Exploring Assessment for Learning Practices in the EMI Classroom in the Context of Taiwanese Higher Education

NAIHSIN LI [a]

JESSICA WU [b]

[a] *The Language Training & Testing Center, Taiwan*
Email: naihsinli@gmail.com

[b] *The Language Training & Testing Center, Taiwan*
Email: jw@lttc.ntu.edu.tw

Abstract

The EMI Classroom Assessment Practices questionnaire was developed and administered to 40 EMI (English as a Medium of Instruction) university teachers in Taiwan with the aim of meeting EMI teachers' needs to conduct learning-oriented classroom assessment. The questionnaire surveyed the teachers in terms of their assessment practices in EMI and their self-perceived skills in these practices. The effects of the medium of instruction (English vs. Chinese) and teacher- and course-related variables related to assessment practices were also examined. The results revealed a high correlation between techniques that were less commonly practiced and those in which the teachers felt less skilled, signaling the possibility that assessment practices in the EMI classroom were dominated by the teachers' familiarity with techniques rather than by instructional objectives. In addition, the teachers perceived themselves to be less skilled in certain learning-oriented assessment practices, such as self-assessment and designing test items that assess higher-level cognitive abilities. Compared with courses taught in Chinese, EMI courses showed a tendency to involve fewer classroom interactions and to assess students' higher-order abilities less often, elements which are critical to the emergence of learning. Although the EMI teachers were aware of students' difficulties with English, these difficulties were not commonly taken into consideration, and accommodations were seldom made. Finally, the teachers' assessment practices were found to be mediated by variables such as discipline, course size, and the teachers' focus of assessment. The findings of this investigation have implications for the development and implementation of useful training programs in support of EMI teachers' professional development in assessment.

Keywords: English as a medium of instruction (EMI), classroom assessment, assessment for learning

Introduction

Assessment is an essential part of learning and teaching. The results can be used to evaluate learning outcomes and teaching effectiveness. More importantly, they can further benefit subsequent learning and teaching by informing teachers and learners of "where the learners are in their learning, where they need to go, and how best to get there" (Assessment Reform Group, 2002). This notion of *assessment for learning* has great potential to address pedagogical issues observed in classrooms where English as an L2 is adopted as the medium of instruction (EMI). This practice of EMI is gaining popularity in tertiary education in non-English-speaking countries as a response to globalization and the position of English as a lingua franca in various professional domains (Macaro, Curle, Pun, An & Dearden, 2018). Therefore, this study examined teachers' actual assessment practices in the EMI classroom and their self-perceived skill levels in these techniques in the context of Taiwanese tertiary education, intending to identify the teachers' needs in implementing learning-oriented assessment practices in the EMI classroom.

EMI in Taiwan

Despite variation in terminology and its form of application, EMI generally refers to the practice of using English as a vehicular language to transmit academic knowledge (Macaro *et al.*, 2018). The use of a lingua franca (i.e., English) as the medium of instruction is believed to create opportunities for universities to recruit international students and faculty, which can catalyze the internationalization of these universities and strengthen their competitiveness locally and globally. There is also an expectation that EMI can enhance domestic students' English abilities and increase their global mobility. Attracted by the possible benefits of EMI, tertiary educational institutions in Taiwan have begun to participate in this trend. In 2004, the Ministry of Education (MOE) began to officially encourage universities and colleges to conduct classes in English to promote the internationalization of higher education (MOE Taiwan, 2009). Later, in 2006, the percentage of EMI courses offered and its annual growth became one of the indices for university evaluation in the *Development Plan for World Class Universities and Research Centers for Excellence*, later retitled as *The Top University Development Plan* (Chen, 2011). As a result, the number of EMI courses in universities alone, not including those in vocational and technology colleges or institutions, showed a growth rate of 18% between 2009 and 2014 (Chung & Lo, 2016). Furthermore, the MOE established a policy of promoting and refining English-taught degree programs in universities and colleges beginning in 2011 (MOE Taiwan, 2017).

While EMI policies are often implemented with the good intentions of bringing both economic benefits for universities and educational benefits for students, in practice, they have actually posed many pedagogical challenges to content teachers (Doiz, Lasagabaster, & Sierra, 2013; Dearden, 2015; Galloway, Kriukow, & Numajiri, 2017). Among these, testing and assessment in an EMI setting is an under-researched area that must be further explored.

Challenges in EMI Assessment

A large-scale investigation on the implementation of EMI in 55 countries around the world identified testing and assessment as a "problematic" area (Dearden, 2015). One major challenge is that students are assessed through a language they are still learning. Therefore, it is very likely that their performance will be affected by their English proficiency and the assessment will fail to truly reflect their acquisition of academic knowledge. Indeed, previous studies have found an L2 effect on students' academic performance. For example, Robinson (2010) showed that young Spanish learners of English

performed significantly better on a math assessment when it was conducted in their L1. In addition, van der Walt and Kidd (2013) found that students who spoke Afrikaans as their home language consistently scored lower than English-speaking students on an English reading comprehension test, even when L1 support was provided. A solution to this challenge demands EMI teachers be clear about their focus of assessment: whether it is on content knowledge, on language, or on both (Coyle, Hood & Marsh, 2010).

Nonetheless, even greater challenges lie in choosing an appropriate assessment method for the instructional objective and implementing it in accordance with the assessment purpose. For example, when the purpose of an assessment is for students to demonstrate their academic knowledge, efforts should be devoted to ensuring that students are not penalized for their English abilities. On the other hand, if improvement in academic English abilities is also deemed important, or at least relevant, to the specific content course, EMI teachers should take the necessary measures to promote students' language learning (Rogier, 2012; Tai, 2015).

Assessment for Learning

The aforementioned challenges can be more effectively addressed if teachers are equipped with competent assessment knowledge and skills and, more vitally, are aware of the close relationship between assessment and learning. In recent years, the notion of assessment has been reconceptualized as an integral part of teaching and learning. To be specific, assessment is not only a measure taken to present the end products of learning (i.e., assessment *of* learning), but is also conceived of as an interactional classroom practice that involves repeatedly gauging learning outcomes while continuously feeding forward to advance learning (i.e., assessment *for* learning; see Black & William, 2009; Davison & Leung, 2009; Jones & Saville, 2016). This concept is properly demonstrated in the Learning Oriented Assessment model proposed by Jones and Saville (2016). This model was constructed on the basis of an extensive literature review of learning and assessment, and greatly overlaps with Black and William's (2009) widely cited framework of formative assessment in the basic conceptualization of learning processes and the particular elements that are essential to promote learning through assessment. However, with its specific focus on language learning, the model described by Jones and Saville (2016) may be more relevant to classroom assessment practices in the EMI classroom, where language learning is often an intended goal, or usually a hidden agenda.

Jones and Saville (2016) propose that the basic elements of classroom interaction are *tasks, goals, scaffolding, feedback* and *emergence.* Teachers should involve learners in a learning *task* whose *goals* are clear to both parties. In the process of approaching the learning goals, the teacher should provide necessary support to *scaffold* the learners' task performance, which includes accounting for their prior knowledge, sharing with them explicit criteria for success, and providing learning support as required. After the task is performed, the teacher should also make evaluations and provide constructive *feedback* that can bridge the gap between the learners' actual performance and the desired outcome. In addition to examining the learners' acquisition of what has been transmitted, the teacher should also provide chances for the learners to "transfer knowledge and use it in new situations" in order to facilitate deep learning and the *emergence* of higher-order skills. Equally important to the aforementioned practices are the teacher's efforts to foster the learners' awareness and engagement with the learning goals. The teacher can do this in various ways; involving students as the assessors of their own performance (i.e., self-assessment) or that of their peers (i.e., peer-assessment) is particularly effective in building learners' autonomy in guiding their own learning through self-regulation and self-evaluation (i.e., *assessment as learning*; Dann, 2014). Although learners take the central role in *assessment for learning*, the teacher plays an important role as a facilitator who helps to make learning more achievable by conducting appropriate assessment practices.

Assessment for Learning in EMI

The practice of assessment for learning is particularly helpful in addressing the learning challenges faced by students in the EMI classroom. For example, a clear mutual understanding of the learning objectives and expectations of students' English performance can alleviate the anxiety felt by students who are preparing and performing an assessment task in English. Moreover, such practices help the teacher decide when and how scaffolding, such as strategies to accommodate students' English difficulties, should be provided. Previous research on EMI assessment has presented several accommodation strategies, such as permitting students to use their L1 to perform the assessment task (Li, 2017), devising bilingual assessments in which L1 resources are incorporated (van der Walt & Kidd, 2013) or lowering the linguistic complexity of the assessment by using simpler English or incorporating visual aids (Kao & Tsou, 2017). If English language improvement is one learning objective, then the teacher should also deliberately maneuver the assessment procedures to achieve this desired outcome. Strategies may include, but not be limited to, making clear the language learning goals to the students, creating opportunities for meaningful use of the language, and providing corrective feedback to English errors.

Most EMI assessment studies have focused on evaluating the effectiveness of EMI programs (Lei & Hu, 2014; Yang, 2015; Li, 2017), and there is a dearth of research discussing classroom assessment in an EMI setting. Kao and Tsou (2017) contributed a book chapter on this issue, investigating the EMI teaching practices of 29 teachers and their perceptions on assessment in the EMI classroom. Through a survey and interviews, their study investigated the grading criteria adopted by these teachers and suggested that their adoption of assessment tools was affected by their beliefs regarding the role of English in the EMI classroom. Despite the informative findings, developing valid grading criteria is only one aspect of assessment practice. To create learning-oriented classroom assessments, teachers should also possess skills related to aligning assessment tools with teaching goals, scoring and interpreting the results to make further judgments and teaching plans, and communicating assessment results with students and providing feedback to lead to further learning (American Federation of Teachers [AFT], National Council on Measurement in Education [NCME], & National Education Association [NEA], 1990). A comprehensive investigation of teachers' skills and practices in all aspects of assessment competence is necessary for us to identify teachers' assessment needs in creating a learning-oriented EMI classroom.

To achieve this purpose, the EMI Classroom Assessment Practices (ECAP) questionnaire was developed and administered to in-service EMI teachers in Taiwanese universities. The questionnaire investigated the teachers' actual use of and self-perceived skillfulness in these assessment techniques, and also explored the roles of students' L1 (Chinese) and L2 (English) in EMI classroom assessment.

The literature on EMI studies shows that the practice of EMI can be highly contextualized (Macaro *et al.*, 2018). Therefore, we also explored how teacher- and course-related variables might differentiate EMI teachers' assessment practices or competencies. Moreover, the choice of the medium of instruction (i.e., students' L1 or L2) has been found to affect assessment practices. For example, Hu and Li (2017) found that teachers tended to ask fewer higher-order questions in EMI than in CMI classes. Hence, we also investigated the effect of the medium of instruction on assessment practices among the teachers who had conducted a similar course in Chinese (CMI).

The research questions addressed in this study are as follows:

(1) What are the actual assessment practices in the EMI classroom and teachers' self-perceived skills in these practices? Is there any gap between these two?
(2) Does the medium of instruction (EMI or CMI) have an effect on assessment practices?
(3) What are the roles of teachers' and students' L1 (Chinese) and L2 (English) in the EMI classroom?
(4) Are there any discernible effects of teacher- and course-related variables?

Method

Participants

The ECAP questionnaire was distributed online through existing networks among EMI teachers in Taiwan. Forty university teachers completed the questionnaire, and their responses were collected and analyzed. All the teachers fulfilled the qualification of teaching at least one EMI course at the tertiary level. Their teaching experience ranged from 1 year to 30 years, with a median of 11 years. However, most of them were novice EMI teachers with less than 5 years of experience in offering EMI courses ($n = 25$).

The teachers were from a variety of departments (see Appendix A). For ease of statistical analysis, they were subgrouped into three major disciplines: Humanities ($n = 14$), Business and Management ($n = 16$), and Science and Engineering ($n = 10$), in line with the division standards of the Ministry of Science and Technology (MOST) in Taiwan. Most of the teachers had received their highest degree in an English-speaking country ($n = 25$) or had lived in an English-speaking country ($n = 33$).

The majority of the courses given by these teachers were lectures ($n = 38$) and were offered to students at the undergraduate level ($n = 30$). The size of the classes ranged from 5 to 200 students. Among the EMI classes, 16 had less than 25 students, 15 had 26–50 students, and 7 had more than 51 students. Two teachers did not report the size of their courses. Most of these classes were composed exclusively ($n = 7$) or mostly ($n = 23$) of Taiwanese students. A summary of the teacher- and course-related information is presented in Appendix B.

ECAP Questionnaire

The ECAP questionnaire (Appendix C) consists of three parts. Part I was designed to collect demographic information on the EMI teachers and information regarding their EMI courses. Part II contained 42 questions intended to collect information on the teachers' practices and skills in assessment. Among these questions, items 1–27 were developed from the Standards for Teacher Development in the Area of Assessment (AFT, NCME, & NEA, 1990), and items 28–42 inquired about the forms of assessment (items 28–40) and students' involvement (items 41 and 42). The teachers were asked to indicate on two Likert scales from 1 to 5 respectively (1) their frequency of applying a specific assessment practice (1 = *not at all used*; 5 = *used very often*) and (2) their skillfulness in the specific assessment technique (1 = *not at all skilled;* 5 = *very skilled*). In addition, if the EMI teacher had taught the same or a similar course in Chinese (i.e., CMI), he or she was asked to indicate whether there was a difference in assessment practices between these courses (1 = *no difference*, 2 = *only in EMI*, 3 = *more often in EMI*, 4 = *less often in EMI*, 5 = *only in CMI*). Teachers who responded with a 2 or 3 on this scale were required to further indicate their main concerns regarding these practices (i.e., *English language, course size/type,* or *both language and course size/type*).

Part III presented questions concerning the role of English in the EMI classroom. Teachers were asked whether the focus of an assessment task was content, English language, or both. They were also asked about how frequently they applied accommodation strategies in the face of students' English

difficulties in an assessment task (i.e., the use of simple English, scaffolding, and permission for students to use L1) and their frequency of providing corrective feedback on students' English errors.

Results

Assessment Practices and Teachers' Self-Perceived Skillfulness in These Practices

First, we investigated the teachers' assessment practices in the EMI classroom by examining their responses to the first Likert scale, regarding frequency (hereafter referred to as the "scale of frequency"). The response rate was 100%. To provide a broad overview, the average ratings of the assessment practices were first categorized into the seven standards (items 1–27), assessment methods (items 28–40), and communication of grading criteria (items 41 and 42). The results, summarized in Table 1, revealed that the average ratings ranged from 3.50 to 4.30. One-sample t tests showed that the teachers applied these assessment techniques significantly more often than occasionally (i.e., a rating of 3) in their EMI classrooms (one-tailed, $\alpha = .05$, $ps < .01$).

However, an item-based inspection revealed that some techniques were less often practiced in the EMI classroom. These included selecting textbook-provided test items for assessment (item 2), adopting a grading model (items 12 and 13), incorporating students' ability and improvement in grade calculation (items 16 and 18), communicating assessment results to other educators (item 26), and assessing students through fill-in-the-blank or short-answer questions, portfolios, or self-assessment (items 31, 38, and 39).

In addition to examining the average ratings across all participants in each assessment practice, we also examined the distribution of the participants' responses. In particular, we investigated which assessment techniques were reported to be less frequently used (i.e., responses 1, 2, and 3 on the scale of frequency) by more than 35% of the EMI teachers. In addition to those found in the statistical tests, other less-frequently performed assessment practices included providing written feedback to students (item 24) and a number of additional assessment methods such as assessing students with tests or quizzes; observations; multiple-choice questions; group or pair class participation; or individual, group, or pair hands-on activities (items 28, 29, 30, 35, 36, 37).

Table 1 Average Rating Scores on the Scales of Frequency and Skillfulness

Categories of assessment practices	Frequency	Skillfulness
Standard 1: Choosing assessment methods appropriate for instructional decisions (items 1 & 2)	3.58	3.81
Standard 2: Developing assessment methods appropriate for instructional decisions (items 3–5)	4.09	4.04
Standard 3: Administering, scoring, and interpreting the results of both externally produced and teacher-produced assessment methods (items 6–7)	3.82	3.85
Standard 4: Using assessment results when making decisions about individual students, planning teaching, developing curriculum, and improving schools (items 8–11)	3.93	3.96
Standard 5: Developing valid pupil grading procedures that involve pupil assessment (peer assessment) (items 12–22)	3.67	3.88

Standard 6: Communicating assessment results to students, parents, other lay audiences, and other educators (items 23–26)	3.45	3.70
Standard 7: Recognizing unethical, illegal, and otherwise inappropriate assessment methods and uses of assessment information (item 27)	4.30	4.10
Methods of assessment (items 28–40)	3.50	3.73
Communication with students about grading criteria (items 41 & 42)	4.25	4.31

Table 2 *Parallel Analysis of Less-Frequent and Less-Skillful Assessment Practices*

Categories of assessment practices	Item Numbers
Standard 1. Choosing assessment methods appropriate for instructional decisions	2
Standard 5. Developing valid pupil grading procedures which use pupil assessment (peer assessment)	12, 13, 16, 18
Standard 6. Communicating assessment results to students, parents, other lay audiences, and other educators	24, 26,
Methods of assessment	28[1], 29, 30, 31, 33[2], 35, 36, 37, 38, 39

Note. [1] Assessment techniques that were less frequently practiced, but that teachers perceived themselves to be skilled in; [2] Assessment techniques that teachers perceived themselves to be less skilled in, but that were frequently practiced in class

Effects of teacher- and course-related variables. Further analyses were conducted to explore whether teacher- and course-related variables distinguished the teachers' assessment practices and their self-perceived skills in these assessment techniques. The variables examined were discipline, teaching experience, EMI experience, experience studying or living in an English-speaking country, course size, course level, and the composition of students in the course. Subgroup performances were compared with non-parametric statistical tests. Only statistically significant results relevant to our research questions are reported below.

Notable differences were found among the assessment practices in different disciplines. For example, the Science and Engineering teachers were found to perform certain assessment practices significantly less often than the Humanities teachers or/and the Business and Management teachers. Of these, the assessment techniques that were particularly relevant to learning included incorporating student engagement (χ^2 (2) = 6.91, $p < .05$) into grade calculation; communicating classroom assessment criteria (χ^2 (2) = 12.26, $p < .005$) or results to students (χ^2 (2) = 7.07, $p < .05$); and assessing students through observation (χ^2 (2) = 6.96, $p < .05$) and individual class participation (χ^2 (2) = 7.66, $p < .05$). Compared with teachers in the other disciplines, the Science and Engineering teachers also used assessment results less often when evaluating student improvement (χ^2 (2) = 6.73, $p < .05$). In terms of the disciplinary difference in the skills of assessment, the Science and Engineering teachers were significantly less skilled than the Humanities teachers and the Business and Management teachers in communicating performance assessment criteria to students in advance (χ^2 (2) = 6.42, $p < .05$).

In addition, the size of a course was found to exert a remarkable influence on EMI teachers' assessment practices. In a comparison across EMI classes of different sizes, teachers of larger classes ($n > 50$) more often assessed students with textbook-provided test items (χ^2 (2) = 7.05, $p < .05$) or multiple-choice questions (χ^2 (2) = 6.42, $p < .05$) and less often provided written feedback to students (χ^2 (2) = 8.53, $p < .05$) or assessed students with essay questions (χ^2 (2) = 8.01, $p < .05$). There was also a tendency for the teachers to assess students' higher-level cognitive abilities less often in larger EMI

classes (χ^2 (2) = 5.66, *p* = .059).

Effects of the Medium of Instruction on Assessment Practices: EMI vs. CMI

Among the 40 participants, 18 reported conducting the same or similar course in both CMI and EMI. While 11 of them (61.11%) indicated no difference in their assessment practices in EMI and CMI, seven participants acknowledged that the medium of instruction affected their assessment decisions to some degree. Subject-based inspections revealed considerable variation among these teachers' assessment practices in classes with different instruction media. For example, although two of the seven teachers practiced all the assessment techniques consistently in CMI classes only, one teacher indicated that she practiced the same assessment techniques more often or only in EMI classes. Another three teachers generally performed similar assessment practices in the EMI and CMI classes; however, EMI, or its accompanying feature of a smaller class size, still affected their assessment practices in some aspects. The remaining teacher demonstrated much greater variation in assessment practices compared with the other teachers, due to the medium of instruction.

Despite the variation across teachers, we also noted some commonalities among their assessment decisions. Specifically, some assessment practices were more likely to be applied in the EMI classroom, such as selecting textbook-provided test items for classroom assessment (item 2), using assessment results to plan teaching (item 8), incorporating students' engagement into grade calculation (item 21), and assessing students through observation (item 29) or presentation (item 40). By comparison, the assessment practices that were conducted less often in the EMI classroom or only in the CMI classroom included choosing or developing appropriate assessment methods for instructional decisions (items 1, 3, 4, and 5), using assessment results when developing curriculum (item 9), and including more varied grading criteria (e.g., items 12, 13, 14, 16, 19, and 21). The teachers were also more likely to provide feedback to students and communicate classroom results to students or other educators in the CMI classroom (items 23–26). In addition, teachers in the CMI classroom tended to assess students through a greater variety of methods. In particular, five of the seven teachers acknowledged that they assessed students' higher-level cognitive abilities less often in the EMI classroom.

Roles of English and Students' L1 in Assessment in the EMI Classroom

In the EMI classroom, English appeared to be not only the main medium of instruction but also the default language of assessment. Most of the teachers (90%) said that they often or very often used English to present an assessment task to students. However, when they did so, they often used simple English. The reasons provided by some of the teachers suggested that this was a strategy meant to accommodate Taiwanese students' difficulties with English, especially when English was the vehicular language as required by the course or due to the enrollment of international students in the class.

While the teachers were aware of students' difficulties with English, only slightly more than half of them very often or often took into account students' English ability when evaluating their assessment performance (*n* = 21). The other half occasionally or never did, mostly because they claimed that English was not their focus of instruction (*n* = 18), although three of them also indicated that the English language was one of their assessment focuses.

In addition, it was not common for the teachers to employ accommodation strategies, such as scaffolding or the use of L1, to compensate for students' English difficulties when performing the assessment task. For example, only less than half of the teachers (*n* = 19) indicated that they often or very often provided scaffolding, such as providing Chinese translations of key terms in a test paper. Similarly, only one-third (*n* = 13) of the teachers said they often or very often allowed students to use

their L1 in the assessment task if students had difficulties performing the task in English. The main reason why the other teachers occasionally or never tried to accommodate students' English limitations in assessment was that English was not their focus of instruction (n = 19–20). The other reasons included insufficient time, the presence of foreign students in the class, and the presupposition that students should have the ability to complete the task in English.

Most of the EMI teachers did not find it necessary to provide corrective feedback on students' English errors in assessment performance. For example, more than two-thirds (n = 27) of the teachers reported that they occasionally or never did so because language was not the focus of instruction (n = 24). However, six among them also indicated that English language was one of their focuses of assessment. Other reasons were also specified. One teacher believed that students could improve their English by using it; therefore, correcting their errors might not be necessary. Another teacher indicated that since language accuracy is secondary in an EMI course, an overemphasis on students' language errors may discourage them. Two teachers said they had insufficient time to correct students' English errors.

Effects of teacher- and course-related variables. Comparing teachers who had the experience of living in an English-speaking country with those without such experience, the former more often took students' English language ability into consideration when evaluating students' performance (χ^2 (2) = 6.54, p < .05). They also more often corrected students' English errors in assessment performance (χ^2 (2) = 7.31, p < .05). In addition, it was found that teachers in the discipline of Humanities significantly more often provided scaffolding to students than teachers of the other two disciplines (χ^2 (2) = 6.18, p < .05). Teachers with longer teaching experience or longer EMI teaching experience significantly more often allowed students to use Chinese in assessment tasks than less experienced (EMI) teachers (teaching experience: χ^2 (2) = 7.67, p < .05; EMI teaching experience: U = 110.00, p < .05).

Teachers' beliefs about their focus of assessment in the EMI classroom also affected their language policy. Although 25 teachers indicated that the focus of assessment was solely on content, approximately one-third of them regarded language as an additional focus of assessment. The comparisons between these two groups showed that EMI teachers whose focus of assessment was both content and language more often took students' English ability into account when evaluating students' assessment performance (U = 87.5, p < .005), provided scaffolding (U = 117.00, p < .05), and corrected students' English errors when evaluating their performance in English (U = 101.50, p < .05). Nonetheless, there were no significant group difference in their use of English, use of simple English, and consent to students' use of L1 in the assessment task (all at p > .05).

Discussion

With the increasing trend of EMI in tertiary education, there is an urgent need to address concerns related to assessment in the EMI classroom. We propose that the practice of *assessment for learning* has the potential to resolve students' learning difficulties in EMI, for it promotes intensive interactions between teachers and students as well as among students, which allows both teachers and the learners themselves to closely monitor progress and resolve learning problems. Moreover, this learning-centered approach encourages greater student engagement with learning goals, which can maximize students' learning achievements. However, the effective implementation of such an approach requires teachers to be well-equipped with assessment knowledge. Therefore, this study was conducted to examine teachers' assessment practices in the EMI classroom and their self-perceived skill in these assessment techniques, with a focus on their needs in conducting classroom assessment to promote learning.

The first principle of a good assessment practice requires the teacher to set clear learning objectives

and align them with an appropriate assessment task. Most teachers indicated that they frequently carried out and were skilled in such practices in the EMI classroom. Nevertheless, cross-referencing with other findings in this study suggested that this might not be the case. First, we found a high correlation between the techniques that were less practiced and those in which the teachers felt less skilled. One possible interpretation of this finding is that through the processes of applying assessment practices that were appropriate for their instructional purposes, the teachers felt that they had become more skillful in these specific assessment techniques. However, it is more likely that the assessment practices employed in the EMI classroom were determined by the teachers' prior familiarity with those types of assessment techniques. That is, some assessment techniques were more frequently practiced solely because the teachers felt more comfortable in using them, rather than because these tasks were more appropriate for the instructional objectives. In fact, the general profile of the EMI teachers' assessment skillfulness showed that they were not well-acquainted with all the possible assessment techniques. For example, subject-based inspection revealed that 14 out of the 40 participants indicated that they were not skilled or only somewhat skilled in half or more of the assessment methods. Three of them even acknowledged that they were not at all skilled in any of the assessment methods. With limited assessment techniques at their disposal, it would be very difficult for them to choose or develop appropriate assessment methods for their instructional purposes, even though nearly half of the 14 participants still claimed that they were skilled in doing so.

In terms of goal-setting, we also noticed a misalignment in terms of the teachers' expectations for the role of English in the EMI class. According to our findings, while most of the teachers reported focusing assessment on content knowledge, nearly one-third stated that the English language was also included as a focus of assessment. However, 40% of these respondents also claimed that English was not the focus of instruction. This misalignment between instructional goals and assessment focus puts students with poorer English ability at a disadvantage and increases their anxiety in learning. If the learning of English is one assessment focus, it is suggested that it be included as one learning objective, on the basis of which the teacher can develop an instructional plan.

One central idea of *assessment for learning* is to engage students in the learning process. This can be achieved by sharing with learners the explicit criteria for success and by encouraging learner autonomy in order to guide their own learning through self-assessment. A close examination of the grading criteria revealed that the teachers not only incorporated skill and knowledge criteria, such as projects, exams, or homework, but also often incorporated behavioral and attitudinal criteria, including classroom behavior, effort, attendance, student engagement, and extra-credit activities (although they did not include students' ability and improvement), suggesting that the teachers emphasized and gave credit to students' learning processes and behaviors. In addition, most of the teachers communicated the assessment criteria to students in advance and inform the students how grades were to be assigned. This emphasis on behavioral and attitudinal criteria helps students engage with the learning process. Despite how these assessment techniques were practiced often by most EMI teachers, the Science and Engineering teachers were found to communicate assessment criteria to students less often. They also less often incorporated student engagement into grade calculation and assessed students' individual class participation. This seems to imply fewer teacher–student, or even student–student, interactions in the Science and Engineering EMI classroom.

Self-assessment is an essential practice that promotes learning, or, more specifically, the *assessment-as-learning* process (Black & William, 2009; Dann, 2014). Nevertheless, it was less often practiced in classroom assessment by most of the EMI teachers (rating = 2.63), along with other assessment tools that also directly reflect learning, such as portfolio assessment and hands-on activities. Self-assessment can encourage students to engage in the learning process by requiring them to fully understand the criteria for success and evaluate learning outcomes. Moreover, with proper guidance, this assessment

practice helps students set their own goals, adopt or refine strategies to meet these goals, and monitor their learning progress (Boekaerts & Corno, 2005). These abilities are known as self-regulation, which is a critical feature of an autonomous learner and which enables learners to maintain their motivation for learning, thus increasing the possibility of academic success (Pintrich, 2004; Heikkilä & Lonka, 2006). Despite its importance, self-assessment was not widely practiced in the classrooms; the teachers also perceived themselves to be less skilled in this particular assessment technique.

In addition to self-assessment, assessment of higher-order abilities is also essential because such abilities are critical for equipping students to solve real-world problems and adapt to a rapidly changing world. Moreover, questions involving higher-level cognitive abilities can not only tap into students' acquisition of knowledge but also force them to integrate what they have learned in order to solve new problems, thereby leading to deep learning, the main goal of *assessment for learning* (Cotton, 1991; Jones & Saville, 2016). However, while most teachers were aware of its importance and frequently performed this technique in their classrooms, they perceived themselves to be less skilled in it. Moreover, this assessment practice was found to be practiced less often, not only in large-sized classes but also in classes where English was the medium of instruction. According to our findings, one-third of the teachers with both EMI and CMI experience indicated that they tended to assess students' higher-level cognitive abilities in the EMI classroom less often. Similar findings were reported by Hu and Li (2017). Their study revealed that teachers asked a significantly higher number of comprehension-related questions in the EMI classroom, and suggested that this phenomenon could result from the English difficulties of teachers as well as students. Regardless of the reasons why the teachers in this study employed this assessment technique in EMI less often, teachers should be provided with further training in this aspect of assessment.

To move students forward in learning, it is important to provide them with scaffolding during the learning process. This is particularly important in the EMI classroom, where EFL learners face the double challenge of acquiring content knowledge in a language with which they are less familiar. Although the teachers in our study were aware of students' English limitations and used simple English to present the assessment tasks, most of the teachers required students to perform these tasks in English, with no allowance for the use of their L1. Moreover, it was not common for the teachers to scaffold student performance when required by how the students performed the assessment task in English or to take into account students' English ability when evaluating their performance. This might cause students to feel highly frustrated with learning because of the constraint of their English ability. In addition, students' performance in the assessment task, which was asserted to reflect their absorption of content knowledge, could be severely confounded by their English ability. Therefore, EMI teachers should be aware of the effect of the medium of instruction on academic performance and the necessity to provide scaffolding or implement other strategies to accommodate students' English difficulties. In particular, they should reconsider the role of students' L1 in relation to English in the EMI classroom, in which students' learning of content knowledge is the main instructional focus and takes priority over the learning of English.

Learning is less likely to occur if the assessment cycle is not completed with the teacher providing feedback to the learners. For example, the Assessment Reform Group (2002) proposes as one of the principles for *assessment for learning* that teachers should provide information regarding learners' strengths and weakness and help them plan their next steps in learning. While most of the teachers in our study communicated assessment results to students and provided oral feedback, many of them either never or only occasionally provided corrective feedback regarding students' English language errors, despite how some of them also considered language as one assessment focus. If language learning is one of the intended outcomes of EMI, teachers should provide corrective feedback, since evidence from empirical studies shows that this feedback, whether direct or indirect, is a prerequisite

for sustained improvement in language accuracy (van Beuningen, de Jong, & Kuiken, 2012).

Assessment in the EMI classroom is challenging, especially when English is deemed the *only* language permitted in a classroom in which English is a foreign language to both the teacher and students. As shown in our study, EMI courses were more likely to involve fewer classroom interactions and activities that assess higher-level cognitive abilities, which are critical for the occurrence and consolidation of learning. Moreover, given that content knowledge cannot be assessed independently from the medium through which it is conveyed, English ability actually interfered with student assessment performance, since many of the teachers did not often provide scaffolding or accommodation strategies and disregarded students' English ability when making such evaluations. These assessment practices have a detrimental impact on students' learning in EMI courses. To solve these problems, teachers should be extremely clear about their instructional purposes and align them properly with assessment practices. In particular, if the English language is not the focus of instruction, it would be inappropriate to include it as the focus of assessment, and efforts should be made to ensure that students are not penalized on the basis of their English ability in their assessment performance. On the other hand, if English is one of the assessment focuses, language-learning goals should be included in the instructional objectives and clearly conveyed to the students. Accordingly, the teachers should conduct pedagogical practices to enhance students' English learning. However, this leads to the questions of whether English language learning should be included as one learning goal in EMI, and what the role of EMI should be in language learning in relation to ESP or EAP. This issue is beyond the scope of this research but is worthy of further discussion.

In addition to the efforts teachers should make regarding their assessment knowledge, school and departments also play important roles in facilitating teachers' practice of *assessment for learning*. For example, this study found that course size exerted a significant effect on assessment practices in the EMI classroom. Some learning-oriented assessments, such as assessing students' higher-level cognitive abilities, are practiced less often in classes containing a greater number of students. This phenomenon is not difficult to understand, since such situations are often burdensome in terms of classroom management, reducing teachers' willingness to conduct learning-oriented assessments, which require constant awareness of each student's learning progress and difficulties. Hence, administrative efforts should be made to reduce the sizes of EMI classes. However, controlling class size is only one factor. When universities implement EMI as part of recent trends, they have the duty to provide the necessary resources and support for both teachers and students.

Implications and Conclusions

The study has limitations in terms of the small sample size and the constraints of using a questionnaire to gather information. Future studies will have to collect responses from more EMI teachers to gain a more complete picture of assessment practices in the EMI classroom and teachers' assessment needs. In addition, research methods other than a questionnaire, such as classroom observations and interviews, can be adopted to obtain qualitative data. Thus, a mixed-methods approach integrating quantitative and qualitative data can be used to provide a better understanding of assessment practices in the EMI classroom. For example, while our study only inquired about the frequency of teachers providing feedback to their students, observational studies can obtain information regarding how feedback is provided and the content of the feedback, both of which are highly relevant to the practice of *assessment for learning*.

Despite the limitations of this study, its findings have implications not only for advancing our understanding of teachers' assessment needs in the Taiwanese EMI context but also for developing and implementing useful training programs in support of EMI teachers' professional development with

regard to assessment. For example, although the results demonstrated that EMI teachers in Taiwan universities generally have basic assessment knowledge, further training is still required to enable their assessment skills to mature, particularly those that are critical to the creation of a learning-oriented classroom environment. In addition, the findings that assessment practices in the EMI classroom were mediated by teacher- and course-related factors such as discipline or class size suggest that professional training should take into consideration these contextual factors and customize training courses to serve the needs of their target populations.

Our findings also provide insights to policymakers and university administrators who aim to promote EMI in terms of the resources and support they should provide. For example, administrative efforts should be taken to limit the size of EMI classes. When the number of students in the EMI classroom is reduced, the teacher can better attend to each student's learning needs and difficulties. However, if controlling class size is not feasible, schools or departments should take measures to help teachers manage their time and workload effectively, such as by providing teaching assistants for their classes. Moreover, leaders at the department and university levels should provide academic resources, such as EAP/ESP courses, that can support students' use of English to absorb and present content knowledge. When necessary, schools should also help to establish a model of collaboration between language and content teachers in devising effective classroom assessments. Such collaborative endeavors, though not yet common in the current practice of EMI, have been reported in the literature and are suggested to have the potential to advance students' academic knowledge while addressing their language needs as well as content teachers' pedagogical needs (Wilkinson, 2013; Iyobe & Li, 2017). Last but not least, schools should provide EMI teachers with supportive resources, such as regular assessment workshops or working groups, to facilitate the teachers' development of assessment literacy and to promote the use of assessment for learning.

References

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). *Standards for teacher competence in educational assessment of students.* Washington, DC: National Council on Measurement in Education.

Assessment Reform Group. (2002). *Assessment for learning: 10 principles.* Cambridge: University School of Education.

Black, P., & William, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5-31. https://doi.org/10.1007/s11092-008-9068-5

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54*, 199-231. https://doi.org/10.1111/j.1464-0597.2005.00205.x

Chen, J.-J. (2011). *A study of the policy performance and influence of "The Top University Development Plan" in Taiwan* (Unpublished master thesis). National Taiwan Normal University, Taipei.

Chung, C., & Lo, M. L. (2016). Prospect and case study of English-medium instruction of transportation courses in Taiwanese universities. *English Teaching & Learning, 40*, 87-121.

Cotton, K. (1991). Close-Up #11: Teaching thinking skills. Retrieved from http://www.qsm.ac.il/userfiles/ershad_tarbawi/general/Teaching_Thinking_Skills.pdf

Coyle, D., Hood, P., & Marsh, D. (2010). *Content and language integrated learning.* UK: Cambridge.

Dann, R. (2014). Assessment *as* learning: Blurring the boundaries of assessment and learning for theory, policy and practice. *Assessment in Education: Principles, Policy & Practice, 21*, 149-166. https://doi.org/10.1080/0969594X.2014.898128

Davison, C., & Leung, C. (2009). Current issues in English language teacher-based assessment.

TESOL Quarterly, 43, 393-415. https://doi.org/10.1002/j.1545-7249.2009.tb00242.x

Dearden, J. (2015). *English as a medium of instruction: A growing global phenomenon.* London: British Council.

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2013). *English-medium instruction at universities: Global challenges.* Bristol, Buffalo, Toronto: Multilingual Matters.

Galloway, N., Kriukow, J., & Numajiri, T. (2017). *Internationalization, higher education and the growing demand for English: An investigation into the English medium for instruction (EMI) movement in China and Japan*. London: British Council.

Heikkilä, A., & Lonka, K. (2006). Studying in higher education: Students' approaches to learning, self-regulation, and cognitive strategies. *Studies in Higher Education, 31*, 99-117. https://doi.org/10.1080/03075070500392433

Hu, G., & Li, X. (2017). Asking and answering questions in English-medium instruction classrooms: What is the cognitive and syntactic complexity level? In J. Zhao & Q. Dixon (Eds.), *English-medium instruction in Chinese universities: Perspectives, discourse and evaluation* (pp. 184-203). New York: Routledge.

Jones, N., & Saville, N. (2016). *Learning oriented assessment: A systemic approach.* Cambridge: Cambridge University Press.

Kao, Y-T., & Tsou, W. (2017). EMI course assessment: A survey study of the issues. In W. Tsou & S-M. Kao (Eds.), *English as a medium of instruction in higher education: Implementation and classroom practices in Taiwan* (pp. 183-205). Singapore: Springer.

Lei, J., & Hu, G. (2014). Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *International Review of Applied Linguistics in Language Teaching, 52,* 99-126. https://doi.org/10.1515/iral-2014-0005

Li, M. (2017). Evaluation of learning outcomes in an education course: Does it work? In J. Zhao & Q. Dixon (Eds.), *English-medium instruction in Chinese universities: Perspectives, discourse and evaluation* (pp. 147-164). New York: Routledge.

Macaro, E., Curle, S., Pun, J., An, J., & Dearden, J. (2018). A systematic review of English medium instruction in higher education. *Language Teaching, 51*, 36-76. https://doi.org/10.1017/S0261444817000350

MOE Taiwan. (2009). *Jiao-yu-bu jiang-li da-xue-yuan-xiao tui-dong guo-ji-hua bu-zhu-ji-hua yao-dian* [Directions governing MOE subsidy to the internationalization of universities and colleges] [Data files]. Retrieved from http://edu.law.moe.gov.tw/LawContent.aspx?id=FL037774#lawmenu

MOE Taiwan (2017). *Jiao-yu-bu bu-zhu da-zhuan-yuan-xiao jing-jin quan-ying-yu xue-wei xue-zhi ban-bei ji-hua yao-dian* [Directions governing the MOE subsidy for colleges and universities to promote and refine English-taught degree programs] [Data file]. Retrieved from http://edu.law.moe.gov.tw/LawContent.aspx?id=GL000575#lawmenu

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Reviews, 4*, 385-408. https://doi.org/10.1007/s10648-004-0006-x

Robinson, J. (2010). The effects of test translation on young English learners' mathematics performance. *Educational Researcher, 39*, 582-590.

Rogier, D. (2012). *The effect of English-medium instruction on language proficiency of students enrolled in higher education in the UAE* (Doctoral dissertation). University of Exeter, United Kingdom. Retrieved from https://ore.exeter.ac.uk/repository/bitstream/handle/10036/4482/RogierD.pdf?sequence=2

Tai, H-Y. (2015). Writing development in syntactic complexity, accuracy and fluency in a content and language integrated learning class. *International Journal of Language and Linguistics, 2*, 149-156.

Van Beuningen, C. G., de Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of

comprehensive error correction in second language writing. *Language Learning, 62*, 1-41. https://doi.org/10.1111/j.1467-9922.2011.00674.x

Van der Walt, C., & Kidd, M. (2013). Acknowledging academic biliteracy in higher education assessment strategies: A tale of two trials. In A. Doiz, D. Lasagabaster & J. M. Sierra (Eds.), *English-medium instruction at universities: Global challenges* (pp. 27-43). Bristol: Multilingual Matters.

Yang, W. (2015). Content and language integrated learning next in Asia: Evidence of learners' achievement in CLIL education from a Taiwan tertiary degree programme. *International Journal of Bilingual Education and Bilingualism, 18*, 361-382. https://doi.org/10.1080/13670050.2014.904840

Author Biodata

Naihsin Li received her PhD degree in language acquisition and learning. Her research interest lies in exploring the underlying processes of language learning and examining the factors affecting the development of first language and second language. Currently she works as a postdoctoral fellow in the Language Training & Testing Center, Taiwan.

Jessica Wu holds a Ph.D. in Language Testing. She is the R&D Program Director at the LTTC. She also serves as an adviser to the government at a national level on the development and administration of L1 tests. She has published numerous articles and book chapters in the field of language testing and has presented her work at conferences around the world. She is currently the President of the Asian Association for Language Assessment (AALA, 2018 - 2019).

Appendix A

Subgrouping of Disciplines

Major Disciplines	Departments	
Humanities (*n* = 14)	• Education • Institute of Applied English • Economics • Holistic Education Center • English Department • General Education • Journalism	• Center for Teacher Education • MA Program of Chinese Arts for International Students • Department of Foreign Languages and Literature (*n* = 2) • English Literature and Language • Teaching Chinese as a Second Language • Japanese Language and Culture
Business & Marketing (*n* = 16)	• Transportation Management • Banking and Finance • International Business (*n* = 3) • International College • Marketing and Logistics • Business Administration • Hospitality Management • Information Management	• Bachelor Program of Management Science and Finance • International Business Administration Program • International Business & Trade Program • Institute of Technology Management • Leisure Industry and Health Promotion • Institute of Marketing Communication
Science & Engineering (*n* = 10)	• Photonics • Medicine • Statistics • Plant Pathology • Chemistry	• Animal Science • Electrical Engineering • Natural Biotechnology • Computer Science and Information Management • Institute of Molecular Biology

Appendix B

Demographic Information

Characteristics	Subgroup	No. of teachers
Teacher-related variables		
1. Discipline	Humanities	*n* = 14
	Business & Management	*n* = 16
	Science & Engineering	*n* = 10
2. Teaching experience	5 years or less	*n* = 12
	Between 6 to 15 years	*n* = 15
	16 years or more	*n* = 13
3. EMI experience	5 years or less	*n* = 25
	6 years or more	*n* = 15
4. Country where the highest degree was received	English-speaking countries	*n* = 25
	Other non-English-speaking countries (including Taiwan)	*n* = 15
5. Length of stay in an English-speaking country	Never	*n* = 7
	5 years or less	*n* = 15
	6 years or more	*n* = 18
Course-related variables		
6. Course level	Undergraduate level	*n* = 30
	Graduate level	*n* = 10
7. Course type	Lecture	*n* = 38
	Seminar	*n* = 2
8. Composition of students	Half or more international students	*n* = 10
	More Taiwanese students	*n* = 30
9. Course size	25 students or less	*n* = 16
	Between 26 to 50 students	*n* = 15
	51 students or more	*n* = 7
	Did not report	*n* = 2

Appendix C

EMI Classroom Assessment Practices (ECAP) Questionnaire

Part II.
Section A: Assessment Practices and Knowledge about Assessments
1   Choosing appropriate assessment methods for instructional decisions.
2   Selecting (and revising if necessary) textbook-provided test items for classroom assessment.
3   Revising previously produced self-made tests to match current instructional emphasis.
4   Developing assessments based on clearly defined course objectives.
5   Matching assessments with instruction content.
6   Defining a rating scale for performance criteria in advance.
7   Recording an assessment result on the rating scale/checklist while observing student's performance.
8   Using assessment results when planning teaching.
9   Using assessment results when developing curriculum.
10 Using assessment results when making decisions (e.g., pass, non-pass) about individual students.
11 Using assessment results when evaluating student improvement.
12 Using a norm-referenced grading model (Students are evaluated in relationship to one another, e.g., the top 10% of students receive an A, the next 30% a B, etc.).
13 Using a criteria-referenced grading model (Students are evaluated against an absolute scale, e.g. 95-100 = A, 88-94 = B, etc. Since the standard is absolute, it is possible that all students could get As or all students could get Ds.

14   Assigning different weights to projects, exams, homework, etc. when assigning semester grades.
15   Incorporating extra credit activities in the calculation of grades.
16   Incorporating students' ability in the calculation of grades.
17   Incorporating classroom behavior in the calculation of grades.
18   Incorporating improvement in the calculation of grades.
19   Incorporating effort in the calculation of grades.
20   Incorporating attendance in the calculation of grades.
21   Incorporating students' engagement in the calculation of grades (e.g., asking/answering questions, participating in discussion, etc.)
22   Assigning semester grades.
23   Providing oral feedback to students.
24   Providing written feedback to students.
25   Communicating classroom assessment results to students.
26   Communicating classroom assessment results to other educators.
27   Protecting students' confidentiality with regard to test scores.

Section B: Forms of Assessment and Student Involvement

28   Using tests/quizzes mainly to assess student progress.
29   Assessing students through observation.
30   Assessing students through multiple-choice questions.
31   Assessing students through fill-in-the-blank or short answer questions.
32   Assessing students through essay questions.
33   Designing test items that assess higher levels of cognitive abilities (e.g., critical thinking).
34   Assessing individual class participation.
35   Assessing group (or pair) class participation.
36   Assessing individual hands-on activities.
37   Assessing group (or pair) hands-on activities.
38   Using portfolios to assess student progress.
39   Using students' self-assessment to assess progress.
40   Using presentations to assess student progress.
41   Communicating performance assessment criteria with students in advance.
42   Informing students in advance how grades are to be assigned.

Part III.
English Language and Content

43   The focus of an assessment task is A. Content B. English Language C. Both
44   Using English to present an assessment task to students.
45   Using simple English to present an assessment task to students
46   Taking students' English ability into account when evaluating students' performance in an assessment task.
47   Providing scaffolding (e.g., Chinese translations of key terms in a test paper) to students.
48   Allowing students to perform an assessment task in Chinese without penalty if they have difficulty performing the task in English.
49   Correcting students' language errors when evaluating their performance in English.