedback focus. Table 4 shows that Cassie’s feedback for classroom functions dropped from 31 times in the baseline phase to 18 times after teacher learning. Feedback for subject

Table 4. *Changes in Feedback Practices*

Providing Feedback	Before CPD	After CPD
Feedback focus	91	169
Classroom function	31	18
Warming up	28	10
Classroom management	0	0
Requests or directions	0	0
Learning check	3	8
Subject core competencies	60	151
Language ability	39	58
Cultural awareness	0	20
Thinking capacity	3	20
Learning ability	18	53
Feedback technique	157	182
Task-referenced	115	153
Person-referenced	40	22
Management	2	7

core competencies increased dramatically from 60 to 151 times, with feedback for learning ability increasing the most, from 18 to 53 times.

Feedback technique. More task-referenced feedback and management were provided to students, while fewer person-referenced feedback practices were found after professional development. The frequency of task-referenced feedback practices increased from 115 times in the baseline phase to 153 times after CPD. The types of task-referenced feedback that increased the most were awareness raising, further probing, recast and supplementing information. Additionally, the frequency of management (e. g., reminding, and task management) also increased after CPD. On the contrary, Cassie provided less person-referenced feedback (e.g., praising and thanking) than the baseline phase.

In the first round of action, Cassie suddenly stopped praising her students, and started to provide task-referenced feedback. The students felt discouraged and hesitated to answer the questions. Cassie was frustrated by her students' reactions in that lesson and felt that she suddenly did not know how to teach after teaching for 13 years. Cassie reflected on the lesson using the classroom video recordings and shared her feelings with the team. In the second round of action, she changed her feedback by using expressions that were more acceptable to the students (e. g., by summarizing, reinforcing, reminding, and providing supplementary information), rather than pointing out mistakes directly. Her students quickly adjusted and became active in the lesson. After class, Cassie talked to several students about their thoughts. Cassie was very delighted that her students thought the feedback she gave in class was helpful. A boy said it aptly in the post-interview: "Although it takes time to get used to, corrections and suggestions for improvement are always more useful than praising only. Good medicine doesn't taste good" (S2-SIPost7). It took time and efforts for Cassie and her students to step out of their routines, but they became more confident in the second round of action.

3.5 Follow-Up Action

No follow-up action was found in Cassie's baseline lessons. During the professional development, she became aware that follow-up actions make the cycle of formative assessment practice complete. Table 5 shows the frequency of follow-up practices increased from 0 in the baseline phase to 47 times after the CPD. More specifically, "action focus" increased from 0 to 31 times, and "action technique" increased from 0 to 16 times. Cassie's actions mainly focused on learning ability because several learning strategies were covered in the two lessons during the action research. However, there were 12 follow-up actions focusing on language ability that were not closely related to the explicit targets of teaching. For example, the target for a task was developing learning strategy of classifying and organizing information. Cassie caught a wrong pronunciation in a student's utterance, and the follow-up action was about pronunciation and vocabulary knowledge.

During the CPD, Cassie started to provide opportunities for the students to take some actions to close a learning gap. Planned follow-up tasks and activities were organized to provide opportunities for improvement and consolidate the learning target. Scaffolding was taken to help students achieve the learning goals. She also provided actions on language accuracy and pronunciation, which were not closely related to the target of learning. It should

Table 5. *Changes in Follow-Up Action*

Follow-Up Action	Before CPD	After CPD
Action focus	0	31
Language ability	0	12
Cultural awareness	0	4
Learning ability	0	15
Action technique	0	16
Language focused	0	5
Scaffolding	0	2
Opportunity for improvement	0	9

be noted that although Cassie had clear teaching goals for each lesson during the action research, she tended to switch among the targets contingently inside the classroom.

4. Discussion

This paper has presented one English language teacher’s growth in assessment literacy for classroom-based formative assessment during a 12-week teacher continuing professional development program. The CPD program was found most useful in changing this teacher’s knowledge and beliefs about formative assessment. It was also encouraging to see the improvement in classroom practices that could be used formatively for student learning. Specifically, towards the end of the 12 weeks, Cassie had a comprehensive understanding of assessment, and knew that assessment could be used to help students learn. She had clear, systematic and feasible long-term, mid-term, short-term English language teaching goals. She clarified learning targets at the beginning of each lesson. She started to provide criteria for success by using checklists. She mentioned that more task-referenced feedback should be provided to her students in class to reach the target of learning. Most importantly, she took follow-up actions and provided opportunities for her students to close the learning gap that was identified in class.

4.1 The Importance of Target-Oriented Formative Assessment

In classroom-based formative assessment, clear teaching and learning objectives play an important guiding role and provide a reference for checking the successfulness of classroom teaching and learning. Therefore, teachers should be clear about the “what” of assessment, which would help them target their teaching and assessment, effectively obtain teaching information, reflect on teaching methods, ensure teaching quality, and help students optimize their learning strategies and improve their learning results.

At the end of the 12-week CPD program, one of the main improvements was that the teacher started to set concrete goals not only for a period of teaching, but also for each lesson. Cassie clarified learning targets to her class and chose, designed, and used assessment tools based on the targets. She also used the targets of teaching, learning, and assessment to guide every step in formative assessment, namely eliciting evidence, interpreting the evidence, providing feedback, and taking follow-up actions. This made it possible to use formative assessment to achieve the targets, rather than completing cycles of formative assessment

practices aimlessly.

Despite the clear targets, however, implementing formative assessment inside the classroom revealed a complex picture. Very often, a particular performance problem such as pronunciation captures the teacher's attention on the spot. Her instantaneous interpretation would lead to feedback and/or follow-up action. In other words, these unplanned, informal, and contingent assessment practices can be regarded as formative assessment, and yet the assessment is not related to the explicit target of teaching for the session. Technically, these formative practices will not help close the gap between what is assessed and the explicit target of the lesson, no matter how many cycles of formative practices are found. That said, not everything that comes to the teacher's attention in class is worth the same treatment in class. An experienced teacher should be constantly picking and prioritizing assessment targets inside the classroom and making decisions in choosing and switching between planned and contingent assessment targets. This is exactly where pedagogical content knowledge comes into effect.

Another problem is the interpretation of learning targets. What constitutes a learning target has been clearly stated in the curriculum standards. However, these become vague in the textbooks and in the classroom. It is not clear how "learning targets" in the textbook like "read about the virtual choir; or listen to people talk about their music preferences" will lead to the students' growth of linguistic and functional competence outlined in the curriculum. It is interesting to note that the "teaching objectives" listed in the Teachers' Books are in general aligned to the curriculum standards. When asked about whether she read the objectives in the Teachers' Book, Cassie said she would normally focus on the instructional suggestions and the answer sheets, but not the "teaching objectives" section.

Placing assessment into the hands of classroom teachers pushes them to take a step back and reflect upon the target of assessment. This in turn raises their awareness about the need to align their teaching, learning, and assessment goals. The explicit focus on assessment targets during the group reflection and planning sessions was useful in helping Cassie make her teaching targets clear to her class in each lesson and in relating these targets to the curriculum standards. This awareness of teaching, learning and assessment integration made it possible for Cassie to deliberately employ formative practices in class towards the second half of the 12 weeks.

4.2 Formative Assessment Cycles

The most observed pattern of Cassie's classroom assessment practice is some sort of elicitation or observation, followed by a quick interpretation with or without an explicit affirmation or praise. In other words, most of the time, she did not complete all the four steps in the formative assessment cycle. For example, when a student failed to answer a question, Cassie would ask another student the same question. In fact, not everything elicited was interpreted; not everything interpreted received feedback; and very few follow-up actions were found. While this is in line with previous research (Gu & Yu, 2020), a closer look at Cassie's assessment practices reveals that follow-up activities were often not needed, and therefore, the usual assumption, the more complete the cycles, the better (Ruiz-Primo & Furtak, 2006),

needs to be revised.

Inside the classroom, it was found to be a common practice for Cassie to re-start the assessment cycle before any follow-up action was taken, just to make sure that the interpretation on the spot (e. g., something has not been learnt) was correct. Repeated assessment like this can be argued to be a good thing, and probably a must, in order to arrive at a cautious and reliable conclusion, before a final decision is made for follow-up actions. After all, information embedded in one or two instances of learning evidence elicited through classroom tasks should not be enough for any definitive judgment of where students are in relation to a learning target. In fixed-format educational tests, a large number of items need to be completed before an inference is made about a student's level of performance based on a cumulative score. In classroom-based formative assessment, while teachers do not often have the luxury of multiple judges, we do need to be sure as far as possible that the judgment and decision we make on the spot are accurate and reliable. One way to ensure this is multiple observations and multiple interpretations.

Whether a step after interpretation in the formative assessment cycle was needed very much depended on the difficulty level of the task being assessed. Most of the time the contingent tasks being assessed in Cassie's classes were very simple, and she quickly arrived at the conclusion that the students had already reached the target, so she immediately moved on to another task. One take-away from this phenomenon is that assessment tasks need to be set at a level slightly beyond the students' current ability, in order for assessment to reveal a learning gap that is worthy of follow-up efforts. Task difficulty pushes the assessment cycle closer and closer towards the learning target, making the formative learning process spiraling cycles. In other words, classroom contingent assessment is happening all the time with the teacher's observations and judgments. However, many observations will end up with a judgment which does not lead to further action on the task. Only challenging tasks beyond the learners' current level of performance can push learners towards the next step of learning; and only completing the entire cycle makes assessment formative.

4.3 The Skill Dimension of Teacher Learning

This study also found that the beliefs about and knowledge of assessment were relatively easy to change. However, teacher learning in assessment literacy is also a matter of skill, and is therefore a slow process, which may even be clumsy at an initial stage. Skill learning (Anderson, 1982) normally begins with a cognitive stage where the facts and rules are encoded as declarative knowledge. It then proceeds to an associative stage when the declarative knowledge becomes proceduralized, successful performance gets reinforced, and errors get reduced. As the new skill gets repeated multiple times, performance is gradually automatized in the autonomous stage. Seeing the teacher's development of assessment literacy as skill learning helps explain why lectures and workshops alone would not work. CPD is, by nature, contextualized within the teacher's daily work. An action research model integrating both declarative and procedural knowledge of assessment during a whole semester allows the teacher to grow her assessment literacy through not just reading and planning, but also trial, error, reflection, amendment, and automatization.

In the slow and tedious process of de-learning, relearning, and habit formation, both teachers and their students may experience emotional frustrations beyond their control, when routines of classroom behaviors change and when they step out of their comfort zones. In one of Cassie's lessons, she tried not to give empty praises and instead focus her feedback on task-specific information that she believed would lead to improved task performance. One girl started to cry after the teacher gave her feedback on what she did wrong. Cassie also voiced her frustration after class and said that she suddenly lost her confidence as a teacher.

The process of teacher learning is also social and dynamic. After the baseline phase, Cassie asked the first author to comment on her lessons and provide suggestions. She talked with the researcher individually for one hour and wrote a detailed reflection journal after thinking and summarizing the experience using the recorded lessons. Cassie also interacted a lot with her students in class and after class. She was willing to listen to her students' opinions and suggestions about some of the changes she had made in class, and then she took the suggestions into account for further improvement. As a result, Cassie achieved visible growth as a teacher at the end of the study. She even started sharing with her colleagues about formative assessment practices a few months after the study ended. One year later, Cassie was promoted due to her outstanding achievements.

5. Conclusion

With formative assessment being explicitly written into curriculum guidelines, there is no doubt that in-service teachers need continuing professional development in developing their formative assessment literacy. Two rounds of action research in this study showed that it can help teachers develop their classroom-based assessment literacy. For thousands of secondary school teachers, a model that supports teacher-initiated reflective practice and community of learning with minimum external guidance and facilitation should be especially valuable. In this sense, we hope this study could serve as a model for secondary school teachers in and beyond China where most teacher education resources focus on pre-service teacher education programs.

Despite the obvious limitation of a one-person case study, for researchers and teachers interested in classroom-based formative assessment, this study provides new theoretical insights as well as practical procedures for the realistic enactment of research results in the classroom. One contribution is the placement of domain-specific learning targets at the very heart of formative assessment, without which formative practices will be meaningless. Another contribution is the emphasis on multiple sources of learning evidence and multiple judgments (informing) before deciding on the necessity and type of feedback and follow-up actions (forming) (Davison & Leung, 2009). Many times, classroom assessment results may not lead to the discovery of a learning gap. In these circumstances, there is no need for formative action. Insights like these give the innovative teacher much more confidence than the demand for complete cycles of formative assessment.

References

American Federation of Teachers, National Council on Measurement in Education, & National Education

- Association. (1990). *Standards for teacher competence in educational assessment of students*. <http://buros.org/standards-teacher-competence-educational-assessment-students>
- Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review*, 89(4), 369-406. <https://doi.org/10.1037/0033-295X.89.4.369>
- Andersson, C., & Palm, T. (2017). The impact of formative assessment on student achievement: A study of the effects of changes to classroom practice after a comprehensive professional development programme. *Learning and Instruction*, 49, 92-102. <https://doi.org/10.1016/j.learninstruc.2016.12.006>
- Andrade, H. L., Brookhart, S. M., & Yu, E. C. (2021). Classroom assessment as co-regulated learning: A systematic review. *Frontiers in Education*, 6. <https://doi.org/10.3389/educ.2021.751168>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. <https://doi.org/10.1080/0969594X.2010.513678>
- Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. Open University Press.
- Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148. <https://doi.org/10.1177/003172171009200119>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement, Issues and Practice*, 30(1), 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.
- Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly*, 43(3), 393-415. <https://doi.org/10.1002/j.1545-7249.2009.tb00242.x>
- Day, C., & Sachs, J. (2005). Professionalism, performativity and empowerment: Discourses in the politics, policies and purposes of continuing professional development. In C. Day & J. Sachs (Eds.), *International handbook on the continuing professional development of teachers* (pp. 3-32). Open University Press.
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2016). Teacher assessment literacy: A review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251-272. <https://doi.org/10.1007/s11092-015-9233-6>
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. <https://doi.org/10.1080/15434303.2011.642041>
- 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. <https://doi.org/10.1080/0969594X.2013.836076>
- Gu, P. Y. (2021). *Classroom-based formative assessment*. Foreign Language Teaching and Research Press.
- Gu, P. Y., & Yu, G. (2020). Researching classroom-based assessment for formative purposes. *Chinese Journal of Applied Linguistics*, 43(2), 150-168. <https://doi.org/10.1515/CJAL-2020-0010>
- Gu, Y. (2012). English Curriculum and Assessment for Basic Education in China. In J. Ruan & C. B. Leung (Eds.), *Perspectives on Teaching and Learning English Literacy in China* (pp. 35-50). Springer. https://doi.org/10.1007/978-94-007-4994-8_3
- Guskey, T. R., & Yoon, K. S. (2009). What works in professional development? *Phi Delta Kappan*, 90(7), 495-500. <https://doi.org/10.1177/003172170909000709>
- Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, 89(2), 140-145. <https://doi.org/10.1177/003172170708900210>
- Herppich, S., Praetorius, A. -K., Förster, N., Glogger-Frey, I., Karst, K., Leutner, D., Behrmann, L., Böhmer, M., Ufer, S., Klug, J., Hetmanek, A., Ohle, A., Böhmer, I., Karing, C., Kaiser, J., & Südkamp,

- A. (2018). Teachers' assessment competence: Integrating knowledge-, process-, and product-oriented approaches into a competence-oriented conceptual model. *Teaching and Teacher Education*, 76, 181-193. <https://doi.org/10.1016/j.tate.2017.12.001>
- Hill, K., & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 395-420. <https://doi.org/10.1177/0265532211428317>
- Kennedy, A. (2005). Models of continuing professional development: A framework for analysis. *Journal of In-Service Education*, 31(2), 235-250. <https://doi.org/10.1080/13674580500200277>
- Kennedy, A. (2014). Understanding continuing professional development: The need for theory to impact on policy and practice. *Professional Development in Education*, 40(5), 688-697. <https://doi.org/10.1080/19415257.2014.955122>
- McMunn, N., McColskey, W., & Butler, S. (2004). Building teacher capacity in classroom assessment to improve student learning. *International Journal of Educational Policy, Research & Practice*, 4(4), 25-48.
- Ministry of Education of the People's Republic of China (MOE). (2020). *General Senior High School Curriculum Standards*. People's Education Press.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48 (1), 4-11. <https://doi.org/10.1080/00405840802577536>
- Popham, W. J. (2017). *Classroom assessment: What teachers need to know* (8th ed.). Pearson.
- Ruiz-Primo, M. A., & Furtak, E. M. (2006). Informal formative assessment and scientific inquiry: Exploring teachers' practices and student learning. *Educational Assessment*, 11(3), 205-235. <https://doi.org/10.1080/10627197.2006.9652991>
- Sachs, J. (2001). Teacher professional identity: Competing discourses, competing outcomes. *Journal of Education Policy*, 16(2), 149-161. <https://doi.org/10.1080/02680930116819>
- Sachs, J. (2016). Teacher professionalism: Why are we still talking about it? *Teachers and Teaching*, 22(4), 413-425. <https://doi.org/10.1080/13540602.2015.1082732>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15 (2), 4-14. <https://doi.org/10.3102/0013189X015002004>
- 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. <https://doi.org/10.1111/j.1745-3992.1999.tb00004.x>
- Taylor, L. (2013). Communicating the theory, practice and principles of language testing to test stakeholders: Some reflections. *Language Testing*, 30(3), 403-412. <https://doi.org/10.1177/0265532213480338>
- Timperley, H., Wilson, A., Barrar, H., & Fung, I. (2007). *Teacher professional learning and development: Best evidence synthesis iteration*. Ministry of Education.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: Teaching, learning and assessment in the classroom*. Open University Press.
- 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. <https://doi.org/10.2307/20466661>
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *Australian Educational Researcher*, 40(2), 241-256. <https://doi.org/10.1007/s13384-013-0089-9>
- Xu, Y., & Brown, G. T. L. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162. <https://doi.org/10.1016/j.tate.2016.05.010>

About the authors

Jiayi LI is currently a PhD candidate at the School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand. Her research interests include language testing and assessment, language teaching methodology, and teacher development. Email: jiayi.li@vuw.ac.nz

Peter Yongqi GU is Associate Professor at the School of Linguistics and Applied Language Studies, Victoria University of Wellington, New Zealand. His research interests include language learning strategies and language assessment. His recent book, *Classroom-Based Formative Assessment* (FLTRP) aims to help language teachers develop assessment literacy for formative purposes. Email: peter.gu@vuw.ac.nz

Chinese Abstracts

中文提要

发展基于课堂的形成性评估素养

顾永琦 (新西兰)惠灵顿维多利亚大学

林志强 (中国)香港浸会大学

摘要: 目前,形成性评估用于改善教与学、提高教育质量的功能日益得到认可,已被写入政策文件和课程标准,并纳入教师教育计划。但就教师而言,在课堂上系统地实施形成性评估仍然是一项艰巨任务,因此,如何提高教师的形成性评估素养成为大家共同关心的问题。本文是本期特刊的介绍,首先,通过强调形成性评估素养所涉及的主要问题对本期特刊的定位及其涵盖范围进行介绍;其次,简要介绍收录于本期特刊的十篇文章。笔者希望本期特刊能够对促进教师形成性评估素养的理解及其发展起到抛砖引玉的作用。

关键词: 形成性评估;教师评估素养;教师教育

未来评估素养展望

苏珊·布鲁克哈特 (美国)杜肯大学教育学院

摘要: 本文对以下两个问题进行了讨论:(1)未来更好的评估与当前的评估环境有何不同?(2)未来评估素养会是什么样子? 答案将基于最近对评估领域研究的大规模回顾,为教师的教学和专业发展工作提供信息。本文将聚焦评估中正在发生以及将来会发生的变化,探究这些变化对教师、学校领导和学生的评估素养需求的影响。

关键词: 评估素养;评估未来;形成性评估

评估素养:改变文化,并让改变适应香港的文化

克莉丝·戴维森 (澳大利亚)新南威尔士大学教育学院

摘要: 近年来,教师评估素养因其在学习和教学中的关键作用而受到广泛关注。随着最近对教师知识和技能建设的重视,相关理论和实证概念化不断涌现。其目的是做出高度情境化、公平、一致和值得信赖的评估决定,为学习和教学提供信息,从而有效地支持学生和教师的学习。本文探讨了在引入校本学习评估作为香港高风险中学评估改革的一部分后,中学英语教师评估素养的变化。本文利用近 4,500 名参加过评估改革头六年(2005-2011年)共同专业发展课程的教师完成的教师问卷调查,探讨了专业发展对教师评估素养的影响。尽管香港有根深蒂固的竞争性应试文化,对课程前后评估的定量和定性数据的分析表明,教师的态度、信心和实践发生了积极变化,特别是在使用评估标准、设计和实施方面适当的评估任务,让学习者更积极地参与评估过程,做出值得信赖的评估决定,并向学生提供有效的反馈和前馈,以改善学生的学习。本文研究的结果表明,如果教师得到很好的支持,评估文化的改变是可能的。本文还讨论了对评估改革和更普遍的教师评估素养发展的影响。

关键词: 评估素养;学习评估;校本评估;语言评估

发展教师课堂形成性评估素养——以一位中学英语教师为例

李加义/顾永琦 (新西兰)惠灵顿维多利亚大学

摘要: 自从Black和Wiliam(1998)发表了极具影响力的研究综述,形成性评估在教育改革中愈发受到重视。21世纪初,中国颁布的英语课程标准明确指出教师应在课堂开展形成性评估。教师评估素养的缺乏制约了课堂形成性评估的开展。本研究设计并实施了为期12周的教师课堂形成性评估素养提升项目,旨在帮助五位中学英语教师提高形成性评估知识、信念和实践水平。项目以合作行动研究的形式开展,研究者与教师组成了实践共同体,以实现持续专业发展。本文以一位教师为例,介绍其课堂形成性评估素养的提高。作者使用NVivo 12对课堂录像和访谈进行转录和编码。基线数据表明,该教师缺乏开展形成性评估所需的知识、信念和实践方法,尤其缺乏明确的教学目标。此外,其评估目标并非提高学生核心素养,而课标明确要求高中英语课程目标是培养和发展学生的学科核心素养。项目结束时,该教师的形成性评估知识、信念和实践方法均得到显著提升;在课堂形成性评

估实践方面,其课堂教、学、评的目标设定与学生核心素养的培养紧密结合,实现更多形成性评估闭环,帮助学生达到学习目标。这些发现为教师提高形成性评估素养提供了令人鼓舞的支持,也证明教师在持续专业发展框架内提高评估素养的可行性。

关键词: 评估素养;形成性评估;教师学习;持续专业发展;合作行动研究

大学英语新手教师测评身份建构——两位新手教师的个案

甘凌 北京工商大学 / 香港浸会大学

林志强 香港浸会大学

摘要: 语言测评素养发展研究日益兴盛,但学界对教师测评身份建构知之甚少。本研究通过叙事框架、访谈和观察,探索两位大学英语新手教师如何构建测评身份。研究发现,两位教师随着语言测评素养的发展,经历了从不认同测评者身份到成为测评者的转变,并以不同轨迹构建其测评身份。一位教师积极主动构建其坚定的测评实践者及探索者身份,但更认同自己为引导者及“磕磕绊绊的新手测评者”。另一位则成为自信的反馈者和宽容的评分者,但更认同自己“人生导师”的身份。两位教师的测评身份建构主要受其测评观念、测评经历、个人因素(如自我效能、主动性、反思)及学校因素的影响。本研究为语言测评身份建构研究以及如何发展成为有效测评者提供启示。

关键词: 新手大学英语教师;测评身份;语言测评素养;职业身份建构

个人、经历和环境因素如何影响英语教师的评估观——一项叙事研究

许悦婷 / 陈嘉雨 / 何莉怡(通讯作者) 华南师范大学外文学院

摘要: 已有研究表明教师的评估观受其个人特征、经历和特定的社会文化背景影响,但这些因素如何影响教师的评估观尚未得到充分研究。本研究基于48名中国高中英语教师的叙述框架数据,探究教师评估观及其影响因素。结果显示,教师普遍认同以促学为导向的评估目的;一系列因素对教师的评估观产生了不同程度的影响,包括个人因素(教师的主观能动性)、经历因素(教师对评估的[反]学徒式观察)以及环境因素(应试文化)。本文进而讨论了教师如何在促学导向的评估目的和问责导向的评估目的之间进行磋商与妥协,以及教师主观能动性、作为被评估者的情感体验和应试文化如何塑造教师评估观。研究对教师评估观研究以及教师评估素养政策、实践和专业发展具有重要启示。

关键词: 教师评估观;影响因素;中国英语教师;叙事研究

教师对评估反馈的不同理解与实践——对教师评估素养的启示

吴瑞萍 新加坡英语学院

陈兴杰 新加坡国立教育学院学习科学与评估学术组

摘要: 评估反馈是教师评估素养的一个重要方面,可以从三个相互关联的维度理解此概念:教师对评估反馈的理解,评估反馈的课堂实践,以及教师如何处理学生参与评估过程中所产生的各种情绪(Pastore & Andrade, 2019)。本研究运用现象学研究方法,探讨了15名新加坡教师对评估反馈的理解与实践。借鉴变异理论的观点,本研究从访谈数据的分析得出了教师对于评估反馈的五种诠释,揭示了教学中的评估反馈并非专属教师的输出,而是师生之间探讨评估反馈的过程。教师对于评估反馈的理解能够从他们提供的不同性质的反馈中体现出来。本文建议将教师提供的评估反馈分成五类,最基础的为强调学生错误的反馈,最理想的为促进学生进行反思的反馈。调查结果显示,教师可以根据教学环境和学生的不同需求,调整评估反馈的性质与作用。本文也讨论了研究结果对于发展教师评估素养的意义,以帮助教师审查他们对评估反馈概念的理解,并希望能帮助教师加强评估反馈的实践,使评估反馈的使用不局限于提高学术表现这一层次,而是使评估反馈成为提升学生自我调节能力的途径。

关键词: 评估反馈;教师评估素养;促进学习评估;反馈伙伴关系

同伴反馈圈:三人小组中口头观察同伴反馈在学生反馈素养发展中的价值

玛丽·杨 (新加坡)东南亚教育部长组织区域语言中心

摘要: 关于同伴反馈有效性、尤其是学习者积极参与反馈并根据反馈采取行动的最佳方式的争论越来越多(Lam, 2017)。研究表明,妥善处理的同伴反馈可以带来许多社会、认知和语言方面的好处(Sun & Doman, 2018)。本文研究了24位来自缅甸的英语教师对于同伴反馈圈(PRC)的看法。在PRC中,学习者以三人一组的形式工作,其中两人对第三个人的表现提供反馈,第三个人倾听、反思,然后才对反馈意见做出回应。结果发现,教师认为这种安排可以增加反馈的数量、多样性和可靠性,营造更有利于反馈的氛围,并有助于发展“软技能/通用技能”。受访者认为,教师和学生对这种新方法不熟悉、时间限制以及可能出现相互矛盾的反馈意见是PRC的潜在困难。