

Integrating CSE and Assessment for Learning in Chinese EFL Writing Classroom

Jiuliang Li
Beijing Institute of Fashion Technology, China
(Email: lijium@hotmai.com)

Received: 26 July 2023; Accepted: 1 October 2023; Published: 26 December 2023
<https://doi.org/10.58304/tc.20240107>

Abstract

The China's Standards of English Language Ability (CSE) is developed to serve as a guide for English language teaching in China. Assessment for learning (AFL) serves as an alternative assessment approach that maximizes student learning. They are closely linked in the sense that both claim to serve the needs of teaching and learning. Just as AFL practices in China are generally lacking, so is CSE in need of more research for use in educational context. In this paper, I describe a two-semester intervention at a Chinese university where I attempt to use a CSE-based scoring rubric in combination with the major AFL activities to foster development of writing skills of students who learn English as a foreign language (EFL) in the college English class. After a review of the related literature, I describe the context in which I implemented the project that incorporated CSE rubric into major AFL activities, i.e. peer/self-assessment, and how I implement AFL guided pedagogical activities. Based on the evidence collected in the interim, a temporary conclusion is drawn that the students, as well as the teacher researcher, reap benefits from the intervention project. Finally, I give some suggestions for future exploration.

Keywords

Assessment for learning, China's Standards of English Language Ability, rubric, peer/self-assessment, EFL writing

Introduction

Writing is possibly the most formidable language skill for L2 students (Hyland, 2003). Facing the various challenges, educators and researchers have been looking for viable ways to promote the development of L2 writing skills. In recent years, regions including the United Kingdom, Australia, and Hong Kong pay increasing attention to assessment for learning (AfL) in curriculum policy statements out of the consideration that AfL is educationally more desirable and plausible than formal testing. In AfL, learning is a goal in its own right, and assessment is a means to achieve that goal (Assessment Reform Group, 2002) by collecting information that enables teachers to modify teaching and encourages students to improve learning (Lee & Falvey, 2014). Lee (2007) rightly proposed that AfL need to be espoused in the second and foreign language writing classroom through process writing, self- and peer feedback, student-teacher conferences, portfolios, etc.

In China, the Chinese Standards of English Language Ability (CSE) (NEEA, 2018) offers another alternative for coping with the challenges confronting the learners and the teachers. Designed within and for China's specific context, CSE is applicable to assessment of English language and can be used as a yardstick for English language teaching and learning. It is a useful tool through which language learners and teachers know where they are and where they need to go along the process of language development. With a range of criteria of English

proficiency that could be employed to inform teaching, learning and assessment, CSE serves as a guideline for English language education in China (Liu, 2017). For example, it could help teachers and authorities develop assessment methods for evaluating teaching and learning (Liu, 2015; Wang, 2018). Liu (2017) maintains that “CSE promotes the implementation of formative assessment where learners take the major responsibility for learning” (Liu, 2017, p.6).

In addition to the shared interests and commonalities briefly discussed above, both AfL and CSE emphasize, among other things, the role of teaching and learning, and the importance of goal and assessment in the language classroom. It is therefore reasonable to infer that AfL and CSE effectively complement each other in promoting learning in writing classroom. However, just as AfL practices in China are generally lacking (Liu & Xu, 2017), so is CSE in need of more empirical research for practical use in educational context (Liu, 2017). In this paper, I describe a two-semester intervention in my English class at a Chinese public university where I attempted to use a CSE-based scoring rubrics in combination with major AfL activities to foster development of my students’ EFL writing skills.

Expected Benefits and Potential Constraints

Expected benefits

AfL encompasses any assessment for which the foremost objective in its design and application is to serve the purpose of improving students’ learning, and such assessment becomes “formative assessment” when the evidence and information are gathered to effectively adapt the teaching plan to learning needs (Black et al., 2004). In AfL, assessment is the focus but learning is the goal (Gardner, 2006). It carries potential for transforming teaching and learning processes in ways that enhance learning outcomes (James & Pedder, 2006). With respect to writing, the potential is assumed to be even greater (see Black & Wiliam, 2006; Clarke, 2005) as AfL stimulates a re-consideration of instruction, classroom practices and processes – e.g. how standards and criteria are developed and presented to learners, how to engage learners actively (e.g., through self- and peer assessment), and how to promote their motivation in learning (Lee & Falvey, 2014).

Over the last three decades, evidence in support of learning has grown about the power of assessment (Wiliam, 2017), especially for AfL. The extensive use of these assessment practices by teachers and learners have proven to be effective in promoting the development of learners’ metacognition, self-regulation, and autonomy (Assessment Reform Group, 2002; Earl, 2013; Willis, 2010). In recent years, there has also been substantial evidence from AfL research demonstrating its positive impact on student learning in terms of both motivation and actual performance (Black & Wiliam, 1998, 2003; Sebba 2006; Sebba & Maxwell, 2005). AfL has been used in China for over 20 years. However, issues remain unclear as to how it should be conducted and interpreted in classroom practice, especially in mainland China.

Potential constraints

Cultural values

Active student involvement is encouraged in writing classroom that emphasizes AfL (Lee, 2017). However, it may not be easy for students in EFL contexts, China in particular, where teaching and learning tends to polarize as knowledge imparting and receiving due to influence of the Confucian heritage culture. The conventional teacher-centered pedagogy suffers a lack of productive teacher-learner dialogues and results in students’ favoritism over teacher feedback (Tsui & Ng, 2000; Yang, Badger, & Yu, 2006). Students also fear to reveal their misconceptions in peer/self-feedback lest they lose the respect from their teacher and peers in the classroom (Yin & Buck, 2015). In addition, their’ concern for face-saving and maintaining

harmony may keep themselves from providing constructively critical feedback to evade tension and disagreement (Carson & Nelson, 2006; Cheng & Warren, 2005).

Pressure from examinations

In China, where an examination culture dominates, the implementation of AfL activities in the educational institutions meets with great impediments (Berry, 2011). Confronted with the pressure of high-stakes external examinations (Hui et al., 2017; Tan, 2016; Yu, 2015), teachers who subscribe to AfL cannot fully implement these methods, or give them up for more summative assessment practices, due to the need of students to prepare for these high-stakes examinations. Wang and Wang (2011), for example, found that when the non-key universities in China under investigation prioritize preparing students for College English Test Band 4/6 (CET4/6) in their curriculum, teachers and students are generally reluctant to perform AfL and its related practices. Students' single-minded emphasis on examinations is also deemed as an obstacle when they fail to recognize the connections between what they learn and what are tested in the exams (Koh et al., 2015; Lin et al., 2013).

School level constraints

There are "entrenched school practices" that form the main hurdle to the full deployment of AfL in writing (Lee, 2017). These include a compulsory policy to stick to the conventional rule that emphasizes marking of students' errors in writing, leading to tension, frustration, and burnout among teachers, as well as negative effects on students psychologically (Lee & Falvey, 2014). Some teachers regard AfL as a good learning strategy, but are reluctant to use it as they are confined to institutional values and school culture that favor testing (Yu, 2015). There is also the ethos, on the part of both learners and instructors, that scoring and grading remain a 'valued part of learning', which, however, is inconsistent with AfL (Azis, 2015).

Class-size and curriculum

AfL would be more effective for smaller class sizes which are considered amenable to the development of more favorable environment for teacher-student and student-student interactions which are crucial to AfL. Unfortunately, in Chinese universities and colleges, class of 50 students or more is common for non-English majors. The plain fact is that the large class size might have adverse influence on AfL practice (Liu & Xu, 2017). In addition, teachers have to cover the jam-packed syllabus under the pressure to cover all prescribed curriculum materials because of syllabus requirement which further debilitates interaction in the classroom (Lee, 2017). Closely related to curriculum and class-size, time constraint constitutes another debilitating factor, given the fact that most AfL strategies, such as peer and teacher feedback, portfolio assessment, etc., are time-consuming (Lee, 2017).

These factors, and many others at play, discourage Chinese English educators from practicing the alternative assessment, which in turn results in the dearth of research on AfL in the EFL writing classrooms and the situation that the existent limited research is mainly found in secondary and college EFL contexts in Hong Kong and Taiwan.

Context of the Study

In the present study, I explored with my students for viable ways of conducting AfL practices in our English class. The 88 non-English majors were in their second year of study. In the first two years on campus they took College English course where they and I, their teacher, met twice a week, each time for 100 minutes (two class sessions with a 10-minute interval). The course consisted of two modules including Reading & Writing, Listening & Speaking. Every volume of the textbook for Reading & Writing consisted of 8 units, each of which contained

two reading texts accompanied by exercises. The texts took various genres (exposition, argumentation, narration, and description). For each unit within the textbook, according to the curriculum, the Reading & Writing module took 4 sessions, and the Listening & Speaking module 2 sessions. There was no separate writing course in the curriculum, and there was certainly not enough room for AfL practice due to the tight schedule. Each unit within the textbook generally provided less-than-one-page “structure analysis” that discussed the genre and pattern of the reading material followed by a writing task to practice the said genre and pattern.

Applying CSE Informed AfL Pedagogical Principles

James and Pedder (2006) have it right that the change of classroom processes means transformation in what teachers and learners do, and the focus is particularly on transformation in pedagogical practice. Lee (2007) early suggested that “A key premise is that the implementation of AfL starts with awareness of the connection between teaching and learning and specifically how assessment can be used to inform teaching and learning” (p. 203). Later she proposed five pedagogical principles that underlie effective AfL practices (2017), which I followed to carry out instruction and assessment. In the interim, I always used the rubrics that were developed based CSE (hereinafter the CSE rubrics) as a guide to inform my instruction in the pre-, during, and post-writing sections. The CSE rubrics served as a map with many signpost descriptors that provided the framework and direction my students and I needed in our process of learning, both in reading the texts as inputs and in writing scripts as outputs.

Pre-writing instructional scaffolding

The essay tasks in the textbook took various patterns and genres including exposition (comparison and contrast, cause and effect), description, argumentation, and narration. Although not explicitly stated, the textbook generally required a five-paragraph essay from the students: introduction, three main ideas, and conclusion. This was more so in the case of exposition and argumentation as was exemplified in the sample essay provided in the textbook. These patterns and structures aligned with those introduced in Langan (2014), which was being used as the writing textbook in some key universities in China. When the textbook did not provide sufficient content on teaching writing, I usually resorted to Langan (2014) for supplemental material to provide more detailed information and a more practical approach to classroom instruction and assessment.

I first presented the writing process (Langan, 2014) to the class to familiarize the students with the process approach of writing which includes pre-writing, writing a first draft, revising, and editing. In particular, I used Cindy’s experience of writing her description essay of “family portrait” (Langan, 2014 p.185-188) as an example to illustrate the whole process of completing a descriptive essay task so that my students would develop a better understanding and knowledge of the whole writing process and how to carry it out in their own task. For the pattern and genre targeted in a given unit in the textbook, I aligned the task to the textual features used in the reading material within the textbook that the students were studying in the said unit. To improve their sentence skills, I encouraged the students to pick out expressions and sentence patterns from the textbook, analyze the surface features and their intended use. If time permitted, I gave the students translation and sentence making exercises that involved these language features, and urged the students to apply them in the writing tasks. These were done out of my understanding that attending to language use would be an important language learning activity to my students, who were still learning their L2. I observed that some students attempted to apply the sentence patterns in their follow-up writing tasks. A good example is the pattern of “be it...or...” which was attempted successfully by one student. Reading helped

the learners recognize the textual features in a given genre, and applied use in their own writing would help them improve awareness and mastery of these features as well as the knowledge therein. In this way, the newly acquired knowledge could be reinforced through imitation, which is one of the advantages that read-to-write activities could offer.

I also made use of student sample essays when the opportunities arose because students would relate to and learn better by examining examples from their peers (Langan, 2014). For example, observation of the student essays revealed that many scripts had no thesis statement and/or topic sentences, which might be a result of the differences in western and Chinese reasoning process, for example deduction vs. induction. Therefore, based on panel discussion with four teacher raters, I selected an exemplar argumentative essay on the topic of animal research from the previous semester written by one of their peers to illustrate what were expected from the students in terms of the basic rhetorical structure.

I also resorted to other sources for information. For the argumentation task, for example, I searched the internet for Toulmin's (1958) argument model and explain it to the class by analyzing the argument structure of student sample essays that included claim, ground, warrant, backing, and possibly, rebuttal. This helped the students improve their knowledge about argumentation which in turn would enhance the persuasiveness of their argument.

Involving students in peer/self– assessment and self-reflection

Review of related literature suggests that although peer/self-assessment has been widely reported in EFL writing classrooms in China, the general picture is that students' engagement and perceptions of peer assessment tend to be low (Liu & Xu, 2017). Effective delivery of AfL relies on the teacher's careful planning of the activity, thoughtful training of students as assessors, and sustained support throughout the process (Roskams, 1999; Wang, 2014; Zhao, 2014).

At the beginning of the semester, I learned through interviews that the students had little experience of the alternative assessments that constitute the key AfL strategies. Some students said that they did some peer/self-assessment in high school, in which, however, they were only required to give a single score, and without using rubrics. It was therefore doubtful to what extent such form of peer/self-assessment was valid.

I conducted training to the students on peer/self-assessment with the aid of CSE rubrics and relevant literature to ensure validity (e.g., Hu, 2005; Langan, 2014; Luo et al., 2014). To begin with, in view of the possible effects from one-shot writing and timed essay writing in traditional L2 school contexts, I stressed to the students that writing was a process and that revising a rough draft three or four times was often "at the heart of the writing process" (Langan, 2014, p. 30). In the training process I (1) presented and explained the CSE rubrics; (2) discussed the potential advantages and problems of peer/self-assessment as well as possible solutions to the identified problems; (3) explained how peer/self-assessment was to be carried out and what the students were expected to do before, during, and after peer/self-assessment.

In order for the students to have sound understanding of the assessment criteria and develop good knowledge of how to evaluate the quality of an essay and give helpful feedback, I asked them to complete an evaluation form developed based on related literature (Baker et al., 2020; Langan, 2014; Luo et al., 2014). In particular, Lee (2017) suggests that evaluation form (feedback sheet) contain a few open-ended questions as well as a rating rubric. As a result, the form that we used contained four parts that asked students to: 1) give and justify their scores in

the four components of the rubrics; 2) describe the strength of the essay; 3) describe the weakness of the essay, and; 4) give suggestion for improvement. This would help enhance their assessment validity which would in turn help ameliorate assessment and feedback quality.

I also asked the students to write reflection in Chinese upon completion of each task, as reflection makes students aware of their experiences in learning processes (Lee, 2017). The general process cycle of the essay tasks is shown in the table below.

Table 1
Essay Task Procedure

Week	Task	Session	Activity
1	1	1st session	Introduction to the task type (in class) and start of an assignment (complete after class)
		2nd session	Peer/self-assessment (complete in class for task 1 and 2; complete after class for task 3 onwards)
2		1st session	Submission of 1st and 2nd drafts together with written peer/self-assessment comments and reflection
		2st session	Return of teacher feedback and review of task
3	2	Repeat the procedure	Repeat the activity
...

Despite a tight schedule, I set aside four class hours for students to practice peer/self-assessment in their first two essay tasks (comparison & contrast, description) so that they would develop better understanding and knowledge of how to conduct these alternative assessment forms. I hoped that in the follow-up tasks they would be in a better position to perform peer/self-assessment on their own.

Teacher providing descriptive and diagnostic feedback

Recognizing that grammatical errors were not suggested to be corrected extensively according to the literature (e.g., Hu, 2005; Lee, 2007, 2017), in addition to the heavy workload it may incur, I provided descriptive and diagnostic feedback to student scripts more on organization and content, less on grammars and vocabulary. This, on the one hand, would increase my feedback efficiency, on the other, served the needs of the students and mine for attending to the content and organization of essays. These features were captured in the component of Coherence & cohesion in the CSE rubrics, which stressed that, among other things, “information and ideas are connected logically and flow together smoothly”, and also in the component of Task fulfillment. I made comments on the margin next to the individual point I underlined in student scripts and gave my overall evaluation and comments at the end. My efforts proved to be fruitful, especially for the students who became aware of their weakness in organization and content. For instance, in the description task, one student organized her main ideas in the description of one of her favorite restaurants in the order of service, setting, and dishes. After reading her essay carefully, I told her an order that I deemed more logical would be the setting first, then the dishes, and finally the service. I also suggested that she put her feeling toward the restaurant together with the service, as her feeling mostly stemmed from the service of the restaurant in the essay. During the interview, the student said that she found those suggestions very helpful.

Creating a supportive classroom culture

To carry out AfL in the writing classroom, “it is important for teachers to provide a secure and supportive learning atmosphere” (Lee, 2017, p.35) and to help L2 students overcome apprehensions about writing. To learn more about the students’ needs hence give better support to the students, I conducted after-class interviews with some students. We talked about their experience of the class and the peer/self-review, difficulties and doubts in particular. These feedbacks informed my follow-up teaching. For instance, one student said during the interview that they tended to focus on the grammar of the peer’s essay and didn’t know how to evaluate and revise the content and rhetorical aspects of the essay. In the next class, I explained to the students that the assessment of the rhetorical aspects of an essay was related to the component of Coherence & cohesion and Task fulfillment in the rubrics, and students could use the descriptors under these components to evaluate whether if an essay met the criteria, if not, then the writers had to revise based on these criteria. As a supplement to the CSE rubrics and for giving students more support for assessment practice, I introduced the four bases of writing proposed by Langan (2014), that is, unity, support, coherence, and sentence skills.

Upon completion of each task when I gave back the students’ scripts and my feedback, I conducted a task review section in the class with a view to enhancing the students’ awareness of what were expected from them out of the whole writing process cycle. I presented good examples of pre-writing, peer/self-assessment form, revision, and reflection. I used sample essays to illustrate to what extent the higher level requirements in the CSE rubrics were met. For example, I presented the skillful use of idiomatic expression (*put the cart before the horse*) in a sample essay to show to the students that the appropriate use of this type of expression would be conducive to the development of an essay toward a higher CSE level in terms of Lexical resources, a component assessing to what extent a script used fixed expressions including proverbs, idioms, and formulaic expressions.

Lee and Falvey (2014) pointed out, “The implementation of AfL has to be supported by pedagogical practices that are sensitive to students’ interests and needs – for example, choice of writing topics and tasks that are of interest and relevance to students” (p. 229-230). This was borne out in my class. One student said in the interview: “Today the essay topic is ‘The challenge of studying abroad’. I happen to have a plan to study abroad. So I give much attention to the task.” With these knowledge and experience in mind, I thought about ways to design pedagogical activities that were meaningful and interesting, inserted them into the already tight class hour, and allowed of enough time possible given that implementing AfL was time consuming.

Disengaging scores from feedback

“In AfL, if feedback is to produce positive impact on students, scores have to be de-emphasized” (Lee, 2017, p.36). In our writing practice the influence of scores cannot completely be removed from using the CSE rubrics. To alleviate the negative influence on students, we used levels (CSE4-7) instead of scores to describe the quality of the essays, and I, as the teacher researcher, only gave qualitative comments to essays. This, I believed, would help direct the students’ attention to how to make improvement in writing rather than how many points they got.

Benefits for Students

AfL underlines the pivotal role that assessment plays in reinforcing and extending learning, pays special attention to teachers, learners, and the classroom, and attaches great significance to the relationship between teaching, learning, and assessment (Lee, 2007). As students are at

the focal point of AfL activities, the value of the benefits they reap from the CSE-AfL practice is of great significance. I summarize the benefits from the following perspectives.

The CSE rubrics guided writing process

According to the interview, the students thought the CSE rubrics helped their writing in some ways. The most important of all was that they knew where they were across the CSE level continuum, which level they should or wanted to achieve, and what they needed to do in order to realize their goal. These were essential elements in the AfL definition (Assessment Reform Group, 2002).

The rubrics were found to facilitate the entire process of writing: pre-writing, writing, and revising. In the prewriting stage, the rubrics provided a “blueprint” for students to follow. Some students would first read through the descriptors of the required genre under the component of “Task fulfillment” to build a framework for writing, which involved knowledge on how to carry out prewriting techniques I presented in the class (e.g., free-writing, clustering, questioning, etc.). While composing, they bore in mind the primary points in the descriptors of the rubrics component with a view to complying with the requirements specified therein. The students believed that doing so may ensure that they were on track and that they fulfilled the task requirements effectively. In the peer/self-assessment stage, the students compared their performance to the individual descriptor under each component to determine to what extent a script met the requirements stipulated in these descriptors. As they were required to justify the scores they awarded, they used the rubric descriptors to guide their evaluation as well as suggestions for revision.

In the revising stage, they paid much attention to the weak areas identified based on the rubrics. If necessary, they would do peer/self-negotiation by referring to the rubric descriptors for confirmation and/or further clarification. To achieve the expected results of revision they often logged onto the internet or used applications in their mobile phones to find appropriate vocabulary and sentence structure with a view to improving the quality of their essays with respect to the component of Grammar and Lexical resources. In the interview they showed agreement on the diagnostic effects of the CSE rubrics as it helped them develop better understanding about the strength and weakness at each component and find viable ways for improvement. These findings echoed with the results of meta-analysis of the formative uses of scales/rubrics conducted by Panadero and Jonsson (2013) who found that learners, when given rubrics and participated in metacognitive activities such as peer- and self-assessment, were able to internalize the assessment criteria and enhance their subsequent learning outcomes.

Peer/self-assessment assessment

From analyzing the data of the interviews and student assessment forms and reflective notes, I learned that the students benefited from the major AfL strategies to varying degrees. In terms of self-assessment practice, through analyzing their own writing against the rubrics, they gradually developed more objective understanding about their limits and strengths in essay writing and language development. They developed the ability to make sound judgment about the quality of a piece of writing, which cast a contrast to the vague impression that formed in the head after reading an essay in the past. The repeated self-assessment practice during the semester gave them the opportunity to ponder over their writing development, within the same and across the different genres. In particular, through comparison of their own performances across different tasks, they gradually acquired the new knowledge of English writing as a scientific discipline, and of themselves as an autonomous language learner.

Peer assessment also proved to be a valuable AfL strategy. The students said that having their paper read by a reader provided a different perspective to their writing. This would facilitate a new understanding about their advantages and disadvantages in English writing, because, among other things, the readers would point out some issues that may not occur to the students if work alone. These issues, especially when raised honestly, often helped reveal the wrong understanding and insufficient knowledge a student held about language learning, essay writing in particular. Peer assessment effectively facilitated mutual-scaffolding among students who would be a valuable pool of resources to each other. Peer assessment, as a form of a social activity, also encouraged students to become more responsible learners. The students said in the interview that before they presented their scripts to their partners, they had to double check it and tried to eliminate all the grammatical and vocabulary errors they can spot to reduce the burden on and trouble to the readers. Bruce (2001) contended that students reap the most benefit when they were prompted to assess their work in individual activities before submission. Checking work for mistakes and errors is the self-assess of cognitive quality and progress and regarded in many self-regulated learning studies as the foundational construct for planning, monitoring, and reflecting upon cognition (e.g., Bandura, 1997; Pintrich, 1999; Zimmerman & Pons, 1986).

Classroom teaching

In each unit of the textbook, I presented supplemental material from my reading of related literature to help my students better comprehend what were entailed in writing of the task genre addressed in the unit. Much of the information I brought was new to the class. One student said in the interview that before they came to the university, the instructors in the past laid much more emphasis on grammar and vocabulary than on content and rhetoric. The information delivered through classroom instruction, together with the CSE rubrics presented, led them to more knowledge of writing, particularly with respect to topic development and coherence and cohesion. Students also felt that in the past they tended to list a series of events in their writing of narrative essays. Now they would think about what type of “human conflicts” (Langan, 2014, p.204) was involved in the story they were prepared to compose. They also learned more about the structure of argumentation, which they seldom wrote about before.

I also made efforts to connect reading and writing in the class to reinforce the knowledge and features of a certain genre of reading that students may find useful in their writing. For example, in the unit where descriptive text was the subject-matter, I asked the students to underline the words in the text that appealed to their five senses (sight, hearing, taste, smell, and touch) to let them know that words like these were helpful for sketching a verbal picture in descriptive writing. I also encouraged the students to underline and memorize the useful structures and patterns in the text so that they would come in handy when the students came across similar context in essay writing. The students admitted in the interview that those structures and patterns impressed themselves upon their memory harder than if they were not emphasized and practiced in class.

Problems Identified

Notwithstanding the advantages and benefits, problems came up as the classes proceed and many were identified through classroom observation, analyzing the interview data, and by reading students’ assessment forms and reflective notes. Some problems were related to the rubrics. For example, some students felt some of the rubric descriptors used terms that were general and vague, and thus of little help in revising an essay. For example, what was “clear and convincing logical argument”? Hence how to make an argument “clear and convincing”? Without clear guidance for revision, the value of using the rubrics in peer/self-assessment

became questionable. Also some students expressed that the rubrics confined their thoughts for writing to only four components.

Some students distrusted their ability to assess their own or other's writing, hence thought they were unable to judge the validity of peer comments. They tended to make little changes in the second draft after the peer/self-assessment session and would rather wait until they received feedbacks from the teacher. Some students were found to over-rely on the CSE rubrics to complete the assessment report form. They mostly used the rubric descriptors to evaluate their peers' essays and gave feedback which was of little value for revision. Some students even openly expressed their disapproval of using the assessment report, saying that they would rather use the margin of the essay paper to give feedback which they thought would be more straightforward. These negative effects on learner motivation call for attention in AfL (Assessment Reform Group, 2002), and teachers should see to it that AfL practices are not "becoming mechanistic, ritualized and ultimately meaningless and boring" (James, 2011, p.29)".

There were also problems pertaining to the school level constraints discussed in the second section of the paper that went beyond my reach. The tight schedule and limited class hours of college English course posed a serious threat to the efficient delivery of AfL strategies. Some students simply did not want to write, assess, and revise because they were fully occupied with their major study, which might be an important cause for low motivation in most college non-English majors.

Conclusion

Despite the problems identified above, a temporary conclusion can be drawn from the analysis that my students, as well as I as the teacher researcher, reaped benefits from the intervention project. For the problems related to the rubrics and motivation, I will think about ways to improve the wording and structure of the rubrics and my instructional practice, as well as my teaching plan to enhance students' participation. Lee and Falvey (2014) have it right that AfL could not be "effectively promoted without significant changes in teacher and student behavior" (p.223). For the problems concerning school curricula, Lee (2007) offers a suggestion that teachers secure institutional support. Maybe in my next attempt at CSE-AfL based writing instruction, I would not ask students to do peer/self-assessment and reflection for each essay task, as suggested by some scholars (Lee 2017; Nielsen 2021), and in view of the students' tight schedule squeezed by their major subjects, the school curriculum of limited class hour versus the requirement to cover a large amount of information and tasks in the textbook, and the teacher's heavy workload to cater to the needs of the large class size.

Taking into consideration of these constraints together with the suggestions the students gave in the interview, I also plan to use automated feedback to help handle part of the essay assessment work, with respect to grammar and vocabulary in particular. As discovered by Huang and Zhang's (2014), automated feedback is most acted upon with respect to mechanic issues compared to teacher feedback and peer feedback (Liu & Xu, 2017).

Finally, I concur with Smith (2011) who says that a prerequisite for AfL to be successfully implemented in the classroom is the teachers' assessment practice which relies on their assessment literacy. In this regard, however, I, and many other Chinese EFL teachers, suffer a serious lack of solid foundation as well as sound knowledge of effective delivery. Confronted with the various constraints while trying to implement AfL in the classroom, I found it difficult to apply the theories I learned through my educational experience into practice. Obviously,

more efforts are needed to further our understanding and implementation of AfL mandates to progress from letter to spirit along the continuum of our professional learning (DeLuca et al., 2019; Marshall & Drummond, 2006).

References

- Azis, A. (2015). Conceptions and practices of assessment: A case of teachers representing improvement conception. *TEFLIN Journal*, 26(2), 129–154.
- Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. Cambridge University Press.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.
- Baker, B., Homayounzadeh, M., & Arias, A. (2020). Development of a test taker-oriented rubrics: Exploring its usefulness for test preparation and writing development. *Journal of Second Language Writing*, 50, 100771.
- Berry, R. (2011). Assessment for learning reform in mainland China, Hong Kong and Taiwan. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 49–61). Springer.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7–74.
- Black, P., & Wiliam, D. (2003). In praise of educational research: Formative assessment. *British Educational Research Journal*, 29(5), 623–637.
- Black, P., & Wiliam, D. (2006). Assessment for learning in the classroom. In J. Gardner (Ed.), *Assessment and learning* (pp. 9–25). SAGE.
- Black, P., Harrison, C., Lee, C., Marshall, E., & Wiliam, D. (2004). September. Working Inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan*, 9-21.
- Brown, G. T. L., Kennedy, K. J., Fok, P., Chan, J. K. S., & Yu, W. M. (2009). Assessment for student improvement: Understanding Hong Kong teachers' conceptions and practices of assessment. *Assessment in Education: Principles, Policy & Practice*, 16(3), 347–363.
- Bruce, L. B. (2001). Student self-assessment: Making standards come alive. *Classroom Leadership*, 5(1), 1–6.
- Bryant, D. A., & Carless, D. R. (2010). Peer assessment in a test-dominated setting: Empowering, boring or facilitating examination preparation? *Educational Research for Policy and Practice*, 9, 3–15.
- Carson, J. G., & Nelson, G. L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1–19.
- Carson, J. G., & Nelson, G. L. (2006). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1–19.
- Cheng, W., & Warren, M. (2005). Peer assessment of language proficiency. *Language Testing*, 22(1), 93–121.
- Clarke, S. (2005). *Formative assessment in action: Weaving the elements together*. Hodder Murray.
- DeLuca, C., Chapman-Chin, A., & Klinger, D.A. (2019). Toward a Teacher Professional Learning Continuum in Assessment for Learning. *Educational Assessment*, 24(4), 267–285. <https://doi.org/10.1080/10627197.2019.1670056>
- Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning* (2nd ed.). Corwin.
- Gardner, J. (2006). *Assessment and learning*. Sage.
- Hu, G. (2005). Using peer review with Chinese ESL student writers. *Language Teaching Research*, 9(3), 321–342.
- Huang, J., & Zhang, W. (2014). The impact of the integrated feedback on students' writing revisions. *Foreign Languages in China*, 11(1), 51–56.

- Hui, S. K. F., Brown, G. T., & Chan, S. W. M. (2017). Assessment for learning and for accountability in classrooms: The experience of four Hong Kong primary school curriculum leaders. *Asia Pacific Education Review, 18*(1), 41–51.
- Hyland, K. (2003). *Second language writing*. Cambridge University Press.
- James, M. (2011). Assessment for learning: Research and policy in the (Dis)United Kingdom. In R. Berry & B. Adamson (Eds.), *Assessment reform in education: Policy and practice* (pp. 15–32). Springer.
- James, M. & Pedder, D. (2006). Professional Learning as a Condition for Assessment for Learning. In J. Gardner (Ed.), *Assessment and learning* (pp.9–25). Sage.
- Koh, K., Lim, L., Tan, C., & Habib, M. (2015). Building teachers' capacity in formative assessment: The Singapore example. *The New Educational Review, 40*, 211–221.
- Laflen, A., & Smith, M. (2017). Responding to student writing online: Tracking student interactions with instructor feedback in a Learning Management System. *Assessing Writing, 31*, 39–52.
- Langan, J. (2014). *College writing skills with readings* (9th ed). Foreign Language Teaching and Research Press.
- Lee, I. (2017). *Classroom Writing Assessment and Feedback in L2 School Contexts*. Springer Nature Singapore Pte Ltd.
- Lee, I., & Coniam, D. (2013). Introducing assessment for learning for EFL writing in an assessment of learning examination-driven system in Hong Kong. *Journal of Second Language Writing, 22*, 34–50.
- Lee, I., & Falvey, P. (2014). Perspectives on assessment for learning in Hong Kong writing classrooms. In D. Coniam (Ed.), *English language education and assessment: Recent developments in Hong Kong and Chinese Mainland* (pp.221–236). Springer.
- Leong, W.S., Ismail, H., Costa, J.S., & Tan, H.B. (2018). Assessment for learning research in East Asian countries. *Studies in Educational Evaluation, 59*, 270–277.
- Lin, T. J., Lee, M. H., & Tsai, C. C. (2013). The commonalities and dissonances between high-school students' and their science teachers' conceptions of science learning and conceptions of science assessment: A Taiwanese sample study. *International Journal of Science Education, 36*(3), 382–405.
- Liu, J. (2015). Standard-based foreign language assessment. *Foreign Language Teaching and Research, 47*(3), 417-425. (In Chinese)
- Liu, J. (2017). China's Standards of English and Its Applications in English Learning. *Foreign Languages in China, 4*, 4 – 11.
- Liu, J., & Xu, Y. (2017). Assessment for Learning in English Language Classrooms in China: Contexts, Problems, and Solutions. In H. Reinders et al. (Eds.), *Innovation in Language Learning and Teaching, New Language Learning and Teaching Environments* (pp.17-37). DOI 10.1057/978-1-137-60092-9_2
- Luo, S., Huang, J. & Ma, X. (2014). *Enhancing Learning through Formative Assessment in L2 Classroom*. FLTRP Basic Education Publishing.
- Nielsen, K. (2021). Peer and self-assessment practices for writing across the curriculum: learner-differentiated effects on writing achievement. *Educational Review, 73*(6), 753-774, DOI: 10.1080/00131911.2019.1695104
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research Papers in Education, 21*(2), 133–149.
- National Education Examinations Authority (NEEA). (2018). China's Standards of English Language Ability. Beijing, China: Higher Education Press & Shanghai Foreign Language Education Press. Retrieved from <http://cse.neea.edu.cn/html1/report/18112/9627-1.html>
- Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129–144.

- Pintrich, P. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, 31(6), 459–470.
- Roskams, T. (1999). Chinese EFL students' attitudes to peer feedback and peer assessment in an extended pair work setting. *RELC Journal*, 30(1), 79–123.
- Sebba, J. (2006). Policy and practice in assessment for learning: The experience of selected OECD countries. In J. Gardner (Ed.), *Assessment and Learning* (pp.185–196). Sage.
- Sebba, J., & Maxwell, G. (2005). Queensland, Australia: An outcomes-based curriculum. In J. Looney (Ed.), *Formative assessment: Improving learning in secondary classrooms* (pp.129–148). Organization for Economic Cooperation and Development.
- Smith, K. (2011). Professional development of teachers—A prerequisite for AfL to be successfully implemented in the classroom. *Studies in Educational Evaluation*, 37(1), 55–61.
- Tan, C. (2016). Tensions and challenges in China's education policy borrowing. *Educational Research*, 58(2), 195–206.
- Toulmin, S. (1958). *The uses of argument*. Cambridge University Press.
- Tsui, A., and M. Ng. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, 9(2), 147–170.
- Wang, S. (2018). The application of China's Standards of English language Ability to college English teaching. *Foreign Language Education*, 39(4), 1-4.
- Wang, S., & Wang, H. (2011). On the state of college English teaching in China and its future development. *Foreign Languages in China*, 5 (4), 11-17.
- Wang, W. (2014). Students' perceptions of rubrics-referenced peer feedback on EFL writing: A longitudinal inquiry. *Assessing Writing*, 19, 80–96.
- Wang, W. (2017). Using rubrics in student self-assessment: Student perceptions in the English as a foreign language writing context. *Assessment & Evaluation in Higher Education*, 42(8), 1280-1292, DOI: 10.1080/02602938.2016.1261993
- Wiliam, D. (2017). Assessment for learning: Meeting the challenge of implementation. *Assessment in Education: Principles, Policy & Practice*. 1-4. <http://dx.doi.org/10.1080/0969594X.2017.1401526>
- Willis, J. (2010). Assessment for learning as a participative pedagogy. *Assessment Matters*, 2,65–84.
- Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing*, 15(3), 179–200.
- Yin, X., & Buck, G. A. (2015). There is another choice: An exploration of integrating formative assessment in a Chinese high school chemistry classroom through collaborative action research. *Cultural Studies of Science Education*, 10(3), 719–752.
- Yu, W. M. (2015). Teacher leaders' perceptions and practice of student assessment reform in Hong Kong: A case study. *Planning and Changing*, 46(1/2), 175–192.
- Zhao, H. (2014). Investigating teacher-supported peer assessment for EFL writing. *ELT Journal*, 68(2), 155–168.
- Zimmerman, B. J., & Pons, M. M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23(4), 614–628.

Jiuliang Li obtained his PhD from the National Key Research Center for Linguistics and Applied Linguistics at Guangdong University of Foreign Studies. He is now an associate professor at the School of Science and Humanities, Beijing Institute of Fashion Technology. His research interests are second language reading and writing, formative and summative assessment.