THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

ARTICLES FOR FACULTY MEMBERS

Title/Author	Aligning a university English language proficiency measurement tool with the CEFR: a case in Malaysia / Baharum, N. N., Ismail, L., Nordin, N., & Razali, A. B.
Source	*Pertanika Journal of Social Science and Humanities* Volume 29 Issue 3 (2021) Pages 157–178 https://doi.org/10.47836/pjssh.29.S3.09 (Database: ResearchGate)

Title/Author	Aligning the language criteria of a group oral test to the CEFR: the case of a formal meeting assessment in an English for occupational purposes classroom / Shak, P., & Read, J.
Source	*Pertanika Journal of Social Science and Humanities* Volume 29 Issue 3 (2021) Pages 133–156 https://doi.org/10.47836/pjssh.29.S3.08 (Database: ResearchGate)

Title/Author	Exploring vocabulary teaching: planning and challenges with the implementation of audio-visual approach in rural East Malaysian primary-level English language classrooms / Yang, B. L. C., Ramanair, J., & Shing, S. R.
Source	*3L: Language, Linguistics, Literature* Volume 29 Issue 4 (2023) Pages 228-243 https://doi.org/10.17576/3L-2023-2904-15 (Database: UKM Press)

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

ARTICLES FOR FACULTY MEMBERS

Title/Author	**Local problems and a global solution: examining the recontextualization of CEFR in Thai and Malaysian language policies / Savski, K.**
Source	***Language Policy*** **Volume 19 (2020) Pages 527-547** **https://doi.org/10.1007/s10993-019-09539-8** **(Database: Springer Link)**

Title/Author	**Policy change implementation: the case of the CEFR in Malaysian ESL classrooms / C. Alih, N. A., Abdul Raof, A. H., & Md. Yusof, M. A.**
Source	***Journal of Nusantara Studies (JONUS)*** **Volume 6 Issue 2 (2021) Pages 296–317** **https://doi.org/10.24200/JONUS.VOL6ISS2PP296-317** **(Database: UniSZA Press)**

Title/Author	**Supporting English as a second language (ESL) teachers' professional development through multi-platforms' online affinity space / Aziz, A. H. A. A., Zainudin, W. Z. B. W., & Rashid, R. A.**
Source	***International Journal of Interactive Mobile Technologies*** **Volume 14 Issue 9 (2020) Pages 258–264** **https://doi.org/10.3991/ijim.v14i09.11158** **(Database: ResearchGate)**

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

ARTICLES FOR FACULTY MEMBERS

Title/Author	TESL in-service teachers' experiences when evaluating an oral history workbook / Aziz, A. A., & Makhtar, R.
Source	*3L: Language, Linguistics, Literature* Volume 27 Issue 3 (2021) Pages 56–72 https://doi.org/10.17576/3L-2021-2703-04 (Database: UKM Press)

Title/Author	The application of language proficiency scales in education context: a systematic literature review / Zhu, A., Mofreh, S. A. M., & Salem, S.
Source	*SAGE Open* Volume 13 Issue 3 (2023) Pages 1-19 https://doi.org/10.1177/21582440231199692 (Database: ResearchGate)

Title/Author	The CEFR-aligned curriculum execution in Malaysia and other countries: a conceptual paper / Bakar, R. A.; A. M., & Bakar, R. A.
Title/Author	*Malaysian Journal of ELT Research* Volume 19 Issue 1 (2022) Pages 1-15 https://doi.org/10.52696/TGCT6849 (Database: ResearchGate)

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Aligning a university English language proficiency measurement tool with the CEFR: a case in Malaysia / Baharum, N. N., Ismail, L., Nordin, N., & Razali, A. B.**
Source	***Pertanika Journal of Social Science and Humanities*** **Volume 29 Issue 3 (2021) Pages 157–178** **https://doi.org/10.47836/pjssh.29.S3.09** **(Database: ResearchGate)**

Pertanika J. Soc. Sci. & Hum. 29 (S3): 157 - 178 (2021)

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Aligning a University English Language Proficiency Measurement Tool with the CEFR: A Case in Malaysia

Nurul Najwa Baharum*, Lilliati Ismail, Nooreen Nordin and Abu Bakar Razali

Department of Language & Humanities Education, Faculty of Educational Studies, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

ABSTRACT

The Common European Framework of Reference (CEFR) describes the capability of learners' language skills at six reference levels. It is internationally recognised as the standard language proficiency framework for describing language learning, teaching and assessment. Many countries, including Malaysia, have attempted and invested tremendous efforts to adopt the CEFR as a reference for language ability at all levels of education. However, there are many ways of adopting CEFR, and it is a continuous process of alignment between curriculum and assessment. In this regard, this study is carried out to examine how a Malaysian university attempts to demonstrate this alignment by correlating the scores obtained from English language proficiency courses in the university, called the English Language Competence Score Average (ELCSA), to a CEFR-aligned English language proficiency test (Linguaskill). The results showed an overall significant positive correlation that varied in strength. The overall correlation was 0.371, a positive but weak correlation whereby the strongest correlation was seen between ELCSA and CEFR Writing score with a correlation of 0.417, which is positive and moderate in strength. Therefore, it could be identified that a score of 3.25 and 3.5 on the ELCSA can be considered equivalent to a Linguaskill score of 160 (CEFR Band B2). It could be considered that the B2 CEFR level could be subdivided into lower and higher B2. However, there is a need to correlate ELCSA with other CEFR-aligned tests and perform further revisions to the English language proficiency programme at the university to successfully benchmark the programme and its assessment tool, ELCSA, with the CEFR.

Keywords: Assessment, benchmarking, CEFR, English language programme, language testing

ARTICLE INFO

Article history:
Received: 16 July 2021
Accepted: 04 October 2021
Published: 30 November 2021

DOI: https://doi.org/10.47836/pjssh.29.S3.09

E-mail addresses:
najwabaharum96@gmail.com (Nurul Najwa Baharum)
lilliati@upm.edu.my (Lilliati Ismail)
nooreen@upm.edu.my (Nooreen Noordin)
abmr_bakar@upm.edu.my (Abu Bakar Razali)
*Corresponding author

ISSN: 0128-7702
e-ISSN: 2231-8534

INTRODUCTION

There are increasing concerns in establishing standards for the English Language in terms of international benchmarking worldwide (Read, 2019); such language benchmarks standard can be an expressive scale of language ability (Inguva, 2018). Establishing these standards can be quite important in securing places in international higher education institutions and for employment in international companies. The Common European Framework of Reference, CEFR, has become an international benchmark for language competency in many countries, even beyond Europe, such as Mexico, Canada, Japan, and Vietnam. Additionally, many international high-stakes tests such as the IELTS, TOEFL and TOEIC have now been aligned to the CEFR, further underscoring how the framework has gained acceptance and credibility worldwide (Don & Abdullah, 2019). However, despite CEFR being adopted worldwide, research has also claimed that the CEFR still lacks links with stakeholders, socio-educational contexts and empirical validation (Ali et al., 2018).

The Malaysian government acknowledges and stresses the mastery of the English language to gain economic and social leverage in the globalised world. Therefore, it is essential to establish standards and benchmarks that are accepted worldwide to measure proficiency levels among Malaysians. The English Language Standards and Quality Council (ELSQC) and the English Language Teaching Centre (ELTC) of the Malaysian Ministry of Education were given the task to align the Malaysian English education curricula and assessment with the CEFR, as well as to develop a roadmap for systematic reform of Malaysia's English language education (Prakash, 2019). Following this educational shift, the primary and secondary levels of education have replaced their English language textbooks with CEFR-aligned textbooks. These actions were also followed by the alignment of SPM and MUET examinations in which the results of the test takers English language proficiency were banded against the CEFR descriptors (Sufi & Stapa, 2020).

One of the key issues that surfaced during the adoption of the CEFR in Malaysia was the fear that the Malaysian National Education Philosophy would be side-lined and European cultural values and elements would instead dominate local and national content (New Strait Times, 2019). However, the ministry has organised programmes for teacher training, curriculum familiarisation and adaptation, as well as continued efforts in providing more resources. The use of of-the-shelf CEFR-aligned textbooks (as textbooks for National primary and secondary schools) that were carefully selected, vetted and revised to suit the Malaysian context, by working closely with the publishers, has proven to be more cost-effective and offers a wider acceptance of other cultures along with providing a variety of ways of using the English language in different contexts (Sani, 2018). Implementing the English Language Education Reform in Malaysia was foreseen to be complex, costly and requires persistent efforts and tremendous patience. However,

all these endeavours are for the national advancements that will benefit Malaysia's current and future generation (Ministry of Education, 2015).

Research Problem

While there are some positive indications in referencing language performance against the CEFR at the pre-tertiary level, as stated in the Cambridge Evaluation study in 2017, the impact of such an initiative is not yet seen at the university level (Zulkefli, 2017). The English Language Education Reform in Malaysia stresses the importance of implementing the CEFR in universities to address problems related to poor English communication skills among graduates, which would inadvertently negatively affect their learning experience, employability potential and realise the national agenda (Ministry of Education, 2015). Based on The Roadmap, it was stated that university students are to possess a CEFR B1 level upon university entrance. The Department of Higher Education, Malaysia, stipulated that 'international students' must also sit for exams that reference the CEFR to fulfil the English requirements for university admission purposes (Jaafar, 2019). Additionally, students are required to reach a proficiency of CEFR B2/C1 upon graduation. In accomplishing the required CEFR condition among university graduates, The Roadmap implies that students' English language proficiency may need to be reassessed by the institution prior to their completion of studies (Sufi & Stapa, 2020).

Furthermore, the Malaysian Qualifications Agency (MQA), responsible for quality assurance and accreditation of Malaysian universities, specified that English proficiency courses that are not equated with the international benchmark, CEFR, can no longer be used to fulfil university requirements (Malaysian Qualification Agency, 2020a). Furthermore, the MQA also stated that a CEFR minimum proficiency of C1, or its equivalent in the relevant language, is required to pursue certain job positions in tertiary level institutions (Malaysian Qualification Agency, 2020b). Therefore, there are increasing attempts, demands, and a heightened level of importance for universities to be CEFR-aligned not only of their courses but also their entry and exit grade requirements of universities.

Noticeably, the alignment of CEFR within the tertiary level of education is underexplored and is an area of concern in which further research is required as it affects the efficiency of the English Language proficiency among university graduates. Given that the CEFR is required by the Ministry of Education Malaysia to be aligned with the curricula in the tertiary level education and considering that the Malaysian government has invested a substantial amount of money and effort in aligning the curriculum, it is important to investigate the alignment of English language proficiency assessments at different levels of education to the CEFR standards. Furthermore, investigating the alignment between university English language proficiency evaluation measures

and the CEFR could provide invaluable information to policymakers and test developers about the predictability and comparative values of the university English proficiency assessment with a well-recognised international standard for language education and assessment, namely the CEFR.

Hence, this paper seeks to fill the gap in the literature by attempting to align the accumulative scores obtained by undergraduate students who took an English proficiency programme in a Malaysian public university with the CEFR scores based on the Linguaskill test students sat for. The English proficiency score selected for this study is the ELCSA accumulative score. ELCSA stands for English Language Competence Score Average, an accumulated score derived from a package of English language proficiency courses in University Putra Malaysia. More specifically, the paper will attempt to firstly examine the relationship between the scores obtained in the ELCSA and the overall as well as individual language skill scores on the Linguaskill test, and secondly, identify the ELCSA score that is equivalent to a CEFR B2 level which has been targeted as the minimum CEFR level for Malaysian university graduates. By doing so, the paper can contribute to a greater understanding and contextualisation of the CEFR. Furthermore, aligning ECLSA scores to the CEFR will help provide comparative scores in ELCSA with Linguaskill. It could then provide indications of test-takers CEFR levels based on ELCSA accumulative proficiency scores and could assist and contribute to the university's benchmarking efforts of an internally developed English proficiency measurement tool with international standards.

LITERATURE REVIEW

Benchmarking

There are various definitions given on the concept of benchmarking in the literature. For example, Bogan and English (1994) stated that benchmarking is the continuous pursuit of best practices. By establishing measurement points, comparisons can be made for reasons of learning, adapting and ultimately resulting in better performance, which is the main purpose of benchmarking (Fisher, 1996). The essence of benchmarking is also inspiring ongoing learning and boosting organisations to be at their best (Zairi, 1996). The intent of benchmarking is to aid organisations in establishing a baseline performance criterion that should be complied with (Nwabuko et al., 2020). Similarly, Keegan and O'Kelly (2012) consider benchmarking as a method of comparison between organisations to obtain insights from each other. Benchmarking is operative in identifying best practices, and these practices are applied for the benefit of the organisation (Alosani et al., 2016).

In education, especially with second language learning, benchmarking is required when measurable standards are set for learning (Inguva, 2018). Benchmarking in assessments ordinarily attends to the purposes of evaluating and monitoring program efficiency, planning curriculum and

instruction, communicating expectations for learning and predicting future performance whereby it would operate finest when it is specifically designed to deliver the data required for enhancements to be made (Herman & Osmundson, 2010). Benchmarking could also provide information about the position of a specific student, class, or institution in terms of ranking (Canadian Language Benchmarks, 2012). There are many views to the term benchmarking as it is used in various contexts. Nevertheless, it could be said that benchmarking is an ongoing process of seeking the best practices by making comparisons and creating points of reference so that the effectiveness of a particular programme could be identified and further improved.

Regarding benchmarking language learning and assessment, the CEFR framework has proven to be an influential baseline for the development of language curricula and assessment around the world (Read, 2019). However, benchmarking curricula to the CEFR has brought a great deal of discussion whereby some countries found it problematic to strike a balance between the appeal of establishing mutual international standards and the importance of representing the unique educational and social contexts of distinct countries in language learning (Read, 2019). In Taiwan, for example, attempts were made to adopt the CEFR, which meant that their recognised tests needed to be calibrated against the CEFR (Wu, 2012). However, Wu (2012) pointed out that there were several problems with the process of calibrating tests to the framework, such as the conceptual difficulty in comparing the results of tests that have been designed differently and the lack of technical expertise to confirm the alignment of CEFR upon their tests. Furthermore, Wu (2012) mentioned an unclear relationship between the assessment of English language proficiency according to the CEFR and the grading criteria used by universities. Additionally, it was reported that Taiwan students did not have the exposure to the language to use it communicatively as described on the CEFR scales (Cheung, 2012).

On the other hand, there were also instances where some researchers suggested developing a new framework of reference altogether. For example, in China, rather than adapting the CEFR, the development of a Common Chinese Framework of Reference for Languages (CCFR) or currently known as China Standards of English (CSE) which has been established without much reference to the other frameworks and with their separate tests as measures of student achievement was proposed (Jin et al., 2017). Meanwhile, there are instances where these efforts to align the CEFR with curricula succeeded. For example, in Japan, a team of language researchers at the Tokyo University of Foreign Studies undertook a project to adapt the CEFR to the Japanese context, which successfully resulted in a version of the framework labelled CEFR-J whereby they added sublevels (A1.1, A1.2 and A1.3) to reflect better the degree of English ability (Markel, 2018).

English Proficiency Courses in a Malaysian University

One of the measures taken by universities to improve English language proficiency among students is to offer a range of English language courses required for students to pass as part of their graduation requirements (Rethinasamy & Chuah, 2011). It is also a measure taken by one of the research universities in Malaysia, Universiti Putra Malaysia (UPM). They revamped their English language proficiency level courses and developed an innovative package referred to as the English Language Experience (ELEx). The purpose of the ELEx package is to engage students with the language in a variety of formal and informal situations as well as involve more student-centred courses and task-based language activities. ELEx consists of three components, namely conventional courses (LPE), non-alphabet grade preparation courses (CEL), and language activities (LAX). The number of English skills courses, CEL courses, and LAX activities that students need to take is determined by the MUET results obtained before students start their studies at UPM. Therefore, students who get low results in MUET need to take more English courses and activities than students who achieve high results.

UPM enhanced the ELEx package by implementing a cumulative, and summative assessment of language performance referred to as the English Language Score Competency Average (ELCSA). The ELCSA is obtained calculating average achievement points for the two components of ELEx, namely conventional courses (LPE) and preparatory courses (CEL). LAX activities are not included in this calculation because they serve as support (scaffold) to forming English language skills by allowing students to use the language and build confidence in its use. The assessment of this English language achievement, named English Language Competence Score Average (ELCSA), will be calculated at the final stage of the study program and will be stated in the student transcript. It is also important to add that the ELSCA is isolated from the existing CGPA. Therefore, it does not interfere with nor affects the student's CGPA. Therefore, ELCSA serves as a cumulative summary of the student's achievement in their English language skills.

As mentioned, the targeted level for university graduates is the B2 level of the CEFR whereby at this level, it is expected that graduates can understand complex texts, tackle other abstract topics, engage in discussions as well as be able to communicate with native speakers with ease (Ministry of Education, 2015). A major motivation in introducing the ELCSA is to provide a measure that can indicate the student's English language performance according to the CEFR bands. The ELSCA scores could act as a comparison point compared to other CEFR achievement tests such as IELTS, TOEFL, Linguaskill and MUET. In addition, it could evaluate the effectiveness of the ELEx package.

The ELEx Program at UPM and the Development of ELCSA

In 2013, the Centre for the Advancement of Language Competence (CALC) in UPM executed its undergraduates' English Language Experience (ELEx). The students must follow a carefully developed set of courses and activities during the whole duration of their educational programme.

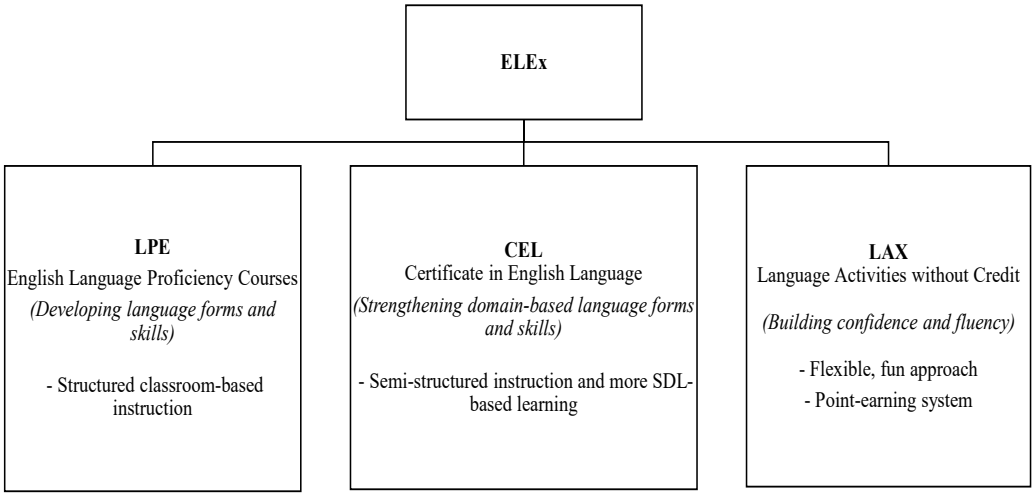


Figure 1. Overview of the ELEx Structure

The components that construct the ELEx are portrayed in Figure 1 (Centre for the Advancement of Language Competence, 2013). As displayed, the ELEx package comprises three significant parts. The LPE component focuses on building the basis of language whereby the knowledge of vocabulary and grammar would be solidified to achieve language accuracy and fluency. Similarly, the CEL component is also constructed to assist in the mastery of vocabulary and grammar. In addition, it also emphasises domain-based learning, whereby it serves to accommodate learning English for general, academic and professional purposes. On the other hand, the LAX component focuses on incidental learning via task-based activities, aiming improve students' confidence and familiarity in using the language.

Recognising that students vary in levels of proficiency from being very limited to very proficient users of English, the ELEx package is designed to cater to students' specific language needs, which is identified based on the levels that they have achieved in their MUET results (Band 1 to Band 6). MUET is a compulsory test that students have to take in order to be admitted into a university. Thus, ELEx provides students of MUET Band 1 or 2 with an intensive programme that aims to supply essential assistance to help foster their confidence in the language while assisting them to meet

their immediate needs required in academic tasks. For those with MUET bands of 3 to 6, the package would provide programmes and courses that aim to enhance further and polish their language competency as well as amplify their confidence and fluency in order to be more linguistically marketable (Abdullah et al., 2015).

The implementation of the ELEx package has proven to show positive outcomes in improving language proficiency, especially among the less proficient students, whereby the students portrayed higher willingness to use the language, which resulted in higher participation and interaction in various contexts (Mustafa, 2018). Although the ELEx package is compulsory for all students, it is seen as more of assistance for students to cope better with their studies rather than an obligation or a test (Sani, 2020). It could be said that this package delivers and is in line with the aspiration of the Ministry of higher education in developing graduates that possess adequate English language abilities. In fact, in the 14th parliament meeting on July 22, 2019, the Ministry of Education mentioned and acknowledged the ELEx package from UPM as one of the government's efforts in assisting youths in mastering the English language (Parlimen Malaysia, 2019)

Since its implementation, the assessment for the ELEx package for each student was evaluated via an alphabetical grade for the LPE component, a 1 to 4 level for the CEL component and a Satisfactory / Unsatisfactory grade for the LAX activities. However, a comprehensive evaluation of the student's English language proficiency was not provided at the point of graduation. Therefore, the English Language Competence Score Average (ELCSA) was established, and the students will obtain scores ranging from 0.0 to 5.0. It is obtained by calculating average achievement points for two components of ELEx, namely the conventional courses (LPE) and preparatory courses (CEL). LAX activities are not included in the calculation as the activities in LAX were for scaffolding purposes that provided opportunities to use the language and build confidence. The division of courses and calculation of the student's ELCSA is shown in Table 1.

Table 1

Courses are taken into account for the determination of ELCSA according to MUET results

MUET band level	CEL courses	Number of CEL courses	LPE courses	Number of LPE courses	Total courses
1 – 2	CEL2102, CEL2103 and one of the courses CEL2105 / 2106/2107	3	LPE2301, LPE2501	2	5

Table 1 *(Continued)*

MUET band level	CEL courses	Number of CEL courses	LPE courses	Number of LPE courses	Total courses
3 – 4	CEL2103 and one of the courses CEL2102 / 2105/2106/2107	2	LPE2301, LPE2501	2	4
5 – 6	CEL2103	1	LPE2402 and / or LPE2502	1 or 2	2 or 3

The scores that will be given for the achievement of each CEL and LPE course are in Table 2 and Table 3.

Table 2

Scores for Certificate in English Language (CEL) Course

Level	Score
1	0
2	3.0
3	3.5
4	4.0

Table 3

Score for courses of Language Proficiency in English (LPE)

Alphabetical Grade	LPE2301*	LPE2501*	LPE2402**	LPE2502**
A	4.0	4.0	4.5	4.5
A-	3.7	3.7	4.0	4.0
B+	3.3	3.3	3.7	3.7
B	3.0	3.0	3.3	3.3
B-	2.7	2.7	3.0	3.0
C+	2.3	2.3	2.7	2.7
C	2.0	2.0	2.3	2.3
C-	1.7	1.7	2.0	2.0
D+	1.3	1.3	1.7	1.7
D	1.0	1.0	1.3	1.3
F	0	0	0	0

Taken only by MUET students 1 - 4

** One or both courses are taken only by MUET Students 5-6

The calculation of ELCSA is based on the total score obtained divided by the number of selected LPE and CEL courses taken (i.e., on average). The average score obtained will determine the level of ELCSA as described in Table 4. As mentioned before, LAX activities are excluded from the calculation because their main purpose is to build confidence in using English. The LPE 2401 course is also excluded because it is in special preparation for students with MUET results 1 and 2. Higher scores are given for LPE2402 and LPE2502 courses as these two courses are high-level courses taken only by MUET students 5 and 6.

Table 4

Scale for English Language Competence Score Average (ELCSA)

Score	Competency	Grade	Estimated CEFR
3.90 above	Excellent	A+	C2
3.725 - 3.89	Very High	A	C1
3.5 - 3.724	High	A-	B2
3.0 - 3.49	Competent	B+	B2
2.5 - 2.99	Average	B	B1
2.5 and below	Low	B-	B1

Linguaskill English Language Proficiency Test

Linguaskill is one of the tests provided by Cambridge Assessment English and has just recently been introduced in Malaysia in 2020. In implementing the CEFR, Cambridge Assessment English played a contributing role and possesses increasing, ongoing and various evidence that supports it to be the embodiment and reflection of the CEFR in multiple aspects (Cambridge Assessment English, 2021b). Linguaskill is a CEFR-aligned, computer-based, multi-level test that assesses one's English language proficiency in writing, reading, listening and speaking (Cambridge Assessment English, 2019). Table 5 illustrates the Linguaskill scores and corresponding CEFR levels.

Table 5

CEFR scores and levels

Cambridge English Scale Score	CEFR Level
180+	C1 or above
160–179	B2
140–159	B1
120–139	A2
100–119	A1
82–99	Below A1

The Linguaskill test provides two test options, namely Business and General, whereby Linguaskill Business assesses the familiarity of the test-taker towards the language of business. At the same time, Linguaskill General would focus on assessing English used in daily life (Cambridge Assessment English, 2019). Linguaskill Business has replaced BULATS that was officially discontinued on December 6, 2019 (Cambridge Assessment English, 2021a). While the Ministry of Higher Education Malaysia (MOHE) had specified several English competency tests (e.g., MUET, IELTS, and TOEFL iBT) that can be recognised by universities to meet English language requirements for student admission, the Linguaskill, Cambridge English Qualifications and OET was also added to the list in 2020 (Cambridge Assessment English, 2021d) Linguaskill reports up to a maximum score of 180+ which is equivalent to C1 or above on the CEFR scale (Cambridge Assessment English, 2021c). Linguaskill was developed by a team of experts and is supported by artificial intelligence. The trial report in April 2016 shows that the Linguaskill test scores are reliable and precise (Cambridge Assessment English, 2016). An analogous measure, the Rasch reliability, was used, and each test obtained a reliability coefficient over .90, which is considered adequate. Whereas the target level of precision was roughly 90% in which most of the tests that failed to reach the target precision were at the extremes of the CEFR: Level A1 or below and C1 or above (Cambridge Assessment English, 2016).

For this research, the Linguaskill General test was used. As mentioned earlier, the Linguaskill General test assesses language used in day-to-day life. The test would include topics involved with studying and working, making plans, travel and technology. Thus, it makes the test suitable for a broad spectrum of organisations, university admissions or exits. The test could also be used for recruitment roles that do not require specialist business terminology; for instance, it would be suitable for employees who are required to showcase their strong command in English to perform their roles effectively.

The Linguaskill General test has three modules which are reading and listening, speaking and writing. The reading and listening tests are adaptive according to the candidate's proficiency level, meaning that each candidate would face a different set of items on their test based on how well they answered the previous question (Cambridge Assessment English, 2018). Although there are not a fixed number of questions, each question the candidates' answer would help the computer understand their level better. The test finishes when the candidate has answered enough questions for Linguaskill to identify their level accurately. The writing test uses innovative auto-marker technology whereby the computer automatically marks it. Meanwhile, a hybrid approach was taken to mark the Linguaskill Speaking test, which uses auto-marking technology and human examiners to ensure efficiency (Xu et al., 2020).

RESEARCH METHODOLOGY

Generally, there are two approaches in which alignment to the CEFR can be adopted: the direct alignment and the indirect alignment (Bruce & Hamp-Lyons, 2015). The direct alignment would require much expertise, resources, and funding which is made possible by large organisations such as Cambridge English Assessment (Ali et al., 2018). Due to time and financial constraints, the direct approach would not be feasible. Alternatively, the indirect approach to the CEFR is adopted by mapping test scores to the CEFR-aligned scores. However,

certain factors regarding the language test such as its purpose, format, test-takers, and the scoring system should be considered before the indirect alignment can be made (Ali et al., 2018). This indirect linkage via 'equation' to an existing test already linked to the CEFR is one of the recommended approaches in the Council of Europe's Manual (Cambridge English Language Assessment, 2011).

According to the Standards for Educational and Psychological Testing (American Educational Research Association, 1999), scores can be considered 'comparable' or 'equivalent' when the test's features are closely similar to each other (Lim, 2017). In this case, this study attempts to uncover the relationship between ELSCA scores and the Linguaskill test scores. Both are designed to measure English proficiency directed towards the goal of real-world applications. Therefore, in order to fulfil the purpose of this research, a quantitative, correlational design was utilised in this study involving the collection of quantitative data followed by a correlational analysis as the study intended to examine the extent to which two or more variables relate to one another (Rockinson-Szapkiw, 2012).

Data Collection and Analysis

The researchers were given access to the participant's ELSCA scores and their full Linguaskill test report, including their overall CEFR score and language skill scores. The researchers used Excel to compile the participant's scores accordingly and then proceeded to use the IBM SPSS Statistics software to calculate the Spearman Rho correlation to uncover the relationship between the two variables. A Spearman Rho correlation was used in this study as it can describe two variables in a monotonic relationship. It should be mentioned that the Spearman Rho correlation seems most befitting as it is suitable for data that is, either ordinal, interval and ratio variables, continuous and non-normally distributed (Schober et al., 2018). In ensuring the standard of quality when assessing the correlational analysis, outliers were addressed and removed. The presence of outliers is common in data collection due to various reasons. It, therefore, is crucial to be dealt with prior to the analysis to ensure the overall reliability of the results (Kwak & Kim, 2017). Additionally, a scatter plot was constructed to observe the relationship between the two variables further, and a trend line was identified.

Sampling

The participants were 197 final year undergraduates from six Science, Technology, Engineering and Mathematics (STEM) based faculties. Table 6 shows the participant's profile.

Table 6

Participants' profile

Participants	Category	Number of test takers	Percentage (%)
Gender	Male	41	20.8
	Female	156	79.1
Age range	21-23	153	77.7
	24 and above	44	22.3
Faculty of	Biotechnology and Biomolecular Sciences	34	17.3
	Computer Science and Information Technology	32	16.2
	Engineering	34	17.3
	Food Science and Technology	34	17.3
	Medicine and Health Sciences	30	15.2
	Science	33	16.7

As shown in Table 6, the participants were 197 final year students (M= 41, F= 156) from six STEM-based faculties in UPM. The purposive sampling method, specifically the Homogenous Sampling, was applied as this sampling form focuses on a particular characteristic of a population where they share similar traits (Etiken et al., 2016). In this case, the participants were chosen according to the following criteria; 1) Participants have completed their undergraduate programmes and therefore also obtained their ELSCA scores. 2) Participants have taken the Linguaskill test and obtained their CEFR band level. 3) Participants were among the STEM-related faculties. The number of participants from each faculty ranged from 30 to 34. The highest number of participants were from the Faculty of Engineering (n= 34), the Faculty of Food Science and Technology (n= 34), and the Faculty of Biotechnology and Biomolecular Sciences (n= 34). It is followed by the Faculty of Computer Science and Information Technology (n= 32) and the Faculty of Medicine and Health Sciences (n= 30). This study specifically chose Science Technology Engineering and Mathematics (STEM) undergraduates because based on their MUET scores, the STEM undergraduate students have varied levels of English language proficiency, which may provide better insights into the correlation between the CEFR and ELCSA

scores. For courses related to English, they are required to meet a MUET band 4 to be admitted in the course (UPM, n.d.). Furthermore, past research has shown that STEM graduates have low employment rates, possibly due to a lack of multiple skills and English proficiency (Murtaza & Saleh, 2018; Thomas, 2019). Additionally, the participants obtained both an ELSCA and a Linguaskill General score, thus allowing the comparison and correlation between ELCSA and Linguaskill.

RESULTS

The performance of the students on the ELCSA and CEFR is presented in Table 7.

Table 7
Performance of Respondents on ELCSA and CEFR

	N	Minimum	Maximum	Mean	Std. Deviation
ELCSA	197	3.000	4.000	3.667	0.211
CEFR	197	122	180	168.43	10.012

The mean ELCSA and CEFR Linguaskill scores were 3.667 and 168.430, respectively. Thus, the Linguaskill score indicates that, on average, the UPM STEM undergraduates had successfully achieved the B2 level as targeted by Malaysia's Ministry of Education.

A correlational analysis between the CEFR and ELCSA scores is presented in Table 8.

Table 8
Correlation between ELCSA and CEFR scores

			ELCSA	Overall (CEFR)
Spearman's Rho	ELCSA	Correlation Coefficient	1.000	0.371**
		Sig. (2-Tailed)	0.000	0.000
		N	197	197

**. Correlation is significant at the 0.01 level (2-tailed)

In fulfilling research objective 2, results show a positive, weak relationship according to the Guilford Rule of Thumb between ELCSA and CEFR scores. In addition, results of Spearman Rho correlation indicated that there was a significant positive association between the overall ELCSA scores and CEFR scores, (r_s (195) = 0.371, $p < .05$).

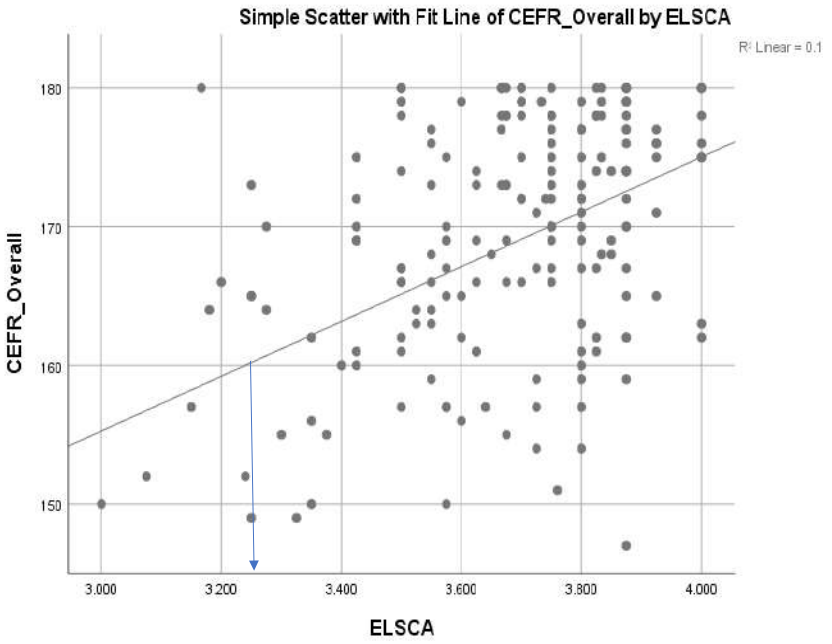


Figure 2. Simple Scatter plot of ELCSA and CEFR scores

Scores on the CEFR and ELCSA were also placed on a simple scatter plot, and based on the trend line in the scatter plot (Figure 2), a score of approximately 3.25 on the ELCSA can be considered equivalent to a Linguaskill score of 160 (CEFR Band B2). However, the trend line does not allow for predicting the C1 Band based on the ELCSA.

The correlations between the language components in the Linguaskill and the ELCSA are presented in Table 9.

Table 9

Correlation between ELCSA and CEFR scores

			Writing (CEFR)	Reading	Speaking	Listening
Spearman's Rho	ELCSA	Correlation Coefficient	0.417**	0.360**	0.249**	0.179*
		Sig. (2-Tailed)	0.000	0.000	0.001	0.012
		N	196	195	188	197

** Correlation is significant at the 0.01 Level (2-Tailed)

Regarding the individual skills, there are correlations of varied strengths between the ELCSA and each of the four skills. There is a positive and moderate relationship between ELCSA and CEFR Writing scores with a correlation coefficient of 0.417. Also, there was a positive, low relationship between ELCSA and CEFR Reading scores with a correlation coefficient of 0.360 and CEFR Speaking scores with a correlation coefficient of 0.249. However, although positive, the relationship between ELCSA and Listening scores was negligible, with a correlation coefficient of 0.179. The relationship that was considered best and strongest was that of ELCSA and writing skills. For that reason, as well as writing being especially important in academic contexts, this relationship is further explored as in Figure 3 in order to determine the ELCSA score that would best reflect a B2 CEFR level.

The scatter plot of scores on the ELCSA, and the Linguaskill Writing skill is presented in Figure 3.

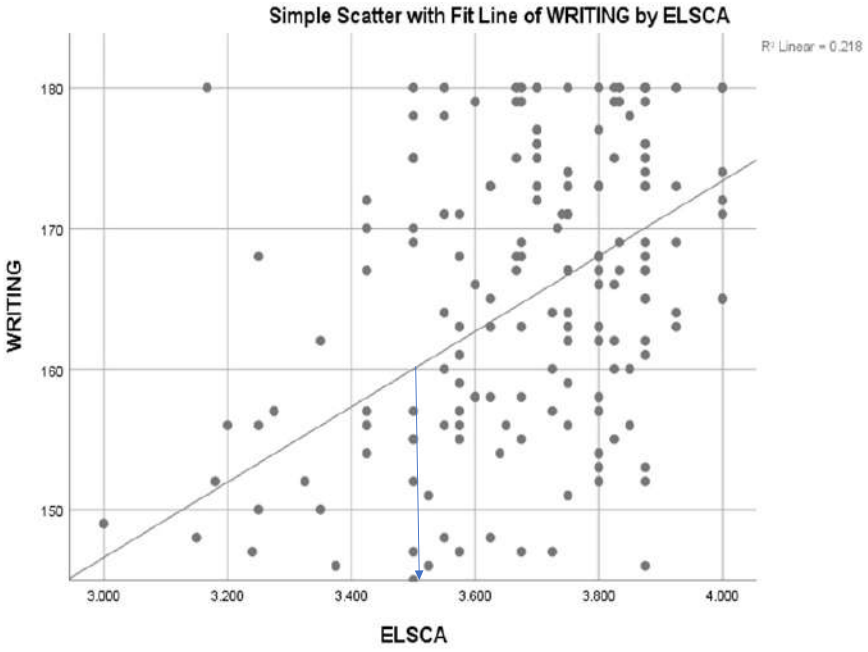


Figure 3. Simple Scatter plot of ELCSA and CEFR Writing scores

As the writing component of the Linguaskill test was the language skill that yielded the strongest correlation with the ELCSA, the scatter plot was used to identify the ELCSA score comparable to a B2 CEFR level. Based on the trend line, the ELCSA score of approximately 3.5 could be identified as equivalent to the CEFR Writing score of 160, which the Linguaskill test specifies as representing the B2 level.

DISCUSSION

The study demonstrates that there is a positive relationship between the ELSCA and CEFR scores. It means that the students that managed to get a high score in ELSCA also managed to get a high CEFR score in the Linguaskill test, and it is likewise for those who received low scores in ELSCA also possessed a low CEFR score. The correlations did, however, differ to some extent in terms of strength. The possible explanation for why the correlations varied in terms of strength is that, though the two scores both measure English proficiency for real-world applications, the two measurements' nature and grading scale differ. While the ELCSA score is cumulative based on language courses taken over time, the Linguaskill test is an English proficiency test. In terms of the grading scales, ELCSA is a score that ranges from 0.00 to 5.00, while the highest possible score obtained in the Linguaskill has a maximum score of 180, which is considered, as equivalent to a C1 and above grade on the CEFR. However, despite the varied strengths of the correlation, the data shows that a positive correlation exists, indicating that the variables move in the same direction.

Furthermore, this paper has shown that it is possible to use an established test that is CEFR aligned as a reference to determine the required scores that match a B2 level in a university English proficiency programme. For example, this study shows that a 3.25 score in the Writing component of the ELCSA corresponds to the B2 CEFR level. In comparison, a 3.5 overall score for ELCSA corresponds to the B2 CEFR level for overall English language proficiency. Thus, it could be assumed that a student who achieves a score of 3.25 in the ELCSA Writing component is at the B2 level of proficiency in terms of writing skills. Also, achieving a score of 3.5 in the ELCSA overall score would mean that a student is at B2 level for overall English language proficiency. This benchmarking is useful, as it can indicate a student's CEFR level using an internally developed university English language programme. Determining the score corresponding to the B2 CEFR level is also important as university students are expected to have a minimum B2 level of proficiency upon graduation. Notably, in so far as the students' performance is concerned, 72.6% of the STEM participants in the study managed to achieve the target that the Malaysian Ministry of Education had set by obtaining the minimum CEFR level of B2 for Malaysian university graduates. The other ten per cent of the participants had exceeded the target and managed to achieve C1, while only 17.3% achieved B1 and fell below the Ministry target.

Previous studies had mentioned that caution should be taken when aligning assessments using CEFR as it was implied that although the different tests use related criteria and are based on descriptors of the same however the perceived equivalence is only assumed (Foley, 2019). Additionally, it should also be considered that even though tests such as IELTS has been aligned to the CEFR, the alignment does not refer to the scores of specific language skill; instead,

it refers to the overall scores (Ali et al., 2018). Nonetheless, a study conducted by Universiti Malaysia Pahang attempted to contextualise the CEFR with their English Writing Language Proficiency Test. Their preliminary analysis has shown that the CEFR-A1 is sufficient in describing their lowest band (Band 1) and that the CEFR C2 and C1 would describe their highest bands, namely band 8 and 9. It was also mentioned that it was necessary to further describe the subcategories of the level of proficiency in order to address all of their bands as their English proficiency test had nine bands altogether (Ali et al., 2018). Therefore, it could be said that, despite being cautious of comparability aspects and over emphasis on standardisation, attempts for an alignment can be made possible. However, it is important to note that fundamentally, the CEFR was originally devised to assist the planning of curricula and that the common reference levels are for further facilitation (Foley, 2019).

CONCLUSION

In seeking to align the accumulative ELSCA scores with the Linguaskill CEFR scores, the authors conclude that there is a positive correlation between the ELSCA scores and the CEFR scores—which shows that there is a possibility in using performance in an English language proficiency programme to predict CEFR levels. Furthermore, this study has also shown that the ELCSA can be used with either the Linguaskill overall score or the writing score to predict and determine CEFR levels, especially to indicate whether or not the student has attained B2 in the CEFR as required by the Ministry of Education for university students upon graduation. Due to this alignment, it can be said that UPM is on the right track in benchmarking its language proficiency programmes with the CEFR. However, it is important to ensure the efficiency of their language programmes and make improvements where necessary.

It is suggested that for future research, attempts should be made to benchmark language programmes in different higher learning institutions to the CEFR. Given that the Linguaskill test is now accepted and adopted in the admission and exit requirements of universities in Malaysia as an alternative to MUET, IELTS, TOEFL and other tests, language centres should consider providing training for students to prepare for such tests or even become centres to carry out the tests. It could further enhance the curriculum of language centres and the practices of language instructors to be more CEFR-aligned. Consequently, this would increase the student's familiarisation with the CEFR and help them develop their language proficiency in line with the CEFR. In sum, this paper contributes knowledge that an alignment between a language proficiency programme of a Malaysian university and the CEFR does exist and that it is pertinent for other institutes to work in unanimity to benchmark their language proficiency programs towards the CEFR so that the level of standards of the English Language in Malaysian universities are acceptable and further credible.

ACKNOWLEDGEMENT

The authors would like to thank the Centre for the Advancement of Language Competence, Universiti Putra Malaysia, and this study's participants for their cooperation.

REFERENCES

Abdullah, M. H., Wah, H. S., & Eng, W. B. (2015). Delivering the ELEx package: An English language experience approach for developing undergraduates' language proficiency. *Journal of Language and Communication, 2*(2), 283-296.

Ali, Z. M., Ali, F., Radzuan, N. M., Alwi, N. N. M., Abu, N. L., & Kassim, Z. (2018). Contextualising the CEFR: the Universiti Malaysia Pahang English language proficiency writing test. In *11th Annual International Conference of Education, Research and Innovation (ICERI 2018)* (pp. 4892-4902). IATED Publisher.

Alosani, M. S., Al-Dhaafri, H. S., & Yusoff, R. Z. (2016). Mechanism of benchmarking and its impact on organizational performance. *International Journal of Business and Management*, *11*(10), 172-183. https://doi.org/10.5539/ijbm.v11n10p172

American Educational Research Association. (1999). *Standards for educational and psychological testing*. American Educational Research Association.

Bogan, C. E., & English, M. J. (1994). *Benchmarking for best practices: Winning through innovative adaptation.* Mcgraw-Hill.

Bruce, E., & Hamp-Lyons, L. (2015). Opposing tensions of local and international standards for EAP writing programmes: Who are we assessing for? *Journal of English for Academic Purposes, 18*, 65-77. https://doi.org/10.1016/j.jeap.2015.03.003

Cambridge Assessment English. (2016). *Linguaskill.* Cambridge Assessment English. https://www.cambridgeenglish.org/Images/429191-linguaskill-listening-and-reading-trial-report.pdf

Cambridge Assessment English. (2018). *Meet the research team behind Linguaskill.* Cambridge Assessment English. https://www.cambridgeenglish.org/Images/521541-linguaskill-research-team.pdf

Cambridge Assessment English. (2019). *Linguaskill simplify your English language testing.* Cambridge Assessment English. https://www.cambridgeenglish.org/Images/466596-linguaskill-simplify-your-english-language-testing.pdf

Cambridge Assessment English. (2021a). *BULATS*. Cambridge Assessment English. https://www.cambridgeenglish.org/exams-and-tests/bulats/

Cambridge Assessment English. (2021b). *Fitness for purpose*. Cambridge Assessment English. https://www.cambridgeenglish.org/research-and-validation/fitness-for-purpose/#cefr-relationship

Cambridge Assessment English. (2021c). *How results are presented*. Cambridge Assessment English. https://www.cambridgeenglish.org/exams-and-tests/linguaskill/information-about-the-test/how-results-are-presented/

Cambridge Assessment English. (2021d). *Malaysia: Ministry of Higher Education recognises English language exams from Cambridge*. Cambridge Assessment English. https://www.cambridgeenglish.org/news/view/malaysia-ministry-of-higher-education-recognises-cambridge-english-exams/

Cambridge English Language Assessment. (2011). *Using the CEFR: Principles of Good Practice.* University of Cambridge. https://www.cambridgeenglish.org/images/126011-using-cefr-principles-of-good-practice.pdf

Canadian Language Benchmarks. (2012). *English as a second language for adults.* Centre for Canadian Language Benchmarks. https://www.canada.ca/

content/dam/ircc/migration/ircc/english/pdf/pub/language-benchmarks.pdf

Centre for the Advancement of Language Competence. (2013). *ELEx Package (English language experience) for undergraduates* Centre for the Advancement of Language Competence. https://calc.upm.edu.my/dokumen/11203_ELEX_Briefing_BI.pdf

Cheung, H. (2012). Academic perspectives from Taiwan. In M. Byram, & L. Parmenter (Eds.), *The Common European Framework of Reference: The globalisation of language education policy* (pp. 224-230). Multilingual Matters.

Don, Z. M., & Abdullah, M. H. (2019, May 27). What the CEFR is and isn't. *Free Malaysia Today*. https://www.moe.gov.my/en/menumedia/printed-media/newspaper-clippings/what-the-cefr-is-and-isn-t-free-malaysia-today-27-mei-2019

Etiken, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11

Fisher, J. G. (1996). *A managers guide to staff incentives and performance improvements.* Kogan Page Publication.

Foley, J. A. (2019). Issues on assessment using CEFR in the region. *Language Education and Acquisition Research Network Journal, 12*(2), 28-48.

Herman, J. L., & Osmundson, D. R. (2010). *Benchmark assessment for improved learning (AACC Policy Brief).* University of California. https://doi.org/10.1037/e685212011-001

Inguva, M. (2018). Effective use of benchmarking: The context of the centre for preparatory studies in Oman. *Arab World English Journal*, 93-104. https://doi.org/10.24093/awej/MEC1.7

Jaafar, M. S. (2019). *Syarat kompetensi Bahasa Inggeris kepada pelajar antarabangsa* [English language competency requirements for international students]. Malaysian Qualifications Agency. https://www.mqa.gov.my/pv4/document/pekeliling/2019/JPT-Penambahbaikan%20Syarat%20Kompetensi%20BI%20Pelajar%20Antarabangsa%20IPTS%2009Dis.pdf

Jin, Y., Wu, Z., Alderson, C., & Song, W. (2017). Developing the common Chinese Framework of Reference for Languages: Challenges at macro and micropolitical levels. *Language Testing in Asia*, *7*(1), 1-19. https://doi.org/10.1186/s40468-017-0032-5

Keegan, R., & O'Kelly, E. (2012). *Applied benchmarking for competitiveness: A guide for SME owner/ managers.* Oaktree Press.

Kwak, S. K., & Kim, J. H. (2017). Statistical data preparation: Management of missing values and outliers. *Korean Journal of Anesthesiology*, *70*(4), 407-411. https://doi.org/10.4097/kjae.2017.70.4.407

Lim, G. S. (2017). *The Occupational English Test and IELTS: A Benchmarking Report.* Cambridge Assessment English.

Malaysian Qualifications Agency. (2020). *Kompilasi dasar jaminan kuasa pendidikan tinggi (2009-2020): Edisi* ketiga [Compilation of higher education power guarantee policies (2009-2020): 3rd ed.]. https://www.mqa.gov.my/pv4/document/compilationPolicy/KOMPILASI%20DASAR%20JAMINAN%20KUALITI%20PENDIDIKAN%20TINGGI%2020092020.pdf

Malaysian Qualifications Agency. (2020). *Malaysian qualification agency.* Programme Standards: Language. https://www2.mqa.gov.my/qad/v2/garispanduan/2020/PS%20Language%20-%20updates%20Oct%202020.pdf

Markel, S. (2018). The CEFR and English education in Japan. *Journal of Policy Studies*, *56*, 33-38.

Ministry of Education. (2015). *English language education reform in Malaysia: The Roadmap 2015-2025.* Ministry of Education Malaysia.

Murtaza, S. F., & Saleh, N. S. (2018). English language use in Malaysian government and private civil engineering workplaces. *International Journal of Education & Literacy Studies*, *6*(3), 84-91. https://doi.org/10.7575/aiac.ijels.v.6n.3p.84

Mustafa, Z. (2018, July 11). Proficient way to better English. *New Strait Times*. https://www.nst.com.my/education/2018/07/389562/proficient-way-better-english

New Strait Times. (2019, June 27). *We can make CEFR work for us*. *New Straits Times*. https://www.nst.com.my/opinion/letters/2019/06/499490/we-can-make-cefr-work-us

Nwabuko, L. O., Igwe, N. J., Okengwu, M. C., Nwabuko, M. A., & Ekere, O. (2020). Benchmark for partnership in human and material resources provision for adult education programmes in the South-East zone of Nigeria. *Global Journal of Health Science*, *12*(5), 46-56. https://doi.org/10.5539/gjhs.v12n5p46

Parlimen Malaysia. (2019). *Mesyuarat Kedua, Penggal Kedua Parlimen Keempat Belas 2019* [Second Meeting in the Second Term of the 14th Parliamentary]. Parlimen Malaysia. https://www.parlimen.gov.my/files/jindex/pdf/JDN22072019.pdf

Prakash, G. (2019, May 23). Education Ministry says using European benchmark for English-language option teachers. *Malay Mail*. https://www.malaymail.com/news/malaysia/2019/05/23/education-ministry-says-using-european-benchmark-for-english-language-optio/1755690

Read, J. (2019). The influence of the Common European framework of Reference (CEFR) in the Asia Pacific region. *LEARN Journal: Language Education and Acquisition Research Network Journal*, *12*(1), 12-18.

Rethinasamy, S., & Chuah, K. M. (2011). The Malaysia university English test and its use for placement purposes: A predictive validity study. *Electric Journal of Foreign Language Teaching*, *8*(2), 234-245. https://doi.org/10.2139/ssrn.2146007

Rockinson-Szapkiw, A. J. (2012). *Step 4: Selecting and justifying your research design.* http://amandaszapkiw.com/artifacts/resources/tutorials/research-process/Step-4-Selecting-and-Justifying-Your-Research-Design.pdf

Sani, R. (2018, February 7). *Building capacity to increase English Proficiency*. *New Strait Times*. https://www.nst.com.my/education/2018/02/332968/building-capacity-increase-english-

Sani, R. (2020, March 18). *Meeting university requirements*. *New Strait Times*. https://www.nst.com.my/education/2020/03/575837/meeting-university-requirements

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763-1768. https://doi.org/10.1213/ANE.0000000000002864

Sufi, M. K. A., & Stapa, M. (2020). Should the CEFR illustrative scales be localised to Malaysian higher education standards? – A conceptual paper. *International Journal of Academic Research in Business and Social Sciences, 10*(9), 885-897. https://doi.org/10.6007/IJARBSS/v10-i9/7882

Thomas, J. (2019, November 28). *Malaysia: Between education and skills*. *The Asean Post*. https://theaseanpost.com/article/malaysia-between-education-and-skills

UPM. (n.d.). *Specific Program Requirements for Bachelor Program*. Universiti Putra Malaysia.

https://fsktm.upm.edu.my/upload/dokumen/BGAKA1_Specific_Program_Requirements_(Bachelor).pdf

Wu, J. (2012). Policy perspectives from Taiwan. In M. Byram, & L. Parmenter (Eds.), *The Common European Framework of Reference: The globalisation of language education policy* (pp. 213-223). Multilingual Matters.

Xu, J., Brenchley, M., Jones, E., Pinnington, A., Benjamin, T., Knill, K., Seal-Coon, G., Robinson, M., & Geranpayeh, A. (2020). *Linguaskill Building a validity argument for the Speaking test.* Cambridge Assessment English. https://www.cambridgeenglish.org/Images/589637-linguaskill-building-a-validity-argument-for-the-speaking-test.pdf

Zairi, M. (1996). *Benchmarking for best practice: Continuous learning through sustainable.* Heinemann.

Zulkefli, A. S. (2017, November 4). M'sian students, teachers show vast improvement in CEFR as per Cambridge Evaluation Study 2017. *New Straits Times*. https://www.nst.com.my/news/nation/2017/11/299126/msian-students-teachers-show-vast-improvement-cefr-cambridge-evaluation

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Aligning the language criteria of a group oral test to the CEFR: the case of a formal meeting assessment in an English for occupational purposes classroom / Shak, P., & Read, J.**
Source	***Pertanika Journal of Social Science and Humanities*** **Volume 29 Issue 3 (2021) Pages 133–156** **https://doi.org/10.47836/pjssh.29.S3.08** **(Database: ResearchGate)**

Pertanika J. Soc. Sci. & Hum. 29 (S3): 133 - 156 (2021)

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Aligning the Language Criteria of a Group Oral Test to the CEFR: The Case of a Formal Meeting Assessment in an English for Occupational Purposes Classroom

Priscilla Shak[1]* and John Read[2]

[1]Centre for the Promotion of Knowledge and Language Learning, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
[2]School of Cultures, Languages and Linguistics, The University of Auckland, 18 Symonds Street, Auckland Central, Auckland 1010, New Zealand

ABSTRACT

The Malaysian Education Blueprint (MEB) 2015-2025 has set in motion efforts from all stages of education to align programs, courses, and syllabuses to the Common European Framework of Reference (CEFR) benchmark. This exercise has brought on major revamps in all aspects of English language education in the nation. This study will present such an undertaking in a public university in Malaysia and detail how the language criteria for an oral group test of an English for Occupational Purposes course have been aligned to the stipulated CEFR level. The actual assessment task involved groups of four or five students conducting a meeting of their established company. Data for the study came from an analysis of the audio recordings of nine group meetings, along with post-assessment interviews and focus group discussions involving three EOP instructors. Based on the data analysis, this study recommends a revised set of language criteria for the assessment. Furthermore, it demonstrates how an alignment of the scoring criteria with the descriptors of the targeted CEFR scale can be achieved through a systematic comparison of the language functions (LFs) produced in the meeting task to the targeted CEFR descriptor scales. The revised language component for the meeting assessment could help ease instructors' assessment of students interactional skills and allow them to gauge

ARTICLE INFO

Article history:
Received: 16 July 2021
Accepted: 04 October 2021
Published: 30 November 2021

DOI: https://doi.org/10.47836/pjssh.29.S3.08

E-mail addresses:
pshak@ums.edu.my (Priscilla Shak)
ja.read@auckland.ac.nz (John Read)
*Corresponding author

ISSN: 0128-7702
e-ISSN: 2231-8534

better their students' attainment of the skills required in a formal meeting context.

Keywords: Assessment criteria, CEFR descriptor scales, EOP, formal meeting, group oral, language function analysis

INTRODUCTION

The English Language Education Reform prompted recent prominent transformations of Malaysia's English language education landscape due to the implementation the Malaysian Education Blueprint (MEB) 2015-2025. The MEB, launched in 2015, is a reform plan spanning all stages of education from preschool to tertiary levels, which has resulted in the unified alignment of the English curricula of these institutions to the Common European Framework of Reference (CEFR) (Council of Europe, 2001). The CEFR includes specifications of six levels of proficiency, each of which has been adopted in the MEB as the aspirational target for one level of education in Malaysia: A1for preschool, A2 for primary, B1 for secondary, B2 for post-secondary, and B2 to C1 for university (Ministry of Education Malaysia, 2016).

The CEFR originated as a project sponsored by the Council of Europe in the late 20th Century to promote language learning among adults who had completed their compulsory education. However, it has subsequently become influential at all levels of education in Europe and many other countries worldwide (Byram & Parmenter, 2012; Read, 2019). It is often seen primarily as an assessment scale, and it does serve as a point of reference for many standardized international tests, including IELTS, TOEFL, and TOEIC (Don & Abdullah, 2019; Abidin & Jamil, 2015). However, it has a much broader scope than that: there are multiple scales in the framework that "are accompanied by a detailed analysis of communicative contexts, themes, tasks and purposes" and the "CEFR is used in teacher education, the reform of foreign language curricula, the development of teaching materials and for the comparability of qualifications" (Council of Europe, 2020b).

There have been numerous critics of the CEFR, both in general terms (Fulcher, 2004; Hulstijn, 2007) and more specifically about problems in defining the B2 level for university admission in Europe and Australia (Deygers et al., 2018a; Deygers et al., 2018b). In addition, closer to home Foley (2019) has raised concerns about how the use of the CEFR as a benchmark has been implemented in various ASEAN countries, including Malaysia. Nevertheless, applied linguists have recognized the appeal of the framework to policymakers as a means of articulating language education goals according to internationally defined levels of proficiency and as a tool for accountability in education. As McNamara (2014) has pointed out, "the functionality of a universal letter/number system to code the six levels is a key feature of the CEFR, which makes it attractive to administrators and policymakers" (p. 227).

In Malaysia's case, policymakers insist that a form of standardization is required,

especially to align English graduates' language proficiency across universities and as a form of quality control. As such, it is the public higher learning institutions' role to help the Ministry achieve this target. Accordingly, this article aims to investigate how the assessment of a specific course at a Malaysian university can be aligned to the CEFR B2 benchmark.

The EOP Meeting Assessment as a Test Task

The context of the present study is a course in English for Occupational Purposes (EOP) at a Malaysian university. The students undertake a group project to establish a company, and they are assessed based on their language performance in the task of a simulated company meeting. The main objective of the EOP course is to improve the students' employability by enhancing their language skills to secure future employment and communicate effectively in future workplaces. These include interviewing, presentation, and meeting skills. Specifically, this study focuses on the formal meeting assessment of the EOP course, which is detailed in the next section.

A review of the literature reveals that the meeting test task is somewhat unconventional. For example, Shehadeh (2017) pointed out that there are relatively few studies that investigated the use of task-based language testing (TLBT) in the English for Specific Purposes (ESP) realm despite both sharing similar underlying principles, which are "goal-oriented," "has a real outcome" and "reflects real-life language use and language need" (Shehadeh, 2018, p. 1).

When learners are engaged in a task, they actively focus on meaning-making through interaction in the target language (Nunan, 1989). At the same time, tasks naturally encourage collaboration between learners (Bruton, 2002). In attempting their tasks, learners interact with one another and engage in collaborative efforts to complete the task assigned as there is a real need to do so for mutual benefits (Nakatsuhara, 2013; Shak, 2014; Shak, 2016; Taylor 1983). Therefore, tasks enable language learners to function in "extended, realistic discourse" and help them learn how to use language appropriately for real communicative purposes (Taylor, 1983, p. 70). According to Skehan (1998), managing tasks engages the "naturalistic acquisitional mechanism" that helps learners to develop language skills (p. 95).

For an assessment task to be authentic, it should "parallel those in the real world" (Messick, 1996, p. 3). It means that a task should simulate the target context as closely as possible. Ellis (2003) also highlighted the need for task-based assessment to represent "real-world" behavior and activities (p. 285). In an earlier study undertaken by the first author to investigate the learners' perception of a task-based group project work related to the current study, it was found that the participants viewed the tasks assigned as comparable to a real-world task (Shak, 2014). In addition, for a test task to be useful, it should be informed by the

real-world language use domain (Bachman & Palmer, 1996). Finally, these authors discussed the notion of 'interactiveness,' which refers to the match between the abilities engaged by the test task and those that learners require in the target language use (TLU) context. Following Bachman and Palmer's visual representation, the TLU domains and tasks for this study are presented in Figure 1.

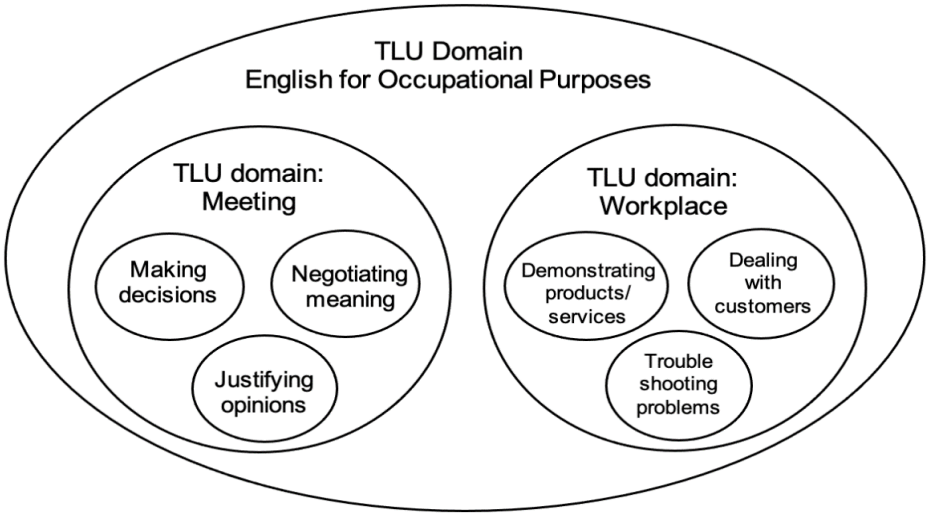


Figure 1. English for Occupational Purposes TLU domain and TLU tasks

As illustrated in Figure 1, the tasks in the TLU domain that apply to the EOP meeting require the test takers to make decisions, negotiate meaning and justify opinions. These functions are among those that are necessary for the successful completion of the meeting assessment task.

Previous studies have highlighted the central role of discourse analysis in offering insights into the nature of interactions in various testing contexts (McNamara et al., 2002; Nakatsuhara, 2013; van Batenburg et al., 2018; Woodward-Kron & Elder 2015). In addition, researchers studying institutional talk have identified formal meeting talk as a genre distinct from other institutional discourse and ordinary conversation (Angouri & Marra, 2010; Asmuß, 2013; Asmuß & Svennevig, 2009; Drew & Heritage, 1992; Svennevig, 2012a; Svennevig, 2012b). Therefore, assessments focusing on this genre should concentrate on its distinctive characteristics and the acquisition of relevant skills to perform the meeting tasks. The appropriate tool for this purpose is Language Function Analysis, which is discussed further in the Data Analysis section below.

The Present Study

The main objective of this study, which is part of a larger-scale project, is to recommend a revised marking scheme for the meeting assessment of the EOP course offered by a Language Centre in a public university in Malaysia. The paper focuses on the alignment of the assessment criteria to the stipulated CEFR B2 level. As such, the paper addresses the following two research questions:

1. What problems did the EOP instructors face when using the existing marking scheme to assess their students' interactional competence?

2. How can the existing marking scheme be revised to align with the CEFR B2 level?

Two sets of qualitative data were obtained from the EOP instructors to address the first research question: individual interviews after the assessment and a Focus Group Discussion (FGD). The synthesized data provided specific details regarding the problems faced by the instructors when assigning marks to their students and their thoughts on the alignment to the CEFR level. For the second research question, results from a Language Function Analysis (LFA) performed on audio recordings of the meeting assessment task were compared to the benchmarked CEFR B2 level descriptor scales for formal discussion (meetings), and recommendations were made based on the findings. The result is a recommended revised version for the language component of the meeting assessment marking scheme.

The EOP Meeting Assessment

The main purpose of the EOP meeting assessment was to evaluate whether the students had acquired the language skills needed to communicate successfully in a meeting setting. In addition, students were tested on their abilities to use language in a formal context and handle such workplace demands in the future. Based on their group project and the roles or positions, each of the students participated in a meeting assessment following a pre-agreed agenda for their group's meeting. The students' main task was to resolve their agenda items to their meeting objective(s). While performing the different roles assigned to them for the meeting test task, students were expected to utilize various language functions such as agreeing, clarifying, suggesting, justifying, negotiating, reciprocating, and interrupting to resolve their agenda items.

The assessment of the meeting task was guided by a marking scheme that contained a list of 16 Likert-type scale items. In accordance with the task-based nature of the EOP group project, the marking criteria focused on the abilities of the students to undertake the meeting task. The evaluation form covered three main components: content and organization (30 marks), presence (20 marks), and delivery, language, and grammar (30 marks). Table 1 lists the items for each of the components. Each item was graded according to a scale of one (very poor) to five (excellent), and each student was assigned individual marks.

While the study was being conducted, the center reviewed all of its English courses

Table 1

EOP meeting assessment's marking criteria

Content and organisation (30%)	Quality of ideas or contents presented in the meeting Sufficient support for ideas Active contribution in the discussion Organized and clear presentation of ideas Perform role assigned effectively Adhere to correct meeting procedures
Presence (20%)	Physical appearance, neatness, and grooming Posture, gestures, mannerism, and movement Eye contact and rapport with group members Listens attentively and shows respect when others are speaking
Delivery, language and grammar (30%)	Enthusiasm and vocal variation (freedom from monotone) Preparation and knowledge of materials (confident and organized) Vocabulary and use of appropriate words (meeting terminologies) Freedom from distracting "uh"s and "like"s Pronunciation, enunciation, audibility, and clarity Grammar

to align them to the Common European Framework of Reference (CEFR) to implement the nationwide English Language Education Roadmap standardization process under the Malaysian Education Blueprint (MEB). As mentioned in the Introduction, part of the MEB requirements is for all English courses in public universities across Malaysia to be aligned to the CEFR's B2 or C1 levels. Given this, the English Language Unit of the Centre determined that the EOP course would be aligned to the CEFR B2 level. This alignment meant that the EOP course would need to produce language learners capable of demonstrating a B2 level of proficiency. As such, it is important that the course assessments could determine whether the learners can perform at this level. Due to this, the assessment criteria of the course would need to be revised according to this benchmark so that an accurate assessment of the learners' proficiency can be correctly mapped to the targeted level.

MATERIALS AND METHODS

The formal meeting assessment involved groups of four or five students. Based on a meeting agenda prepared by the students in advance, each group member was assigned an agenda item based on their role in the project. It provided an information gap as each student had information not available to the others. Following formal meeting conventions, a chairperson was appointed for each group to lead the meeting. Each group was given between 20 to 25 minutes to complete the task. In total, nine meeting groups were audio-recorded.

Each test-taker was awarded individual marks based on the three main rating criteria: a). content and organizations, b). presence, and c). delivery, language, and grammar (Table 1). This paper will focus on the third criterion, the delivery, language, and grammar component.

Participants

In total, 42 second-year undergraduates taking the EOP course and three full-time EOP instructors participated in the study. The student participants had scored Band 1 or 2 in the Malaysian University English Test (MUET), which is a prerequisite for university entrants. The instructor participants recruited the student participants (30 females and 12 males) from their respective classes. Each instructor recruited three groups from their classes. All the instructors were experienced in teaching the EOP course.

Procedures

Each meeting assessment session was attended by the instructor (as evaluator), one group of students (as test-takers), and the first author (as non-participant observer). All the assessment sessions were audio-recorded, as it is less intrusive than video recording for data collection during an assessment event. All the audio files were downloaded into the NVivo 12 software and transcribed orthographically using the transcribe feature of the software. In total, nine transcripts were obtained and analyzed.

All the instructors' post-assessment interview sessions were conducted the week after the meeting assessments. For the post-assessment interviews, a set of semi-structured questions was utilized (Appendix A). Questions relevant to this part of the study included the instructors' feedback regarding their students' performance and their difficulties assigning marks. In total, 136 minutes of recorded data were obtained. In addition, all instructor participants attended a focus group discussion (FGD) as a follow-up to their post-assessment interviews. The FGD was conducted to obtain collective input from the instructors to identify similar issues faced in assigning marks and discuss possible solutions to the problems faced. The FGD lasted for approximately 1 hr 48 min. Appendix B shows the FGD questions.

Data Analysis

The Language Function Analysis (LFA) procedures reported here are situated within a larger project focusing on using group oral assessments in the EOP classroom. For the LFA, both the audio recording and verbatim transcriptions were used concurrently. Therefore, it was necessary to identify the language functions (LFs) that required extensive re-listening and re-reading, and contextual information was essential. The O'Sullivan et al. (2002) Observation Checklist was utilized as an initial operational coding guide (Table 2) to ensure systematic coding of the LFs. Although developed for "real time" use in the Cambridge Main Suite examination paired

speaking test, the successful application of O'Sullivan et al. (2002) checklist was also reported in other studies of oral group tests (Brooks, 2003; Nakatsuhara, 2013).

To ensure that the LFs were coded reliably, the first author and a second coder specializing in English language testing coded all nine transcripts. In instances where there was coding disagreement, specifically those associated with codes where the kappa values were below 0.4, indicating less to a fair agreement (Fleiss et al., 2003; Landis & Koch, 1977; Sim & Wright, 2005; Vierra & Garrett, 2005), the items were further examined and discussed. Upon reaching a final consensus, the kappa values for these items were recalculated. The overall Cohen's kappa value for all of the codes for all the sources is 0.94. Thus, it indicates a high level of inter-coder reliability. In addition, for all codes, average kappa values between 0.71 to 1.0 were obtained.

For the instructors' post-assessment interviews and the focus group discussion (FGD), the audio files were transcribed verbatim orthographically in Word document file format (.docx). The transcripts were then uploaded to NVivo and prepared for coding. Several rounds of close and repeated reading were done before the data were segmented and subjected to thematic analysis coding, allowing researchers to focus on the content highlighted by the participants (Zacharias, 2012). Fereday and Muir-Cochrane (2006) refer to this as "a form of pattern recognition within the data" (p. 82), thus enabling the authors to focus on the specific theme of interest. After the initial coding, the codes and categories were further refined for final data coding before the data was reported.

For the instructors' post-assessment interviews, the themes were coded under two main categories. The first category coded was the challenges in group discussion assessment, which was further sub-coded into i) the scripted discussion; ii) quantity versus quality; iii) role assignment; iv) personality and v) proficiency. The second category coded focused on the challenges posed by the marking criteria. Similarly, for the FGD, the two main categories identified in the post-assessment interviews were used in the NVivo coding. The sub-themes coded under the theme of the challenges in group discussion assessment were i) the scripted discussion, ii) role assignment, iii) monopoly of talk, and iv) proficiency.

Meanwhile, the sub-themes coded under the theme of the challenges in group discussion assessment were i) generic language component, ii) group collaboration, and iii) interpretation of the assessment items. For this study, codes related to the language component of the marking criteria were highlighted in the results section. Data obtained from the post-assessment interviews and the FGD were instrumental in providing the writers with the directions in which the revised assessment criteria should take; most importantly, they need to move towards a more CEFR-aligned format.

RESULTS

Table 2 presents the range of language functions and corresponding percentage of

test-takers use. Additional LFs not found in the original checklist (O'Sullivan et al., 2002) are shown in bold italic typeface. For example, eight additional LFs under *Interactional* functions were identified, while four additional functions under the *Managing interaction* functions were found.

Table 2

The percentage of test-takers for each of the language functions used

Informational functions	%	Interactional functions	%	Managing interaction	%
Expressing opinions	90.5	Asking for opinions	61.9	Reciprocating	42.91
Providing information	83.3	***Asking for confirmation***	59.5	***Nominating***	33.3
Elaborating	76.2	***Confirming***	59.5	***Concluding***	26.2
Justifying opinions	71.4	***Commenting***	54.8	Changing	23.8
Suggesting	66.7	Agreeing	54.8	***Interrupting***	21.4
Describing	31.0	Negotiating meaning	52.4	Deciding	19.0
Staging	14.3	Asking for information	50.0	***Prompting***	4.8
Speculating	14.3	***Acknowledging***	47.6	Initiating	4.3
Summarizing	14.3	***Instructing***	33.3		
Comparing	7.1	***Assisting***	33.3		
Expressing preferences	4.8	***Assuming responsibility***	26.2		
		Modifying	16.6		
		Disagreeing	9.5		
		Granting permission	9.5		

*Additional LFs in ***bold italics*** typeface

As can be seen in Table 2, the meeting assessment elicited the highest number of *Interactional* functions (14 LFs), followed by *Informational* functions (11 LFs) and *Managing Interaction* Functions (8 LFs). It demonstrated the propensity of the meeting test task to elicit the desired functions, which in turn indicated the overall effectiveness of the group oral in prompting interaction among the meeting participants. Thus, it

can be regarded as validating the use of the task to assess the test-takers interactional competence.

Apart from that, the additional LFs identified under the *Interactional* and *Managing Interaction* functions were also unique to the test task, which exemplifies how a specific-purpose assessment task could elicit LFs distinct from other types of group interaction. As presented in this section, identifying the LFs elicited from the test task is crucial in recommending a revised language component for the meeting assessment. It will be addressed further in the Discussion section.

The Instructors' Perspectives

This section presents the data collected from the three EOP instructors' post-assessment interview and focus group discussion (FGD) sessions. It primarily discusses the instructors' concerns regarding their difficulties in evaluating their students' interactional skills and assigning student marks. The instructors' post-assessment interviews were necessary to gain their feedback based on their assessed groups and their personal opinions regarding the assessment task. Meanwhile, the FGD was utilized to obtain collective input regarding what the instructors recognized were the main assessment issues regarding the use of the meeting test task. It was especially useful to gauge their views on what needed to be done to improve the meeting assessment further. The results in this section are based on the synthesized findings.

As the meeting discussion was individually assessed, Instructor 2 expressed that some students did not "care about other people" but focused only on speaking during their turns. As such, interaction and input to each other's topics were minimal, and the desired scaffolding did not occur. These test-takers, it seemed, focused only on presenting their ideas, and, as soon as they had voiced their opinions, they ceased to contribute. "When they're not speaking, you know that they're not in the meeting already… Only doing their part, and that's it", said Instructor 2. Although she observed such behavior, Instructor 2 could not penalize her students as such criteria were not stipulated in the marking scheme. Nevertheless, it was an issue for Instructor 2 as she could not adequately assess her students' interactional skills.

Since the meeting assessment was meant to gauge the test-takers abilities to engage in group interaction, they needed to be involved in the co-construction of the interaction rather than merely presenting their ideas. Therefore, the existing marking criteria that focus on language and grammar components are not particularly relevant for assessing the test-takers interactional abilities. For example, one component focused on vocabulary use, specifically meeting terminologies and useful meeting expressions, but that did not cover the test-takers abilities to use such expressions to co-construct the discussion by continuing, elaborating, negotiating and sustaining the topics being considered.

Both Instructor 1 and Instructor 2 agreed that aligning the existing marking scheme to the CEFR would help improve the validity of the marking scheme in assessing the test-takers interactional skills more effectively and fairly. Instructor 1 believed that the test-takers language abilities could be better gauged if they were assessed based on more specific criteria and "not just by performing [the meeting task]." It implies that the test-takers performance should not be judged solely based on their language abilities to complete their own assigned role but also the means through which they collaborated with the others to accomplish the joint task.

Instructor 2 stressed the need to assess both language and meeting management skills as "they are inter-related. Because if you are able to conduct the meeting, definitely, you have a certain degree of language ability in order to carry out all the procedures, convey ideas clearly and understand others." Hence, in her opinion, the assessment criteria should take these aspects into account. As East (2016) has argued, although to a certain extent, task completion is dependent on linguistic abilities, it may not be a sufficient criterion to assess proficiency in this specific context, where proficiency also involves the ability to engage and interact with each other's thoughts and opinions in order to reach a consensus.

For Instructor 3, the existing marking scheme did not pose any problems for her. She typically adhered to it fairly strictly and would award marks based on the criteria stipulated. Hence, she did not assess components absent from the marking scheme. Interestingly, this was an aspect that she did not realize and only became aware of when attending the FGD. It illustrates how relevant interactional skills might have been neglected in these oral assessments as the focus was just on the linguistic aspects of the test-takers abilities. Nevertheless, Instructor 3 agreed that alignment to the CEFR would entail some revisions to the existing language criteria and believed this move would be more positive.

Overall, although all the instructors agreed that the existing marking scheme allowed them to gauge the competencies required to perform the meeting task and could provide information regarding the test-takers abilities to participate in the discussions, the criteria lacked focus on the use of specific language functions, especially those associated with the group interaction in a meeting. This aspect could be improved with alignment to the relevant CEFR scale.

As the study was being undertaken when the alignment of the EOP course to the CEFR had been proposed in line with the Ministry's standardization exercise, there was increased awareness on the instructors of the need to comply with this requirement. As a result, both Instructor 1 and Instructor 2 could pinpoint the specific table for the Formal discussion (Meetings) scale in the CEFR. Table 3 shows the illustrative descriptors for spoken interaction in that context.

Table 3

CEFR's formal discussion (meetings) illustrative descriptors scale (Council of Europe, 2020a, p.78)

	Formal discussion (Meetings)
C2	Can hold their own in a formal discussion of complex issues, putting an articulate and persuasive argument at no disadvantage to other participants. Can advise on/handle complex, delicate, or contentious issues, provided they have the necessary specialized knowledge. Can deal with hostile questioning confidently, hold on to the turn and diplomatically rebut counter-arguments.
C1	Can easily keep up with the debate, even on abstract, complex, unfamiliar topics. Can argue a formal position convincingly, responding to questions and comments and answering complex lines of counter-argument fluently, spontaneously, and appropriately. Can restate, evaluate and challenge contributions from other participants about matters within their academic or professional competence. Can make critical remarks or express disagreement diplomatically. Can follow up questions by probing for more detail and can reformulate questions if these are misunderstood.
B2	Can keep up with an animated discussion, accurately identifying arguments supporting and opposing points of view. Can use appropriate technical terminology when discussing their area of specialization with other specialists. Can express their ideas and opinions with precision and present and respond to complex lines of argument convincingly. Can participate actively in routine and non-routine formal discussion. Can follow the discussion on matters related to their field, understand in detail the points given prominence. Can contribute, account for, and sustain their opinion, evaluate alternative proposals and make and respond to hypotheses.
B1	Can follow much of what is said related to their field, provided interlocutors avoid very idiomatic usage and articulate clearly. Can put over a point of view clearly, but has difficulty engaging in debate. Can take part in a routine formal discussion of familiar subjects clearly articulated in the standard form of the language, or a familiar variety that involves exchanging factual information, receiving instructions, or discussing solutions to practical problems. Can follow argumentation and discussion on a familiar or predictable topic, provided the points are made in relatively simple language and/or repeated, and opportunity is given for clarification.

Table 3 *(Continued)*

	Formal discussion (Meetings)
A2	Can generally follow changes of a topic in formal discussion related to their field, which is conducted slowly and clearly. Can exchange relevant information and give their opinion on practical problems when asked directly, provided they receive some help with formulation and can ask for repetition of key points if necessary. Can express what they think when addressed directly in a formal meeting, provided they can ask for repetition of key points if necessary.
A1	No descriptors available
Pre-A1	No descriptors available

DISCUSSION

As the authors were made aware of the need for the EOP course to align to the CEFR B2 benchmark, careful consideration was given to meeting this requirement. Hence, in making recommendations for improvement, the authors decided to incorporate the relevant CEFR scale for formal discussion and meetings into the assessment scheme to illustrate what the test-takers should do at the B2 level. However, it has to be pointed at this juncture that a higher number of the LFs produced by the test-takers corresponded more closely to the descriptors below the dividing line after the second statement in the B2 level descriptors. It indicated that the test-takers were likely to be at the lower range of B2 performance, which was to be expected as it represented a more realistic target for Malaysian students with MUET Band 1 and 2 scores. Nevertheless, there were also instances where the more proficient test-takers could produce LFs that reflected higher-level descriptors. Therefore, it indicated that the meeting assessment task was able to elicit LFs beyond B2 level performance. However, as the EOP course has been benchmarked at the B2 level, the revisions were made based on comparison to this level of descriptors.

In order to incorporate elements of the CEFR descriptors into revised language criteria for the meeting test, the authors examined the LFs generated from the meeting assessment, specifically those that yielded higher percentages of test-taker use (ranging from 50% to 90.5%) and compared these to the CEFR descriptors. Table 4 illustrates this comparison.

After examining the corresponding LFs to the CEFR descriptors, the recommended revisions for the language and delivery components were put forth and presented in Table 5 to replace the existing delivery, language, and grammar components of the meeting assessment (Table 1).

Table 4

CEFR B2 descriptors scale for formal discussion and meeting and the corresponding language functions

Level	Descriptors scale for formal discussion and meetings	Corresponding Language Functions
B2	Can keep up with animated discussion, accurately identifying arguments supporting and opposing points of view. Can express his/her ideas and opinion with precision, present and respond to complex lines of arguments convincingly. Can participate actively in routine and non-routine formal discussion. Can follow the discussion on matters related to his/her field, understand in detain the points given prominence by the speaker. Can contribute, account for, and sustain his/her opinion, evaluate alternative proposals and make and respond to a hypothesis.	(Dis)agreeing Supporting Negotiating meaning Expressing/Asking for opinions Justifying opinions Suggesting Asking for confirmation/ Confirming Elaborating Commenting Asking for/Providing information

Table 5

Recommended revisions for the language and delivery components

Language and Delivery
Can present with confidence and enthusiasm (vocal variation, e.g., freedom from monotone).
Can use accurate vocabulary and grammar (appropriate meeting terminologies and sentence structure).
Can speak with correct pronunciation (enunciation, audibility, and clarity).
Can speak fluently (free from lengthy/frequent pauses and distracting fillers, independent of notes).
Can contribute ideas and suggest alternatives.
Can respond to ideas by (dis)agreeing, commenting, confirming, and negotiating meaning.
Can sustain discussion by elaborating, supporting, and justifying opinions and/or arguments.

As presented in Table 5, the recommended version incorporates 'can do' statements, characteristic of the CEFR. These statements correspond to the B2 level of the CEFR's formal discussions and meetings scale. In this revised version, four of the descriptors from the original CEFR list are integrated. Where broader behavioral features are indicated in the CEFR, they are represented more explicitly in the revised version of the marking scheme. For example, at the CEFR B2 level, students 'can keep up with animated discussion, accurately identifying arguments supporting and opposing points of view' (Table 4). These skills are represented in the revised version's abilities to 'present with confidence and enthusiasm' and sustain the discussion by 'elaborating, supporting, and justifying opinions and/or arguments.' It is also worth pointing out that the recommended version does not emphasize accuracy in grammar and pronunciation. Not because these are not important but mainly because these features could be better tested through the other types of assessment that the test-takers have to perform in the EOP course, such as the test, presentation, proposal, and portfolio tasks. As such, the assessment of the meeting task should concentrate more on the abilities of the test-takers to perform interactional functions in such a setting. As Galaczi and Taylor (2018) have recommended, CEFR descriptors should be further refined to meet stakeholder needs. In the case of this study, one of the considerations for the revision of the assessment criteria is the concept of test localization, which "stipulates that for a test to be valid, its design and development must take into consideration the population, context, and the domain in which the test is used" (Abidin & Jamil, 2015, p. 1).

This study has utilized the qualitative bottom-up approach to gain insights into the language produced by the test takers to substantiate the recommendations for a revised marking scheme. At the same time, the post-assessment interviews and FGD with the instructors revealed concerns about the marking scheme and the need to align it with the benchmarked CEFR level, which has illuminated aspects that required improvement.

One of the main aims of language proficiency testing in ESP is to assess test-takers performance based on a simulated setting to predict their capacities to tackle such real-world demands in the future (Basturkmen & Elder, 2004; Douglas, 2000; Woodward-Kron & Elder, 2015). The results of the LFA indicated that, in addition to the LFs found in the assessment of dyads, the group format could generate a wider range of LFs, which lends support to its use for assessing the interactional competence of language learners. Most importantly, the group meeting task could generate language functions that reflect those in natural workplace settings. It is an important aspect of the EOP course as students are exposed to realistic and meaningful interaction. When "the language learners are functioning in the target language in situations similar to the ones they experience every day, they may start internalising English and their motivation may increase" (İlin, 2014, p. 2).

As illustrated in this study, identifying LFs in a meeting setting is instrumental in informing the design of revised marking criteria for the language component of the meeting evaluation form. The recommended language descriptors make it easier for the instructors to evaluate a student's performance. However, as the stakeholders require, they align with the CEFR's formal discussion and meeting descriptors. Despite skeptics' claims, the CEFR can serve as a rich resource for rating scale development and adapted to various testing conditions (Deygers & Van Gorp, 2015; North, 2014; Weir, 2005a; Weir, 2005b; Abidin & Jamil, 2015).

CONCLUSION

This study has illustrated how the language criteria of an EOP meeting assessment can be aligned to the CEFR by demonstrating in detail the steps involved in the alignment process. Qualitative data obtained from the EOP instructors' post-assessment interviews and FGD were utilized to identify the specific issues they faced while assigning students marks to help determine areas requiring revision. In addition, the LFA provided empirical evidence of the LFs elicited by the task. It enabled them to be compared to the CEFR descriptors, which led to the recommended revised criteria.

The methodological implication of the study is that data from the corpus of students' meeting assessment events are a rich and viable resource for the alignment of assessment criteria with the objective and learning outcome of a course. By examining in-depth what was produced by the test-takers in an actual assessment event and comparing this to the targeted performance descriptors, CEFR-compliant assessment criteria could be devised to ensure that the assessment method correlates with the desired level of performance. In this case, the LFA was useful to help gauge the effectiveness of the meeting test task to elicit the desired language output and served as an effective method to map the elicited output to the CEFR's B2 level descriptors for formal meetings and discussions. The result was the recommended CEFR-aligned marking criteria for the language component as presented earlier.

The limitation of this study is that data were collected from just a small number of instructors. Despite this, feedback from these experienced instructors indicated that they were aware of the shortcomings of the assessment scheme utilized then. Another shortcoming is that the trial of the revised assessment has yet to be undertaken. Nevertheless, the proposed revised criteria presentation to the three instructor participants and preliminary discussions indicated that the recommended version would likely ease the challenges of grading the students. In addition, the resulting assessment marks would better reflect the students' interactional abilities. Another limitation concerns the focus of the recommended revisions based on the B2 level descriptors. It has to be acknowledged that it is possible for other lower (B1 below) or higher levels (C1 and C2) LFs can manifest during the formal

meeting assessment. Nevertheless, as highlighted earlier, since the Centre has determined the EOP course to be aligned to the B2 level, the main focus of the revisions in this study was placed on this level's descriptors. Nonetheless, similar processes may be adopted for the other CEFR level descriptors in other contexts based on the steps undertaken in aligning the marking criteria detailed in this study.

An area worth exploring in the future is the trialing and implementing this revised marking scheme to gauge its effectiveness and a further detailed examination of other assessment criteria to enhance further the overall assessment of the students' interactional abilities.

ACKNOWLEDGEMENTS

This paper is based on a completed doctoral study with financial support from the Ministry of Higher Education Malaysia and Universiti Malaysia Sabah.

REFERENCES

Abidin, S. A. Z., & Jamil, A. (2015). Toward an English proficiency test for postgraduates in Malaysia. *SAGE Open, 5*(3), 1-10. https://doi.org/10.1177/2158244015597725

Angouri, J., & Marra, M. (2010). Corporate meetings as genre: A study of the role of chair in corporate meeting talk. *Text and Talk, 30*(6), 615-363. http://doi.org/10.1515/TEXT.2010.030

Asmuß, B. (2013). Conversation analysis and meetings. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 2006-2008). http://doi.org/10.1002/978140598431.wbeal0210

Asmuß, B., & Svennevig, J. (2009). Meeting talk. *Journal of Business Communication, 46*(1), 3-22. http://doi.org/10.1177/0021943608326761

Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.

Basturkmen, H., & Elder, C. (2004). The practice of LSP. In A. Davies & C. Elder (Eds.), *Handbook of applied linguistics* (pp. 672-694*)*. Blackwell. https://doi.org/10.1002/9780470757000.ch27

Brooks, L. (2003). Converting an observation checklist for use with the IELTS speaking test. *Cambridge ESOL Research Notes, 11*, 20-21.

Bruton, A. (2002). From tasking purposes to purposing task. *ELT Journal, 56*(3), 280-288. https://doi.org/10.1093/elt/56.3.280

Byram, M., & Parmenter, L. (Eds.) (2012). *The Common European Framework of reference: The globalisation of language education policy*. Multilingual Matters. https://doi.org/10.21832/9781847697318

Council of Europe (2001). *Common European Framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

Council of Europe. (2020a) *Common European Framework of reference for languages: Learning, teaching, assessment - Companion volume.* Council of Europe Publishing.

Council of Europe (2020b). *The Common European Framework of reference for languages*. Council of Europe Publishing.

Deygers, B., & Van Gorp, K. (2015). Determining the scoring validity of a co-constructed CEFR-based rating scale. *Language Testing, 32*(4), 521-540. https://doi.org/10.1177/0265532215575626

Deygers, B., Carlsen, C. H., Saville, N., & Van Gorp, K. (2018a). The use of the CEFR in higher education: A brief introduction to this special

issue. *Language Assessment Quarterly, 15*(1), 1-2. https://doi.org/10.1080/15434303.2017.1421957

Deygers, B., Carlsen, C. H., Saville, N., & Van Gorp, K. (Eds.) (2018b). Special issue: Language tests for academic enrolment and the CEFR. *Language Assessment Quarterly, 15*(1), 1-108. https://doi.org/10.1080/15434303.2017.1421957

Don, Z. M., & Abdullah, M. H. (2019, May 17). What the CEFR is and isn't. *Free Malaysia Today*. https://www.freemalaysiatoday.com/category/opinion/2019/05/27/what-the-cefr-is-and-isnt/

Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge University Press.

Drew, P., & Heritage, J. (1992). Analyzing talk at work: An introduction. In P. Drew, & J. Heritage (Eds.), *Talk at work: Interaction in institutional settings* (pp. 3-65). Cambridge University Press. https://doi.org/10.1075/fol.1.2.08ade

East, M. (2016). *Assessing foreign language students' spoken proficiency: Stakeholder perspectives on assessment innovation*. Springer.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods, 5*(1), 80-92. https://doi.org/10.1177/160940690600500107

Fleiss, J. L., Levin, B. A., & Paik, M. C. (2003). *Statistical methods for rates and proportions* (3rd ed.). Wiley-Interscience.

Foley, J. (2019). Issues on assessment using CEFR in the Region. *LEARN Journal: Language Education and Acquisition Research Network, 12*(2), 28-28.

Fulcher, G. (2004). Deluded by artifices? The Common European Framework and harmonization. *Language Assessment Quarterly, 1,* 253-266. https://doi.org/10.1207/s15434311laq0104_4

Galaczi, E., & Taylor, L. (2018). Interactional competence: Conceptualisations, operationalisations, and outstanding questions. *Language Assessment Quarterly, 15*(3), 218-236. https://doi.org/10.1080/15434303.2018.1453816

Hulstijn, J. H. (2007). The shaky ground beneath the CEFR: Quantitative and qualitative dimensions of language proficiency. *Modern Language Journal, 91*, 663-667. https://doi.org/10.1111/j.1540-4781.2007.00627_5.x

İlin, G. (2014). Student-teacher judgements on Common European Framework: Efficacy, feasibility and reality. *Journal of Language and Literature Education, 9*, 8-19. https://doi.org/10.12973/jlle.11.221

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. https://doi.org/10.2307/2529310

McNamara, T. (2014). 30 years on–evolution or revolution? *Language Assessment Quarterly, 11*, 226-232. https://doi.org/10.1080/15434303.2014.895830

McNamara, T., Hill, K., & May L. (2002). Discourse and assessment. *Annual Review of Applied Linguistics, 22*, 221-242. https://doi.org/10.1017/s0267190502000120

Messick, S. (1996). Validity and washback in language testing. *ETS Research Report series, 1996*(1), i-18. https://doi.org/10.1002/j.2333-8504.1996.tb01695.x

Ministry of Education Malaysia. (2016, April 27). Executive summary: *Malaysia education blueprint 2015-2025 (Higher Education)*. Kementerian Pendidikan Malaysia.

Nakatsuhara, F. (2013). *The co-construction of conversation in group oral test*. *Language Testing and Evaluation, 30*. Peter Lang.

North, B. (2014). *The CEFR in practice*. *English Profile Studies, 4*. Cambridge University Press.

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.

O'Sullivan, B., Weir, C. J. & Saville, N. (2002). Using observation checklists to validate speaking-test tasks. *Language Testing*, *19*(1), 33-56. http://doi.org/10.1191/0265532202lt219oa

Read, J. (2019). The influence of the Common European Framework of Reference (CEFR) in the Asia-Pacific region. *LEARN Journal, 12*(1), 12-18.

Shak, P. (2014). Incorporating task-based group project work in English for Occupational Purposes Course: The instructors' perspectives. *MANU Journal, 21*, 77-97.

Shak, P. (2016). Taken for a ride: Students' coping strategies for free-riding in group work. *Pertanika Journal of Social Science & Humanities, 24*(1), 401-414.

Shak, P. (2019). *Towards a framework for effective group oral assessment in the ESP classroom* [Unpublished Doctoral dissertation]. University of Auckland.

Shehadeh, A. (2017). Foreword: New frontiers in task-based language teaching. In M. Ahmadian & M. Mayo (Eds.). *Recent perspectives on task-based language learning and teaching* (pp. vii-xxi). De Gruyter Mouton. http://doi.org/10.1515/9781501503399-015

Shehadeh, A. (2018). Task-based language assessment. In J. I. Liontas (Ed.), *The TESOL Encyclopedia of English Language Teaching*. (pp. 1-6). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118784235.eelt0379

Sim, J., & Wright, C. C. (2005). The Kappa statistic in reliability studies: Use, interpretation, and sample size requirements. *Physical Therapy, 85*(3), 257-268. https://doi.org/10.1093/ptj/85.3.257

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Svennevig, J. (2012a). Interaction in workplace meetings. *Discourse Studies*, *14*(1), 3-10. http://doi.org/10.1177/1461445611427203

Svennevig, J. (2012b). The agenda as a resource for topic introduction in workplace meetings. *Discourse Studies*, *14*(1), 53-66. http://doi.org/10.1177/1461445611427204

Taylor, B. P. (1983). Teaching ESL: Incorporating a communicative, student-centered component. *TESOL Quarterly, 17*(1), 69-88. https://doi.org/10.2307/3586425

van Batenburg, E. S. L., Oostdam, R. J., van Gelderen, A. J. S., & de Jong, N. H. (2018). Measuring L2 speakers' interactional ability using interactive speech tasks. *Language Testing, 35*(1), 75-100. https://doi.org/10.1177/0265532216679452

Vierra, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: The Kappa statistics. *Family Medicine, 37*(5), 360-363.

Weir, C. (2005a). Limitations of the Common European Framework for developing comparable examinations and tests. *Language Testing, 22*(3) 281-300. https://doi.org/10.1191/0265532205lt309oa

Weir, C. J. (2005b). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.

Woodward-Kron, R., & Elder, C. (2015). A comparative discourse study of simulated clinical roleplays in two assessment contexts: Validating a specific-purpose language test. *Language Testing, 33*(2), 251-270. https://doi.org/10.1177/0265532215607399

Zacharias, N. T. (2012). *Qualitative research methods for second language education: A coursebook*. Cambridge Scholars Publishing.

APPENDICES

Appendix A

Post-assessment interview questions (adapted from Shak, 2019)

1. What do you think of your students' overall performance for the meeting assessment?
   Potential prompts:
   a) Are you happy with the performance of the groups?
   b) Are you happy with the students' performance?

2. For the formal meeting assessment, were there any successful group discussions that stood out?
   Potential prompts:
   a) Why was/were the discussion(s) successful?
   b) What did the students do to make the discussion successful?

3. Did any of the students perform well beyond your expectation of him/her?
   a) Why was the student's/students' performance successful?
   b) How did this affect your marking?

4. During the meeting assessment, were there any students who performed badly?
   a) Why were the students' performance less successful?
   b) What did the students do/fail to do?

5. Do you think the group discussion assessment format is suitable for assessing your students' language skills?

   Follow-ups if YES:
   a) Why?
   b) How?

   Follow-ups if NO:
   a) Why?
   b) What method(s)/format(s) would you suggest instead?

6. In your opinion, is the use of the group discussion assessment fair for the students?

   Follow-ups if YES:
   a) Why?
   b) Please elaborate on why you feel that it is a fair assessment.
   c) What do you do to ensure that the students are assessed fairly in the group assessment?

   Follow-ups if NO:
   a) Why?
   b) Please elaborate on why you feel that it is not a fair assessment.
   c) What do you think can be done to improve the fairness of the group discussion assessment?

7. During their group assessment, the students were assigned different roles. Do you think this will favor some students (i.e., the chairperson of the meeting) while placing the others at a disadvantage?

Follow-ups if YES:

a) Why?
b) How do you think this can be prevented?

Follow-ups if NO:

a) Why?

8. For the group assessment, is there a specific marking scheme that you adhere to? (Refer to marking scheme)
   a) Did you follow the marking scheme strictly when assessing your students? Why? If not, how did you do it?
   b) How did you use the marks sheets? Do you go according to the list of items in the score sheet?
   c) Do you think the marking scheme reflects the objectives of the meeting discussion assessment? How so? If not, how do you think this can be done?
   d) Do you think the marking criteria allow for effective assessment of the specific language skills required to perform the group discussion task?
   e) Do you think that the marking criteria are suitable for assessing the individual language abilities of the students?
   f) Do you think that the marking criteria are fair for all students?
   g) Do you agree with all the items in the marking scheme? Explain.
   h) Did you face any problems while using the marking scheme? Please explain.
   i) Did you have any difficulty assessing all the students within the duration of their group assessment?
   j) How did you ensure that the assessment was done within the timeframe for each of the students?
   k) In your opinion, how can the marking scheme be improved?

9. The course outline specified groups of four students for the group project. In groups where there were more/extra member(s),
   a) How had the extra student affected the assessment process?
   b) How did you manage the assignment of marks in bigger groups?

10. What did you pay attention to when assigning marks to your students? (eg. Language/performance/cooperation)

11. How did/would you assess students who were quiet during the assessment?
    a) Those who are naturally quiet
    b) Those who are weak in the language
    c) Those who cannot get a word in because of other members who manipulate discussion
    d) Those who chose not to contribute when given a chance (the free-rider?)

12. How did/would you assess students who manipulated most of the talk time during the assessment to get a higher score?

13. How did you use your knowledge of your students to help you in assigning their marks?

14. How did you ensure that everyone gets the marks they deserved and that you have marked them fairly?

15. Were your marks set by the end of the assessment? Did you review your marks? How did you do this?

16. What are your suggestions to make the group assessment process more effective?

17. Do you have anything to add?

Appendix B

Focus group discussion questions (adapted from Shak, 2019)

1. What do you think about the topic that has brought us here today (meeting assessment)?

2. I understand that in this Centre, the course chairperson makes most of the decisions about the course design. What are the roles of the other instructors of the course in the decision-making process?
   Items covered:
   - Course design
   - Course assessment
   - Course content

3. In your opinion, what are the major problems in implementing the group discussion assessment format?
   Items covered:
   - Time constraints
   - Numbers of students in a group
   - Students who free-ride (or do not contribute much to the discussion).
   - Students who monopolize the discussion
   - The different personalities
   - The marking scheme
   - The allocation of marks (individual versus group marks)
   - Whether the marks reflect the individual student's language abilities
   - Whether the marks given are generalizable to other settings. (i.e., whether being able to perform well in the group discussion assessment means being able to perform in other oral tasks competitively as well)

4. What do you think can be done to overcome the problems you (the instructors) face?

5. Could you provide any suggestions on how the group discussion assessment process can be improved?
   Items covered:
   - Planning
   - Strategies to ensure fair evaluation of the students
   - Marking scheme/criteria
     - Task versus construct considerations
     - How to ensure that the student's skills can be captured and are reflected in their scores

- How to ensure that the marking sheet is practical for use for the group discussion assessment

6. Do you have anything to add?

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Exploring vocabulary teaching: planning and challenges with the implementation of audio-visual approach in rural East Malaysian primary-level English language classrooms / Yang, B. L. C., Ramanair, J., & Shing, S. R.**
Source	***3L: Language, Linguistics, Literature*** **Volume 29 Issue 4 (2023) Pages 228-243** **https://doi.org/10.17576/3L-2023-2904-15** **(Database: UKM Press)**

3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies
Vol 29(4), December 2023 http://doi.org/10.17576/3L-2023-2904-15

Exploring Vocabulary Teaching: Planning and Challenges with the Implementation of Audio-Visual Approach in Rural East Malaysian Primary-Level English Language Classrooms

BRYAN LIM CHEN YANG
Faculty of Education, Language and Communication
Universiti Malaysia Sarawak, Malaysia
bryanlim1991@gmail.com

JOSEPH RAMANAIR
Faculty of Education, Language and Communication
Universiti Malaysia Sarawak, Malaysia

SOO RUEY SHING
Faculty of Education, Language and Communication
Universiti Malaysia Sarawak, Malaysia

ABSTRACT

Vocabulary is key to effective communication. Previous studies revealed that teachers often overlook vocabulary teaching compared to other language skills such as speaking, listening, grammar, reading, and writing. The integration of technology into the Malaysian education system has raised concerns about how vocabulary lessons are planned and executed in English classrooms, particularly in rural Sarawak schools in Malaysia. Studies in the past found that teachers have been hesitant to use Common European Framework of Reference for Languages (CEFR) aligned resources despite knowing about their availability. This research was conducted to determine the teachers' vocabulary lesson planning and the challenges teachers experienced when using the audio-visual (AV) approach to teach vocabulary in the primary-level English language classroom. Based on the purposive sampling technique, three primary-level English language teachers from one school in a rural area of Bau, Sarawak, participated in this study to provide rich information relevant to the research questions. Data were collected through semi-structured interviews, classroom observation and document analysis, and were analysed thematically based on Braun and Clarke's (2006) thematic analysis method. The findings indicated that teachers referred to CEFR-aligned documents for lesson planning especially when selecting resources and using the framework as a benchmark for practice. The findings indicated that teachers were facing challenges in two primary areas: infrastructure challenges related to information technology and policy challenges related to the CEFR. The results from this study provide practitioners with practical insights into how rural teachers implement the AV approach to teaching vocabulary in their classrooms.

Keywords: Vocabulary teaching; audio-visual approach; primary-level English language classroom; CEFR; rural school

INTRODUCTION

Vocabulary is the building block for effective communication. Insufficient vocabulary can make communication difficult (Richards & Schmidt, 2002; Sovakandan et al., 2017). Therefore, vocabulary is the key to learning the English language, which serves as the foundation for developing other language skills such as reading, writing, listening, and speaking (Mohd. Nor et al., 2015; Susanto, 2017). However, in reality, researchers found that vocabulary teaching has

always been outshined by the attention given to teaching grammar, reading, and writing skills (Kalajahi & Poursahian 2012; Lu, 2017; Maizatulliza & Kiely, 2018).

Teaching English in Malaysia can be challenging, especially for teachers in rural schools in Sarawak where English is not the learners' first language. The differences in culture and background between the teachers and students in these schools often affect the pedagogical decisions made by the teachers (Chan & Kapong, 2021; Ler, 2012; Musa et al., 2012).

According to Kiss and Rimbar (2017), the Malaysian Ministry of Education (MOE) has provided resources like English textbooks to teachers. However, these textbooks are difficult to use in rural Sarawak schools due to their foreign cultural content being incompatible with the local setting. Thus, teaching vocabulary based on the textbook was a challenge for the teachers, as opposed to teaching learners in urban schools, which have a more supportive environment, such as access to more up-to-date textbooks and technologies that can support English learning (Kiss, T., & Rimbar, H., 2021). The integration of technology into the Malaysian education system has raised concerns about how vocabulary lessons are planned and executed in English classrooms, particularly in rural Sarawak schools in Malaysia. The findings from this study are significant in providing practitioners with insights into how rural teachers implement the AV approach to teaching vocabulary in their classrooms.

Traditionally, Scott & Nagy (1997) argue that vocabulary teaching was based on the definition approach. Other researchers find it a passive approach because teachers focused on providing students with word definitions (Zeta et al., 2019). In recent years, technological advancement has contributed to the popularity of the audio-visual (AV) approach among Malaysian English teachers to improve students' language acquisition (Bahagian Pembangunan Kurikulum [BPK], 2017).

Generally, AV is derived from the combination of two words: audio, which refers to "what we can hear," and visual, which refers to "what we can see" (Anas & Zakaria, 2019). In education, teachers often use the AV approach to engage learners through both audio and visual resources (Swaran Singh et al., 2021), which targets both the senses of hearing and sight (Anas & Zakaria, 2019). AV approach was also adopted to increase the learners' experience in the classroom (Swaran Singh et al., 2021). In line with the introduction of the Common European Framework of Reference (CEFR) into our Malaysian education system in 2017, AV resources aligned with the Teacher's Guidebook and Student's Book were provided by MOE Malaysia (BPK, 2017). The CEFR-aligned AV resources were saved in the form of compact discs (CDs) to aid English teachers in lesson planning and practice.

Studies in the past indicated that teachers have been hesitant to use CEFR-aligned resources, such as CDs, despite knowing about their availability (Kiss, T., & Rimbar, H., 2021; Ngu & Azlina, 2019; Utami, 2015). According to Bayuong et al. (2019), even with the availability of AV resources, integrating information technology (IT) in primary schools in Malaysia, especially in rural areas like Sarawak, has proven to be challenging. Past research has found that English teachers from rural schools were facing challenges such as the lack of resources (Kiss, T., & Rimbar, H., 2021; Utami, 2015), poor network coverage (Ngu & Azlina, 2019), lack of equipment and facilities (Wazeema & Kareema, 2017; Yonas et al., 2020), insufficient access to technology (Nawai & Nur Ehsan, 2020) and lack of teacher's training (Aminuddin & Azman, 2017; Azman et al., 2018).

According to Nurul Farehah and Mohd. Sallehudin (2017), the main reason that teachers were faced with these challenges was due to the lack of CEFR experts to develop and manage the CEFR-aligned resources, and also the lack of training for teachers to implement CEFR-aligned

resources in actual classroom practice. Therefore, it is crucial to understand how teachers plan to use the CEFR-aligned resources for vocabulary teaching and the challenges faced in the implementation process.

Recent studies by Alanazi (2019) and Phan (2021) have found that teachers' lesson planning plays a decisive role in the success or failure of a lesson. Effective lesson planning assists teachers to minimise the possible challenges in the classroom. How teachers teach is influenced by their Pedagogical Content Knowledge (PCK) (Shulman, 1987; 2004). Shulman developed PCK in 1986 to bridge teachers' pedagogical and content knowledge that influences teachers' lesson planning, thus contributing to the success or failure of each lesson (Muhammad & Keily, 2018; Philip et al., 2019). This is especially crucial for vocabulary teaching as vocabulary was only presented in the form of a wordlist in the CEFR-aligned documents (BPK, 2017) without any specified instruction provided. Thus, how to teach vocabulary in the classroom depends on the teachers' PCK.

Therefore, the purpose of this study is to examine primary-level English teachers' vocabulary lesson planning through the use of the AV approach in the rural classroom. The study aims to achieve two main objectives. Firstly, it aims to examine teacher planning for using AV to teach vocabulary in primary-level English language classrooms. This leads to the first research question: "How do teachers plan to use AV to teach vocabulary in the English language classroom?" Secondly, it aims to explore the challenges faced by teachers in using the audio-visual approach to teach vocabulary in the primary-level English language classroom. This leads to the second research question: "What are the challenges teachers encounter when using the AV approach to teach vocabulary in a primary-level English language classroom?" The findings from this study will be able to contribute to the limited research on vocabulary teaching through the CEFR-aligned curriculum in the Malaysian context.

METHODOLOGY

RESEARCH DESIGN

This study adopts a qualitative approach to investigate teachers' use of the AV approach to plan vocabulary lessons, as well as uncover the challenges teachers face. The value of the qualitative approach is to provide a richer and deeper understanding of how people perceive actions, events, behaviours, and relationships (Yin, 2011).

Since this is a qualitative study, the findings were not meant to be generalised to the entire population, but to provide a deeper understanding of how teachers adopt the AV approach for vocabulary lesson planning, as well as the challenges that they encounter. This qualitative research involved data collection, analysis, and interpretation via interviews, observations, and document analysis, which will be addressed further in the following section.

CONTEXT OF STUDY

The sample school for this study was a National Primary School known as Sekolah Kebangsaan (SK) in the rural area of Bau, Sarawak. There are 369 students enrolled at this school, which has the largest student population in the Bau district at the time of the study. The school had 31 teachers at the time of data collection, five of them were English language teachers. English was a compulsory subject in the school which adopted the CFER-aligned curriculum. The English

language proficiency of learners in this school was average. The school was equipped with the basic IT infrastructure which needed to be shared among teachers.

SAMPLING

The purposive sampling technique was adopted to select the participants for this study. In qualitative research, purposeful sampling is widely used because it enables the researcher to find and choose a participant who may offer rich information relevant to the research question (Cresswell & Clark, 2011; Patton, 2015). Purposive sampling allows the sample size to be flexible depending on the resources, time available, and the research objectives (Bernard, 2002). Given the purpose of this study, three primary factors for the selection of participants were identified, namely (1) teaching English, (2) at least five years of teaching experience in English, and (3) English options.

Three out of five primary-level English language teachers who fit the selection criteria participated in this study to provide rich data relevant to answering the research questions. The other two teachers did not meet the requirements of having at least five years of experience teaching English, and one was unable to participate owing to an extended leave of absence during the data collection period. The three teachers who fulfilled the selection criteria were labelled as A1, A2 and A3. The demographic information of the participants is shown in the following Table 1.

TABLE 1. Demographic Information of Study Participants

Teacher Code	Option / Major	Subject Taught	Experience in the Teaching Field	Experience in Teaching English
A1	English	English	8 years	8 years
A2	English	English	32 years	32 years
A3	English	English	17 years	17 years

DATA COLLECTION METHOD

For this study, data were collected through three methods, which were semi-structured interview, document analysis, and observation.

The semi-structured in-person interviews with the school's teachers were undertaken for this study. Within a week before and after the classroom observation, an interview session with the corresponding teacher will be conducted. The duration of each interview session was approximately forty minutes. The research question served as the basis for the interview questions. With the interviewee's permission, the audio recordings of the sessions were made, and the researcher took notes throughout the interview. A verbatim transcription was drafted following the interview to look for recurring themes or patterns.

For the document analysis procedure, the document collected was the teachers' lesson plan to support their responses from the interview. A lesson plan was collected from each teacher before the observation session to allow the researcher to review the teacher's planning before the implementation process.

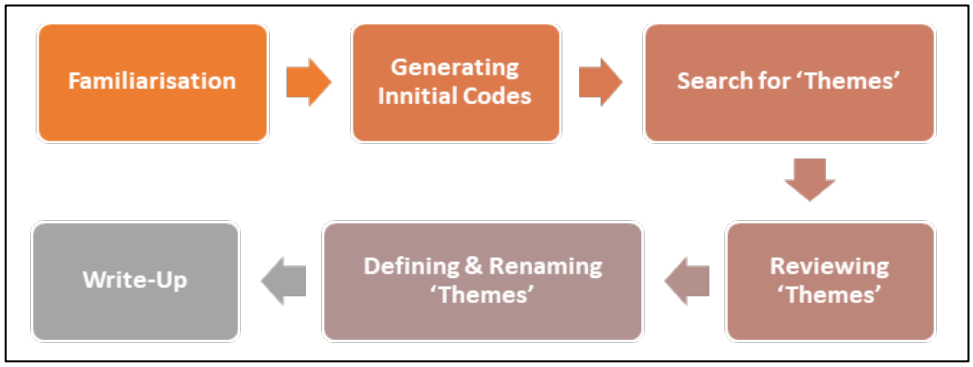
The purpose of the classroom observations was to see if the teacher's lesson was implemented as planned and discover the challenges teachers faced when teaching vocabulary. For every teacher, a sixty-minute observation session was conducted in the classroom. The classroom observation was both audio recorded and written in the observation field-note template with the teacher's permission.

DATA ANALYSIS

For this study, thematic analysis was used as its analytical framework (Braun and Clarke, 2006; Boyatziz, 1998). According to Braun and Clarke (2006), thematic analysis engaged researchers in the process of seeking themes within the collected data, thus using those themes to address the research question.

This study used Braun and Clarke's (2006) thematic analysis method, which consists of six steps, as indicated in Figure 1.

FIGURE 1. Braun & Clarke (2006) Thematic Analysis Method



Before the analysis, data from the interview was transcribed verbatim and thematically examined for this study. Then, by referring to Figure 1, step one began with rereading and getting familiar with the data collected from the interview, observation and document analysis. Step two involved generating a list of initial codes for the data and coding the data. Step three involved sorting the different codes into categories and searching for themes. Step four required the researcher to review the themes. Step five was defining and naming themes. Finally, step six involved a discussion of the emerging themes to answer the research questions.

RESULTS

This study sets out to examine the teachers' planning for vocabulary teaching through the AV approach. The findings were reported in two areas, which are teachers planning to use AV approaches to teach vocabulary and the challenges they encountered.

During the interview, the teachers were asked about their planning, and all of them cited CEFR. Therefore, the reporting of the findings centred on CEFR and the teacher's understanding of it. The two major themes that emerged were resources and benchmarks based on CEFR requirements.

RESOURCES

Planning was predominantly based on the support of CEFR-aligned resources which was divided into printed resources and electronic resources (e-resources). This section reports the findings based on the type of CEFR-aligned resources that teachers referred to in planning for vocabulary lessons.

In terms of printed resources, the finding revealed that all the teachers cited the CEFR-aligned curriculum documents provided by MOE Malaysia, such as *Dokumen Standard Kurikulum dan Pentaksiran* (DSKP) and scheme of work (SoW), for lesson planning, this includes vocabulary lesson. As mentioned by teacher A3 during the interview:

> "We have the DSKP and SoW for every year. DSKP is the curriculum document and SoW is the scheme of work for teacher's reference. We refer to it when doing our yearly plan and daily lesson plan. The DSKP is details of every topic for each year. SoW comes with suggested teaching activities that we can use and it helps us with writing lesson plans."
>
> (A3)

Teacher A3 referred to the DSKP and SoW for her daily lesson planning. She explained that DSKP gives detailed information on each topic for each level, whereas SoW includes suggested instructional exercises for teachers' use. All the teachers agreed with teacher A3 that they depended on the DSKP and SoW as the main reference for lesson planning.

Furthermore, the teachers understood the flexibility of the suggested activities in the DSKP and SoW, which they could modify to suit the learners' needs as well as to suit various learning situations. As shared by teacher A1 during the interview:

> "Erm, basically SoW is the Scheme of Work prepared by the Ministry of Education Malaysia or the panel of CEFR. They have prepared the SoW as a guideline for us, so we just follow that. But if we want to improvise, we can improvise, and if we want to change the activities, we also can change."
>
> (A1)

Teacher A1 was aware that the curriculum documents provided by MOE Malaysia were only guidelines, not rigorous regulations to be followed strictly. She understood that teachers were allowed to modify the suggested instructional activities in the SoW to meet the needs of their learners and the classroom environment. This indicated that there was room for teachers to be creative and enhance their lesson preparation. All other teachers responded similarly.

Similarly, to cross-check the teacher's responses during the interview sessions, document analysis was used to compare the teacher's lesson plan to the DSKP and SoW. Findings indicated that the majority of teachers either implemented the activities exactly as recommended by the curriculum papers or modified them to fit the levels and needs of their learners. This is especially noticeable while planning to teach vocabulary because neither the book nor the curricular materials that correlate with the CEFR included information on how to teach vocabulary.

Findings show that teachers understood that DSKP and SoW focused on language skills such as listening, speaking, reading, writing and language arts. Therefore, the teachers perceived

that vocabulary was not a stand-alone content or language skill, and it needed to be integrated when teaching other language skills as mentioned. As suggested by teacher A3 during the interview:

> "Oh, ya [yes]. Like [on] Monday [we are] supposed to [do] listening [skill], [on] Tuesday [we teach] speaking, [on] Wednesday [we are] supposed to [teach] reading, [on] Thursday erm [it's teaching] writing and [on] Friday [it is] supposed to be language arts. So, when I am teaching vocabulary right, it is integrated in… like on Monday when I [am] teaching listening right, so the vocabulary is inside the listening lesson also."
>
> (A3)

Teacher A3 explained how she integrated vocabulary into her lesson planning for other language skills. For example, she integrated vocabulary when planning a lesson for listening skills. Her decision to integrate vocabulary shows her understanding of vocabulary being the building block for learning other language skills. The other participants responded similarly regarding this matter.

Another type of printed resources cited by teachers were the CEFR-aligned student's book and teacher's guidebook provided by MOE Malaysia. During the interview, the teachers cited the Teacher's Guidebook for lesson planning. As reviewed by teacher A2 during the interview:

> "There is the textbook and the teacher's book. I will go and refer to the teacher's book first, we have the teacher's guide book. From the teacher's book, there will have the guidelines on how you want to answer, how you want to start your lesson ah. From lesson 1, lesson 2, lesson 3 we follow that guideline. Everything is written in the Teacher's Book."
>
> (A2)

Teacher A2 used the CEFR-aligned Teacher's Guidebook as her primary resource for lesson planning since it provided teachers with the suggested instructional activities which linked to the CEFR-aligned Student's Book. She also mentioned that the teacher's guidebook recommended some questions and answers on the topic to help teachers design lessons. Similarly, all of the other teachers agreed that the Teacher's Guidebook provides clear instructions to aid lesson planning. Teachers frequently cross-checked the vocabulary included in the Student's Book with the corresponding tasks from the Teacher's Guidebook when planning a lesson. As mentioned by teacher A1 during the interview:

> "Usually for vocabulary, I look at the words in the pupil's book first, then I refer to the suggested activities in the SoW or the teacher's guidebook. If suitable, I use, if not sometimes I refer to Teacher Fiera or Teacher Ash too for the activities and the audio-visual materials."
>
> (A1)

Teacher A1 reviewed that she usually cross-checks the vocabulary listed in both the CEFR-aligned Student's and Teacher's Guidebook. She reviewed that if the vocabulary listed or activities suggested in the book were unsuitable for her learners, she would source alternatives online through Teacher Fiera or Teacher Ash's blog. This indicated that the teachers were able to adapt the suggested instructional activities from the CEFR-aligned Teacher's Guidebook into their actual classroom practice. Likewise, teacher A3 agreed with this statement.

Next, in terms of electronic resources (e-resources), CD was one of the e-resources provided by MOE Malaysia. The CDs provided match the content and activities from the CEFR Teacher's Guidebook and Student's Book. During the interview, although the teachers provided positive feedback on the content saved in the CDs, most of the teachers admitted that they seldom use it. As reviewed by teacher A2 during the interview:

> "Actually, everything is provided for the textbook and teacher's book. If you look at the textbook, they have the logo of a CD there, and then they can find its audio in the CD. The CDs are paired with the student's book. There is audio inside, I think there were some short videos in it too. But I am not too sure as I seldom use it."
> (A2)

Teacher A2 agreed that the CDs provided by MOE Malaysia were compatible with the activities listed in the CEFR Student's Book and Teacher's Guidebook. The CD logo shown alongside the suggested activities in the book indicates the compatible AV resources, such as audio recordings or short videos. Even though teachers were aware of the benefits of the CD, teachers "seldom use" it due to certain constraints which will be discussed later in the challenges section.

When the teachers' lesson plans were analysed, only teacher A1 used the content from the CD in her lesson planning. She integrated the use of audio recordings from the CD to help with pronunciation of the vocabulary words related to "objects in the classroom." Neither teacher A2 nor A3 showed any sign of CD usage when their lesson plan and observation data were analysed. This corresponded to the teachers' response during the interview that they "seldom use" the CEFR-aligned CDs.

Findings revealed that teachers preferred using online resources over CDs when planning vocabulary lessons due to the accessibility of those materials. Several popular sources were recommended by the teachers during the interview, including Telegram Group, YouTube, Facebook, Teacher Feira, and Teacher Ash blogs. The most popular e-resources among teachers were obtained from the Telegram Group, which teachers used as a platform to share teaching and learning materials. As mentioned by teacher A3 during the interview:

> "When I plan for teaching vocabulary, the teaching materials, I don't like use the CD. I always just use the one download from Teacher Feira Telegram group. Telegram we can type there and then there we can find the audio, videos, flashcards and sometimes worksheet also have. Then, we can use for the lesson. Sometimes the teachers share their lesson plan too in there. It makes it easier for planning the how to teach the words."
> (A3)

> "The videos online are easier to download, and also, it's up-to-date, not like CDs, the video not updated unless change new CD. But with videos online, it is always up-to-date."
> (A3)

Even though AV materials were provided in the CDs, Teacher A3 preferred to source teaching resources online rather than using the CDs. The Telegram group was her main source of information as there were a variety of teaching resources available in Teacher Feira's Telegram Group such as audio, videos, flashcards and worksheets. She also reviewed that sample lesson plans were sometimes shared by English teachers in the telegram group. She also reviewed that online resources were kept up-to-date as it provides flexibility for teachers to update or modify the teaching materials easily rather than the pre-saved materials on the CD. Similarly, teachers A1 and A2 also preferred the use of online resources for lesson planning due to their accessibility, as well as being user-friendly.

In reality, findings showed that the availability of the resources determined the teachers' choice of approach for vocabulary lesson planning. From this study, the AV approach was the most popular choice among teachers. Thus, teachers preferred to use AV materials to engage learners' sense of sight and hearing when teaching vocabulary. As teacher A1 had mentioned during the interview:

> "Okay. For me, when I use the word cards, it's not really works. So, I have to start with the pictures first, then I say the words, they repeat. After I introduce the words with the pictures so they can remember the words. The next lesson, they will remember the words because they can see the pictures, they can hear the words, they can say the words, and they can see the spelling."
>
> (A1)

Based on the findings, the primary factor that influenced teacher A1's vocabulary lesson planning was her prior experience of using the AV approach in the classroom. According to her, using the AV approach was able to provide opportunities for learners to "see the pictures" and "see the spelling" through the sense of sight, as well as being able to "hear the words" and "say the words" through the sense of sound simultaneously. Thus, the variety of AV resources available allows teachers to be creative in vocabulary lesson planning. Teachers A2 and A3 also responded similarly regarding this matter.

Apart from that, the video was another popular choice of AV resource among teachers. Findings from the interviews indicate that all three teachers find videos effective for teaching complicated vocabulary, especially when they find it hard to explain verbally or visually alone. As mentioned by teacher A1 during the interview:

> "Erm audio-visual…. picture cards with the recording, yes. But video I seldom use, unless very complicated topic and words because pupils can see and hear at the same time. Picture card is good for simple vocabulary."
>
> (A1)

Teacher A1 admitted that she only uses videos when teaching complicated topics with difficult vocabulary mainly due to the characteristics of videos which accommodate both audio and visual elements. Videos, she believed help learners to visualise and understand the words better. Teacher A2 and teacher A3 responded similarly on this matter.

BENCHMARK

In general, CEFR language proficiency levels were used as a reference to develop the levels of achievement *(Tahap Penguasaan)* as documented in the Standard Curriculum for Primary School *(KSSR or Kurikulum Standard Sekolah Rendah)* in Malaysia (BPK, 2017).

Findings from this study indicated that teachers were aware of the CEFR proficiency levels being the benchmark for international language proficiency levels. All the teachers were able to list out the six language proficiency levels listed in CEFR in general. For instance, as mentioned by teacher A3:

> "CEFR is the Common European Framework Reference for language languages. Scales A1 A2 which is for beginners or basic, B1 B2 for independent, and C1 and C2 for proficient level. So, we have to plan activities to see which scale they can reach to."
>
> (A3)

Teacher A3 was able to list out the six CEFR language proficiency levels ranging from A1 to C2. She was aware that scales A1 and A2 refer to the basic learner; scales B1 and B2 refer to the independent learner; and scales C1 and C2 refer to the proficient learner. She understood that these CEFR language proficiency levels served as a reference for planning and designing instructional activities aiming to reach the targeted language proficiency level. However, when teachers were asked to explain how the CEFR proficiency levels were implemented, their explanations shifted to the KSSR levels of achievement, which are Tahap Penguasaan (TP), ranging from 1 to 6.

> "Okay. CEFR is erm… CEFR is erm… an international standard that is used by our Ministry of Education of Malaysia to improve English language proficiency of Malaysian students. The benchmark for English standard in Malaysia is like 6 levels, there are 6 levels. Erm.. but in school, we have our own benchmark *tahap penguasaan satu, dua, tiga, empat, lima dan enam [achievement level one, two, three, four, five and six].* Level 3 or TP 3 - *Tahap Penguasaan tiga [achievement level 3]* is the minimum requirement. We refer to the TP when deciding what to teach and how to assess the pupils"
>
> (A1)

Based on the findings, although Teacher A1 had the basic knowledge that the CEFR proficiency level was used as the international benchmark by MOE Malaysia when implementing it in the classroom, she explained that TP was used as the benchmark in school. This shows that the teacher viewed CEFR language proficiency levels and KSSR levels of achievement as two different things. In reality, the KSSR level of achievement is derived from the CEFR language proficiency levels. Teachers A2 and A3 also show having limited knowledge regarding CEFR language proficiency levels and they referred to the KSSR level of achievements instead for lesson planning and assessment in the classroom.

CHALLENGES IN USING THE AV APPROACH TO TEACH VOCABULARY

Findings show that teachers were faced with challenges in terms of planning how to use an AV approach to teach Vocabulary. Based on the data, the challenges that teachers faced were categorised into two main themes, which are the infrastructure and the policy challenges.

INFRASTRUCTURE

The first theme that emerged from the findings was the infrastructure-related challenges. It will be reported based on two categories, which are the lack of information technology (IT) infrastructure, and convenience of access.

The findings indicated that the lack of IT infrastructure in school was one of the main challenges faced by the teachers. During the interview, teachers A1, A2 and A3 expressed their desire to use the AV approach to teach vocabulary as they were aware of its benefits, however, they were faced with challenges due to the lack of IT infrastructure provided in school. As mentioned by teacher A1 during the interview:

> "Sometimes I feel like showing them videos, but *susah juga kena bawa laptop dan pasang LCD* [*this is not easy as I have to bring the laptop and put on the LCD]…Ya*, [Yes] but I think, if possible *la*, video would be a good choice."
>
> (A1)

Teacher A1 expressed her desire to "use videos to teach" in the classroom, but she did not have the IT infrastructure needed to operate it. If she wanted to show videos, she needed to bring the laptop and Liquid Crystal Display (LCD) projector to the classroom and assemble it herself. As stated by teacher A2 in the interview sessions:

> "If LCD and speaker is provided, of course, I would want to use audio-visual because it will make the class lively. Students can see it with their own eyes, and easier to understand. If [I am] teaching new vocabulary, just play the video, *murid akan tengok dan faham sendiri, kan senang [students can watch and understand], ...*But we don't have that here, so…"
>
> (A2)

Findings show that teacher A2 believed that the AV approach helped to make teaching and learning in the classroom lively, whereby learners will be able to "see" with their own eyes, thus making vocabulary learning easier. Teacher A1 and A3 shared similar opinions during the interview. However, they seldom had a chance to use AV resources such as videos due to the lack of IT infrastructure in the classroom. To apply the AV approach, they had to bring a portable LCD into class as an alternative. During the interview, teacher A2 shared:

> "If in the classroom, we got that limitation. Usually, we can paste the white paper, then bring in LCD to project on the white paper as screen. That's why we seldom use video to teach, very *susah [difficult]* and waste a lot of time. Mostly use flashcards and audio recording."
>
> (A2)

Teacher A2 shared her challenge when using the AV approach. As an alternative, she had chosen to "paste the white paper" as a temporary screen for the portable LCD that she brought into the classroom. She reviewed such action as "difficult" and time-consuming. Similar opinions were voiced by the other teachers during the interview too.

The second challenge found in relation to IT infrastructure was the issue regarding convenience of access. Based on the findings, the convenience of access can be divided into two subcategories, which are the issue of not having a laptop with a CD player and time consumption.

During the interview, teachers mentioned that they had trouble accessing the AV materials on the CD since their laptops did not include a CD player. This was one of the reasons teachers decided to source the AV materials online instead. As responded by teacher A3 during the interview:

> "...and also because my laptop cannot play the CD, so I just search from Youtube, it's easier and it's the same. I no need to waste time to find ways to play the CD."
>
> (A3)

Teacher A3 addressed the challenge she faced for not being able to play the CD provided with her laptop without a CP player. So, she decided to source for the AV materials from YouTube instead. Similarly, teachers A1 and A2 also preferred to source AV resources online, especially through the "Telegram Group" which they believed to be more convenient than wasting time looking for alternate ways to access the information on the CD.

Undoubtedly, the infrastructure challenges will affect the teachers' lesson planning as well as the choice of approaches used to teach vocabulary.

POLICY CHALLENGES

The second theme that emerged from the finding was the incompatibility of the learner's competency with the policy associated with the CEFR. Therefore, the policy challenges reported in this section were divided into two: (1) the issue of cultural unfamiliarity, and (2) the issue of the learner's language proficiency level.

The issue of cultural unfamiliarity appeared in two areas for this study: the content and the accentedness, which was unfamiliar to the teachers.

In terms of content, all three teachers agreed that the CEFR-aligned curriculum covered a wide range of topics with foreign cultures to which they were unfamiliar. Thus, it is a problem for teachers to plan how to teach foreign cultures as they find it difficult to relate to our local Malaysian culture. As shared by teacher A1 during the interview:

> "Level 2, it is very difficult because we have to learn other cultures, foreign cultures. For example, when..when we go to the topic about food, some of the foods mentioned in the book are not the food we have here, for example tortillas. So, it's hard to relate it to Malaysia *sedangkan [whereas]* in the book we talked about the foreign country and foreign culture, so it's quite difficult."
>
> (A1)

It was challenging for Teaching A1 to relate the foreign culture covered in the Level 2 syllabus to the local culture. This was a big challenge for her as it was difficult to explain foreign content which is very different from his own culture. For example, it was difficult for her to help the learners understand what is a "tortilla" because it was not available locally. Teachers A2 and A3 responded similarly on this issue.

Besides the content, the unfamiliarity of the accentedness was another challenge faced by the teachers. The AV materials provided by MOE Malaysia were developed by Cambridge which integrated the foreign culture and accent into its content. Teacher A2 were concerned about the native accent presented in the CEFR-aligned materials provided. As mentioned by teacher A2 during the interview:

> "If for the slow students, maybe they don't really understand because have foreign slang and speak very fast…We have to tell them and stress on what are being told in the audio. We repeat it with our own voice."
>
> (A2)

Teacher A2 expressed her concern regarding the use of the AV materials as it was recorded with a British accent in which the pronunciation of the English words may differed. Teacher A2 was concerned that the slow learners would struggle to keep up with the teaching process due to a lack of understanding of the native accent and slang, paired with native speakers speaking at a fast pace. Therefore, the teachers needed extra effort to repeat the content in their own words and stress the main points for the learners to understand.

The issue of the learner's language proficiency level was another sub-theme for policy challenges. According to the findings, all of the teachers saw the inequalities in the language skill levels of the learners as part of the issue when adopting CEFR-aligned content. This was primarily due to learners having different standards and learning at different rates. As mentioned by teacher A3 during the interview:

> "Nowadays, I think it is quite challenging. Challenging in a way that if we see from the student's background as this language is not their mother tongue. So, it is very challenging, erm how to say, it's the responsibility for the English teacher to teach this language to them."
>
> (A3)

Teacher A3 addressed the challenge she had faced in teaching learners from different backgrounds who speak different languages. She understood the responsibility of teaching the English language falls on the shoulders of teachers, especially when the learners do not come from an English background. All the teachers responded similarly regarding this issue.

DISCUSSION

The findings were discussed based on two major themes, which are related to the resources that supported vocabulary lesson planning, and the challenges encountered by teachers in using the AV to plan and teach vocabulary.

Lesson planning plays a decisive role in the success or failure of a lesson (Alanazi, 2019; Phan, 2021). With the introduction of CEFR in our Malaysian English language curriculum, a variety of printed and e-resources were provided by MOE Malaysia to assist teachers with the implementation of the AV approach in the actual classroom (BPK, 2017).

The findings of this study reported that the teachers' vocabulary lesson planning was predominantly based on the CEFR-aligned resources available. All three teachers cited the printed resources, such as curriculum documents (the DSKP and Scheme of Work), as their main reference for lesson planning. Based on their lesson plans, the teachers were discovered to be either directly incorporating the suggested activities into their lessons or choosing to modify them. This was consistent with the findings of Ngu and Azlina (2019), who discovered that the teachers' lesson planning and classroom practises were positively impacted by the CEFR-aligned curriculum. However, previous research has found that teachers do not see vocabulary as a priority in their classroom (Maizatulliza & Kiely, 2018), as vocabulary only appeared in the DSKP and SoW as a wordlist with no further instruction (BPK, 2017; Maizatulliza & Kiely, 2018). As such, teacher pedagogical content knowledge (PCK) was the basis for their decision-making and lesson preparation about vocabulary. The PCK of the teacher was crucial in bridging the gap between the pedagogical knowledge of "how to teach vocabulary" and the content knowledge of "what is vocabulary" (Pompea & Walker, 2017; Shulman, 1987; 2004). The teachers had to rely on their PCK to integrate the AV resources into the vocabulary lessons because there was not much guidance on teaching vocabulary in the curriculum documents. In relation, the findings of this study show that teachers decided to incorporate vocabulary instruction into teaching other language skills such as listening, speaking, reading, and writing. This shows that the teachers understood that vocabulary acts as a building block for learning other English language skills (Mohd. Nor et al., 2015; Sovakandan et al., 2017).

As mentioned earlier, in contrast to the urban schools, Kiss and Rimbar (2021) have addressed the increasing challenges faced by rural school teachers to adopt the CEFR-aligned resources in their classes. The lack of locally relevant content in the CEFR-aligned teaching resources was the main reason leading to the challenges faced by teachers when they plan to use the resources for vocabulary teaching in rural classrooms. The CEFR-aligned resources provided by MOE Malaysia, including the teacher's guidebook, student book, and AV resources, largely reference other cultures, which leaves a lack of locally relevant cultures. This is similar to Ngu and Azlina's (2019) research. The tools' broad usage of foreign cultures can be confusing to teachers and students, which makes adaption in the local classroom challenging. Accordingly, the teachers in this study were still dealing with the difficulties of cultural unfamiliarity and unfamiliarity with accentedness due to a lack of localised context, which is similar to the issues that teachers had previously faced and were identified in studies by Nurul Farehah and Sallehhudin (2017) and Ngu and Azlina (2019). Therefore, the findings from this study reflected the need to train more CEFR experts who are able to cater to the context of the Malaysian classroom while at the same time not losing the essence of the global context in CEFR. Similarly, Nurul Farehah and Sallehhudin's (2017) study has also proposed that Malaysia needs more CEFR experts to create locally aligned resources.

Additionally, as technology integration becomes more prevalent in Malaysian educational systems, an increasing number of e-resources—like CEFR-aligned AV resources—are being produced by subject-matter experts and kept on CDs (Ngu & Azlina, 2019). Even though MOE Malaysia offers AV resources in that format, the study's findings nevertheless demonstrated that teachers were reluctant to use the CDs. Consistent with Ngu and Azlina's (2019) study, the results indicated that most teachers opted to obtain AV resources online for vocabulary lesson planning, instead of using pre-saved resources on CDs. Nonetheless, teachers did recognise the advantages of using AV to teach vocabulary, especially those related to foreign cultures and difficult for teachers to explain orally. However, to maximise the impact of AV, educators needed to prepare ahead of time. The teachers involved in the study were still facing challenges in utilising the AV resources that had been included on CDs because they lacked expertise in doing so (Nawai & Nur Ehsan, 2020; Nurul Farehah & Mohd. Sallehudin, 2017). This addressed the necessity that policymakers provide suitable resources to support teachers in teaching vocabulary, particularly concerning other cultures.

In terms of challenges encountered by teachers in using the AV approach, findings from the previous studies addressed the physical challenges faced by teachers, such as the lack of resources (Utami, 2015;), poor network coverage (Ngu & Azlina, 2019), and lack of equipment and facilities (Wazeema & Kareema, 2017; Yonas et al., 2020). Even though MOE Malaysia has provided support to teachers by preparing AV resources to assist teachers in implementing the CEFR-aligned curriculum, it is not adequate to accommodate all the schools in the country (Ngu & Azlina, 2019), especially in rural schools in Sarawak. Likewise, the teachers in this study who taught in the rural school had similar challenges, such as a lack of IT infrastructure and convenient access. In reality, no technology can be applied to all situations because every classroom is unique (Mutanaga et al., 2018). As a result, teachers must be aware of the constantly changing connections between content, pedagogy, and technology. To overcome these challenges, policymakers should take into account the need to provide IT support and improve IT infrastructure to assist teachers in their lesson delivery. Previous studies made similar recommendations, urging schools to provide supportive environments for technology integration in the classroom (Ajloni, 2019). I hope that the teachers' acceptance and readiness to teach vocabulary using the AV approach will increase.

Future research ought to investigate how policymakers may support teachers' professional dispositions in planning and delivering CEFR-aligned curricula in schools, especially to improve vocabulary acquisition because vocabulary is the foundation for learning the English language. Understanding the perspectives of educators is crucial to assisting policymakers in planning effective professional development sessions for teachers.

CONCLUSION

This study provides insight for other researchers regarding teachers' vocabulary lesson planning. The outcomes of the study shed light on how teachers plan to use the AV approach to teach vocabulary in a primary-level English language classroom. The findings showed that when teachers were given the autonomy to plan and decide how to approach vocabulary teaching, it depended on the teachers' PCK. The findings of this study suggested that teachers understood the flexibility of the CEFR-aligned resources which serve as a guideline to help teachers integrate vocabulary into teaching other language skills. Findings also show that teachers still faced challenges when adopting the AV approach to planning and teaching vocabulary in a rural school

in Sarawak. Even though they were aware that MOE Malaysia provides CEFR-aligned AV resources, teachers at the rural school were found to be dealing with a lack of IT infrastructure and ease of access. This study also reflected on policy challenges that occur in the classroom, such as the issue of cultural unfamiliarity and the learner's language proficiency level. The findings of this study provide useful insights for practitioners on how teachers in rural areas adopt the AV approach to teaching vocabulary in their classrooms. The findings also provide policymakers with information about the policy and infrastructure challenges that teachers face in vocabulary teaching. Finally, the result of this study could contribute to the limited research on vocabulary teaching through the CEFR-aligned curriculum in the Malaysian context.

ACKNOWLEDGEMENT

The authors acknowledge Universiti Malaysia Sarawak (UNIMAS) for the financial support.

REFERENCES

Alanazi, M. H. (2019). A Study of the Pre-Service Trainee Teachers' Problems in Designing Lesson Plans. *Arab World English Journal. 10(1).* 166-182

Aminuddin, A., & Azman, M. (2017). The use of ICT in English language learning in Sarawak secondary schools: A case study. *International Journal of Education and Development, 3(1),* 1-11.

Anas, F. A., Zakaria, N. Y. K. (2019). *ESL Learners' Perceptions on English Subtitled Audio-Visual Materials.* Creative Education, 10, 2796-2803. https://doi.org/10.4236/ce.2019.1012206

Azman, A., & Yusof, N. M. (2018). The impact of technology on teaching and learning in Sarawak schools. *International Journal of Education and Development, 5(2),* 115-124.

Bahagian Pembangunan Kurikulum. (2017). *Bahasa Inggeris Sekolah Jenis Kebangsaan: Dokumen Standard Kurikulum dan Pentaksiran*. Kementerian Pendidikan Malaysia

Bayoung, P. D., Rhessma, N.C.A, Theresa, C.L. & Yunus, M. M. (2019). Vocab-T: Enhancing Young Learners' Vocabulary. *International Journal of Academic Research in Business and Social Sciences, 9(1)*, 210-220. https://doi.org/10.6007/IJARBSS/v9-i1/5390

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage Publications.

Braun, V. & Clarke, V. (2006) *Using thematic analysis in psychology.* Qualitative Research in Psychology, 3 (2). pp. 77-101. Retrieved from: http://eprints.uwe.ac.uk/11735

Chan, S. L., & Kapong, J. M. (2021). Personal Growth Through English Proficiency: A Case Study in Sarawak. *Iasaýı Ýnıversıtetiniń Habarshysy*, *122(4),* 85–93. https://doi.org/10.47526/2021-4/2664-0686.08

Creswell, J. W., & Clark, V. L. P. (2011). Designing and conducting mixed methods research (2nd ed.). Thousand Oaks, CA: Sage.

Kalajahi, S. A. R. and Pourshahian, B. (2012). *Vocabulary learning strategies and vocabulary size of ELT students at EMU in Northern Cyprus*. English Language Teaching, 5(4). 138-149.

Kiss, T., & Rimbar, H. (2021). *English language teacher agency in rural Sarawak: Exploiting teaching materials.* The English Teacher, 48(1), 142-158.

Ler, E. C. (2012). *Cultural factors affecting English proficiency in rural areas.* Advances in Language and Literary Studies, 3(1), 1–23. http://dx.doi.org/10.7575/aiac.alls.v.3n.1p.1

Lu, D. H. (2017). *Teachers' Beliefs and Practices: ESL Teachers' Perceptions of Vocabulary Instruction*. Culminating Projects in English, 93. https://repository.stcloudstate.edu/engl_etds/93

Muhamad, M. & Kiely, R. (2018). Understanding Teachers' Pedagogical Knowledge in ESL Vocabulary Teaching. *Journal of Arts & Humanities. 7(1)*. 2167-9053. http://dx.doi.org/10.18533/journal.v7i1.1328

Musa, N.C., Lie, K.Y., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA Online® Journal of Language Studies, 12 (1).*

Nawai, R., & Nur Ehsan (2020). Implementation challenges of Common European Framework Reference (CEFR) in a Malaysian Setting: Insights on English Teachers' Attitude. *International Journal of Academic Research in Business and Social Sciences*, *10(7),* 28–41.

Ngu, M. K. & Azlina. (2019). *English Language Teachers' Perceptions On The Implementation Of Cefr-Aligned Curriculum Among Primary Schools In Malaysia.* Seminar Wacana Pendidikan 2019(SWAPEN 2.0*)*. eISBN 978-967-13352-8-4 212.

Nurul Farehah Mohamad Uri & Mohd Sallehhudin Abd Aziz. (2018). Implementation of CEFR in Malaysia: Teachers' Awareness and the Challenges*. 3L: The Southeast Asian Journal of English Language Studies, 24*(3). 168-183. http://doi.org/10.17576/3L-2018-2403-13

Patton, M. Q. (2015). Qualitative research and evaluation methods (4th ed.). Thousand Oaks, CA: Sage.

Phan, T. T. N. (2021) 'An investigation into difficulties in designing lesson plans experienced by pre-service teachers at the University of Foreign Language Studies – the University of Danang', *Borneo International Journal of Education (BIJE), 3*(2021). https://doi.org/10.51200/bije.v3i.4109

Philip, B., Tan, K.H & W. Jandar. (2019). *Teacher Cognition in ESL Context: A Pedagogic Perspective in Malaysian Classroom.* Bangi. UKM Press.

Pompea, S., & Walker, J. (2017). The impact of pedagogical content knowledge on teacher candidates' beliefs and practices in astronomy. *International Journal of Science Education*, *39*(14), 1943-1963.

Richards, J. C., & Schmidt, R. (2002). *Longman dictionary of language teaching and applied linguistics*. London; New York: Longman.

Scott, J. A, & Nagy, W. E. (1997). Understanding the definitions of unfamiliar words*. Reading Research Quarterly, 32,* 184-200.

Shulman, L. (1987). *Knowledge and teaching: Foundations of the new reform.* Harvard Educational Review, *57*(1), 1-23.

Shulman, L. S. (2004). *Knowledge and teaching: Foundations of the new reform.* San Francisco, Jossey Bass.

Sovakandan, H., Jaganathan, P., & Hussin, F. M. (2017). Investigating low proficiency ESL students' perception of the use of i-Think Maps in writing practices. *Malaysian Journal of ELT Research*, *14*(2), 1-13.

Susanto, A. (2017). The Teaching of Vocabulary: A Perspective. *Jurnal KATA, 1*(2).

Swaran Singh, Charanjit. (2021). Review of Research on the Use of Audio-Visual Aids among Learners' English Language. *Turkish Journal of Computer and Mathematics Education (TURCOMAT). 12.* 895-904. 10.17762/turcomat.v12i3.800

Utami, D. (2015). *Teaching Vocabulary Using Audiovisual Aids At The First Grade Of Sdn 2 Sawahan In 2014/2015 Academic Year: A Classroom Action Research*. Corpus ID: 150331088

Wazeema, T. M. F., & Kareema, M. I. F. (2017). *Implication of multimedia audio-visual aids in the English language classroom*. Proceedings of 7th International Symposium, SEUSL, 7th & 8th December 2017.

Yin, R.K. (2011). *Qualitative Research from Start to Finish.* New York, NY: The Guilford Press.

Yonas, A. & Senapathy, M. (2020). The Factors Impede Teachers' Effective Use of Audio Visual Materials to Teach English Speaking Skills: ABBA PASCAL Girls' School in Focus, Wolaita Zone, Southern Ethiopia. *Randwick International of Education and Linguistics Science (RIELS) Journal, 1(3)*. 439-448. https://doi.org/10.47175/rielsj.v1i3.157

Zeta, Lydia & Subon, Frankie & Unin, Norseha. (2019). Using Language Games for Vocabulary Retention in a Rural Primary School in Sarawak. *International Journal of Service Management and Sustainability. 4*. 10.24191/ijsms.v4i2.8067.

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Local problems and a global solution: examining the recontextualization of CEFR in Thai and Malaysian language policies / Savski, K.**
Source	***Language Policy*** **Volume 19 (2020) Pages 527-547** **https://doi.org/10.1007/s10993-019-09539-8** **(Database: Springer Link)**

Language Policy
https://doi.org/10.1007/s10993-019-09539-8

ORIGINAL PAPER

Local problems and a global solution: examining the recontextualization of CEFR in Thai and Malaysian language policies

Kristof Savski[1]

Received: 12 October 2018 / Accepted: 11 October 2019

Abstract
Since its publication in 2001, the Common European Framework of Reference for Languages (CEFR) has become a highly influential means of describing language proficiency. Its spread has, however, been marked by contradictions, since the framework has been appropriated in the service of a variety of different policy agendas. In this paper, I argue that such contradictions are indicative of broader ideological contrasts, which may impact how the framework is implemented at the local scale. By drawing on critical discourse analysis and conceptual history, I analyse a set of recent language policy texts from Thailand and Malaysia, two Asian contexts where CEFR has recently been introduced, to examine how such global ideological struggles connect with local agendas. I find that CEFR has in these multilingual contexts been embedded into a bilingual policy agenda which foregrounds the national language (Thai or Bahasa Malaysia) and English while backgrounding other languages. This means that CEFR was detached from the agenda of the Council of Europe, with the recontextualization of CEFR shown to have been a selective process in which the only part to be consistently transferred were the CEFR levels, which were in this decontextualised form presented as a transnational standard. I argue that these patterns are indicative of a struggle between the global agenda of ELT and its roots in the ideology of neoliberalism, that underlies much of the worldwide spread of CEFR, and a local nationalist agenda attempting to appropriate the framework for its own purposes.

Keywords CEFR · Recontextualization · Globalization · Thailand · Malaysia

✉ Kristof Savski
kristof.s@psu.ac.th

[1] Faculty of Liberal Arts, Prince of Songkla University, 15 Karnjanavanich Rd., Hat Yai 90112, Thailand

Introduction

In 2013 and 2014 respectively, the Malaysian and Thai governments announced that they would begin using the Common European Framework of Reference for Language (CEFR) in the development and implementation of future language policies. By doing so, both governments appear to be following a worldwide trend, with the framework having long ago transcended European borders and become a globalized language policy instrument (Byram and Parmenter 2012). At the local scale, CEFR has seen widespread use in the design of curricula and in the development of teaching materials and tests. In parallel, the framework has also become increasingly associated with the global influence of major ELT textbook producers and testing organizations, which have made significant use of CEFR in the development and marketing of their products (Littlejohn 2012). Additionally, the framework remains associated with the original agenda underlying its development, that of European integration, though this in itself is a site of struggle between humanist and neoliberal language policy discourses (Krzyzanowski and Wodak 2011).

The aim of this paper is to examine how the recontextualisation of CEFR in non-European contexts takes place against this complex discursive background. The paper begins by considering, from a discursive perspective, how globalization affects language policies. I continue by discussing the trajectory of CEFR, starting with its inception in the European context while also discussing its use in globalized English language teaching products. I then examine the recontextualisation of the framework in Thai and Malaysian language policy, focussing in particular on the key concepts that underpin the policy agendas in which CEFR is integrated in either context. A data-set of recent policy texts from both contexts is analysed with the aim of addressing the following research questions: What local language policy agendas is CEFR integrated into and what are the key concepts and conceptual relations that underpin those agendas? What elements of CEFR are recontextualised and what conceptual relationships are they placed into? How are ideological struggles surrounding CEFR negotiated in policies at the local scale?

Globalized language policies: scale and recontextualisation

Attention in ethnographic and discourse-analytic work in sociolinguistics (see e.g. Heller 2011) has recently shifted to the examination of relationships between different contexts of language use. To aid in the theorisation of such a conceptual refocus, Blommaert (2007) proposes the concept of scale, which he sees as complementing the existing horizontal dimension (represented in concepts like language community) with "a vertical dimension of hierarchical ordering and power differentiation" (2007, p. 4). Issues of power are thus central to analysis of scales, though it is, as pointed out by Canagarajah (2018), overly simplistic to

assume that a higher or lower position in such a hierarchy would automatically correspond to a higher or lower level of agency for the actors involved. Instead, the examination of scales continues to have an ethnographic orientation, stressing the need to examine particular nexuses of practice and describe the configurations of power that language mediates within them (e.g. Hult 2010; Pietikäinen 2010; Savski 2018).

The concept of scale has also seen use in language policy research, particularly in studies pursuing either a discursive or ethnographic approach, or a combination thereof (Barakos and Unger 2016; Wodak and Savski 2018). For policy analysis, the concept of scale is seen to add not a complementary dimension but to bring to the forefront the need to interlink various analytical levels in an area which is, in many cases, inherently hierarchical. This is particularly central to the examination of policies associated with state authority, since their creation, interpretation and implementation often takes place in conditions typified by the existence of both agentive opportunities and structural constraints (Jessop 2007; Savski 2016). This means that the trajectory of policies across spatiotemporal scales is often determined by the ways in which specific actors interpret them while negotiating constraints of different types (Hornberger 2005; Johnson 2013). In Savski (2018), I thus examined how actors participating in a committee meeting in the Slovene parliament negotiated subject positions imposed by institutional practices (e.g. committee chair vs. member of the public) as well as those imposed by the broader discourse surrounding language policy in Slovenia (e.g. linguist as expert vs. politician as non-expert) while attempting to agree on a mutually acceptable set of amendments to a language strategy. I showed how such lines both governed specific interactions at the meeting but also how actors were able to subvert and exploit them by switching between different subject positions in order to achieve their goals.

In this paper, I analyse the trajectory of a transnational language policy text as it is interpreted and appropriated by actors at the national level in two contexts. To examine this trajectory, I take a discursive approach by focussing on recontextualisation, which I understand to refer to the creation of intertextual and dialogical relationships through the transfer of specific elements of a given text to another context (Maybin 2017; Reisigl and Wodak 2015; see also Wodak and Fairclough 2010). Such a transfer is seen to be potentially determined both by agentive opportunities and structural constraints, with complex configurations of power determining what elements are transferred across the boundaries between particular scales. Furthermore, as such transfers involve the negotiation of boundaries between the power relations, practices, discourses and ideologies that characterise particular scales, they may lead to shifts in the meaningfulness of the recontextualised elements. Kulsiri (2006) for instance examines how elements of educational policy were recontextualised from the US state of Louisiana to Thailand, highlighting how the values of learner-centredness that dominated the original policy were complemented by elements conforming to the more teacher-centred practices traditionally found in Thai education.

This example also highlights the fact that policies which transcend the borders of a single nation-state and become globalized in the sense that they have become embedded in broader transnational and transcultural flows of people, technology,

information, finance and ideas (Appadurai 1990) are not examples of simple transfer but rather of hybridization. The globalization of policy takes the form of a continuous tension between antagonistic tendencies toward greater heterogenization, decentralization and particularization from the local scale and toward more homogenization, centralization and universalization from the global scale (Wodak and Fairclough 2010). Since this antagonism is often dependent upon the specific discourses and political economic conditions of each specific local context, the global spread of policies is often defined by its non-simultaneous and asymmetric nature (Krzyzanowski and Wodak 2009). Such local negotiation of global policies may often lead to unlikely compromises between opposing agendas, as highlighted by the aforementioned example from Kulsiri (2006).

Such agentive negotiation is, however, subject to the structural constraints presented by political-economic relations. A key point to underline is that policies which transcend the national level are often either expressly created with the intention of influencing decisions across a variety of potential polities (as is the case with CEFR) or have become associated with such an agenda despite having originally been created for use in a specific local context [as was the case with the Louisiana curriculum discussed by Kulsiri (2006)]. In other words, language policies associated with processes of globalization are instruments of transnational governance through which powerful global actors attempt to influence local policy (Holzinger and Knill 2005; Roger and Dauvergne 2016). While this may take the form of outright coercion or imposition of reform, as in the case of the imposition of austerity measures on European Union members in the wake of the Eurozone crisis, instruments of 'soft power' are the more usual form of transnational governance in fields like education and language policy, with league rankings, independent ratings and 'good practice' examples being common catalysts for policy change (De Costa et al. 2019; Rutkowski 2007). CEFR is also such an instrument of transnational governance, though highly specific in the sense that while it was developed for use in a clearly delimited set of polities (European nations) it has since also become associated with broader global agendas. Its trajectory toward globalization is discussed in the following section.

CEFR as a globalized language policy

The Common European Framework of Reference for Languages (CEFR) was published by the Council of Europe (CoE) in 2001, the direct result of a decade-long effort initiated in 1991 at a symposium in Rüschlikon, Switzerland (reported on in North 1992). As summarized by Trim, the main outcome of this symposium was a commitment to develop a framework which would 'assist learners, teachers, course designers, examining bodies and educational administrators to situate and coordinate their efforts' (2012, p. 29). Over the following years, the proficiency descriptions that have now become the most familiar element of CEFR were developed through a process in which teacher perceptions were collected and used to create series of descriptors calibrated to the now familiar six levels (with A1 and A2 representing 'basic', B1 and B2 'intermediate' and C1 and C2 'advanced proficiency';

the calibration process is described in detail by North 2014). While these levels were nominally new, they represent an example of evolution rather than revolution, with several of the levels having been previously described, starting with B1—originally referred to solely as Threshold (Van Ek 1975). The very fact that this level was described first is indicative of the motivation underlying it, with its intention originally having been to describe the abilities an L2 speaker needed in order to survive in a context where that language was in dominant use (ibid.). Such a need stemmed directly from the European post-war period, when the establishment of early organisations such as the European Community for Coal and Steel began to push transnational workforce mobility to the forefront of the policy agenda, with issues related to the linguistic and cultural integration of such economic migrants also becoming relevant (Tabouret-Keller 1991).

While the historical background of CEFR may thus be attributed to a relatively practical set of needs for a set of transnational guidelines for language learning, the framework was at the time of its publication attached to a much broader language policy agenda promoted by CoE. The centre-pieces of this agenda were the twin concepts of plurilingualism and pluriculturalism, which feature heavily throughout parts of the original publication (Council of Europe 2001). In the framework, they are seen as reflecting 'the fact that as an individual person's experience of language in its cultural contexts expands, […] he or she does not keep [different] languages and cultures in strictly separated mental compartments' (ibid., p. 4). In other words, the CoE agenda with which CEFR has become associated is one which argues for a fluid approach to defining language competence, one which focusses on the integration rather than the segregation of languages (for further discussion, see Piccardo 2010, 2013 and Savski 2019). By setting this agenda, CoE and CEFR can thus be firmly positioned at the centre of a significant shift in sociolinguistic thought, one which has emphasized the need to rethink established monolingual models of language and the speaker by, for instance, bringing to the fore concepts such as translanguaging (García 2009). Despite its significance in this respect, however, CEFR has also garnered significant criticism (e.g. Shohamy 2011; Pilkinton-Pihko 2013) for failing to incorporate plurilingualism and pluriculturalism more explicitly into the reference levels which have since become its most prominent part, a shortcoming addressed with the recent publication of a set of new reference level descriptions for mediation and plurilingual and pluricultural competence (Council of Europe 2018).

The impact that CEFR has had at the national level in Europe is significant, though this success may not only be attributable to the influence of the Council of Europe but also to the policy actions of the European Union (EU). Almost immediately after the publication of CEFR, the European Council adopted as part of its resolutions from the 2002 Barcelona summit the policy recommendation that all EU citizens should learn two foreign languages in addition to their first language and also called for the 'establishment of a linguistic competence indicator' (European Council 2002, p. 19). It is thus unsurprising that CEFR was soon adopted by the EU and recommended to its members (Jones and Savile 2009), becoming thus associated with a conceptualization of language learning that generally foregrounds its economic rather than cultural benefits (Krzyzanowski and Wodak 2011). The joint influence of both CoE and EU has meant that CEFR has had a broad impact on

language education policies across Europe, having in particular a central role in the reform of curriculum development and assessment practices in different nations. What these applications (described in edited collections such as Byram and Parmenter 2012) have highlighted is the flexibility with which CEFR is interpreted by its users in different contexts, a pattern aligned to the intentions of CoE that the framework be used in a flexible and descriptive rather than a rigid and prescriptive manner (Trim 2012).

While the 'European' trajectory of CEFR is in itself complex, a study focussing on how the framework impacts policies outside Europe needs to take into account a further type of user of the framework, namely major ELT textbook producers and testing organizations. Since its publication, CEFR has seen increasingly broad use by organizations such as Cambridge English and the British Council as a means of describing the difficulty levels of exams and textbooks, with the framework thus to an extent displacing the previous system of labels like 'basic' and 'lower intermediate'. A significant body of research has developed around such uses, with a variety of studies reporting on the different procedures relied on to achieve alignment between the relatively open-ended descriptions provided by CEFR and the more detailed specifications used in the design and evaluation of test tasks (e.g. Martyniuk 2010). Through this connection, CEFR has come to be associated with an agenda whose key presupposition regarding how language is to be taught and learned are potentially contradictory to the concept of plurilingualism that the framework was intended to advance, with international English exams and textbooks continuing to largely be associated with a monolingual learning model directed toward the acquisition of native-like proficiency in standard English (Hamid 2014; Shohamy 2011).

It is against this history of antagonism that the appropriation of CEFR beyond European borders takes place, though there is as yet little literature examining the ideological underpinnings of such uses. It is perhaps indicative that while some examples of such uses have sought to apply CEFR in explicit support of plurilingual language policies (see e.g. Arnott et al. 2017; Piccardo 2014), most applications have focussed specifically on English education (Byram and Parmenter 2012). The framework has also seen some adaptation, particularly in Asia, with localised versions having been developed for ELT purposes in Japan (Negishi 2012) and China (Jin et al. 2017). The focus of these uses and adaptations on ELT suggests a shift with regard to the ideological underpinnings of the framework, but raises questions regarding the extent to which different features of the discourse surrounding CEFR are recontextualised to the local scale and how they interact with local discourses. These questions are addressed in the following sections.

Methodology for examining CEFR in Thailand and Malaysia

Context and data

This paper presents the results of two case studies of CEFR recontextualization at the local scale. The two case studies, Thailand and Malaysia, exhibit significant parallels: The nations are in the same geographical region, share a border and are also

embedded in the same process of regional integration through their membership in the Association of South East Asian Nations (ASEAN), which has gradually seen a shift toward greater mobility in the region. Both are also categorised as nations with developing economies, typical of the region but atypical with regard to context where CEFR had been developed. The way in which the framework has been rolled out in the two also appears to bear many similarities, with the process starting almost concurrently (2013 in Malaysia, 2014 in Thailand) and with several policy texts produced since.

For the purposes of this research, recent educational policy texts in which CEFR was recontextualised were collected, which yielded a data set of nine texts, four from the Thai context[1] and five from the Malaysian (see Table 1 for a full list). A broad categorisation of the documents according to form and function indicated that various types of comparison were possible:

(a) During the time frame under examination, the government of each nation had drafted and published a macro-level education strategy. The two texts in question, M-Blueprint and T-Plan, exhibit numerous similarities and parallels and therefore allowed for closest comparison.
(b) Each government also produced meso-level policy texts more specifically aimed at English language education, to which CEFR was seen to be relevant. Here, however, significant differences are to be found—while the Malaysian government produced strategies similar in structure to M-Blueprint, the Thai government produced a localised version of CEFR. Here, direct comparisons were less feasible.
(c) In each context, a pair of teaching manuals was also created to facilitate the implementation of CEFR. These four texts again bear many similarities and allowed for a direct comparison.

In summary, significant parallels as well as contrasts were observed when examining the data available from both contexts. While differences in how processes of policy development and appropriation unfold in different contexts are natural, a significant departure in this case is in the timelines followed. In the Malaysian case, the policies were developed in a linear sequence, with the top-level strategy published first and the lower-level documents coming after. In Thailand, however, the sequence was inverse, with the teacher manuals being the earliest publications and the overall strategy coming last.[2] The main effect of this was that the texts contrasted in their use of intertextuality—while the Malaysian texts were generally presented as a top-to-bottom chain, the Thai documents were generally independent of each other.

[1] Of the documents presented from the Thai context, *T-Plan* and both teaching manuals were in Thai and were interpreted and analysed with the help of research assistants.

[2] The reasons for this are related to the timelines imposed by previous policy, in this case the expiration of the previous National Education Plan.

Table 1 Overview of analysed policy texts

Policy text genre	Thailand	Malaysia
Macro-level education strategies	2017: National Education Plan (T-Plan)	2013: Malaysia Education Blueprint (M-Blueprint)
Meso-level policy documents specific to English language education	2016: Framework of Reference for English Language Education in Thailand (FRELE-TH)	2015: English Language Education Reform in Malaysia: The Roadmap (M-Roadmap) and English Language Education in Malaysia: An Agenda for Reform (M-Agenda)
Micro-level teaching manuals	2014: Manuals for primary and secondary level (T-PrimMan and T-SecMan)	2016: Manuals for primary and secondary level (M-PrimMan and M-SecMan)

Analysis

The framework guiding the analysis of the policy texts was the discourse historical approach to critical discourse analysis (DHA; see Reisigl and Wodak 2015). In line with DHA, I examined each text with regard to the discursive strategies used within it, by which I refer to "more or less intentional plan[s] of practice […] adopted to achieve a particular social, political, psychological or linguistic goal" (Reisigl and Wodak 2015, p. 33). In this analysis, discursive strategies were described in terms of the different nexuses they formed between particular concepts, seen as particularly meaningful ideas which become associated with discursive struggle and attain considerable mobilizing power (Koselleck 1982, 2004). In this way, the analysis drew on the analytical framework of conceptual history (Ger. *Begriffsgeschichte*; Koselleck 2002) and its previous applications in DHA research on policy (Krzyzanowski 2016; Krzyzanowski and Wodak 2011). In line with *Begriffsgeschichte*, attention was paid to the identification of key concepts in each set of policies and to the establishment of semantic relationships between them. In particular, the following types of relationships were examined:

(a) Key socio-political concepts (Ger. *Grundbegriffe*) which act as ideological cornerstones in a given discourse;
(b) Neighbouring or complementary concepts (Ger. *Nebenbegriffe*) which facilitate the creation of ideological-conceptual networks;
(c) Oppositional or counter-concepts (Ger. *Gegenbegriffe*) which illustrate how the boundaries of ideological-conceptual networks are drawn through exclusion.

While such an approach allows for the identification of key concepts in texts and discourses, it is important to point out that any conceptual relations encoded in text are context-bound and thus likely to shift once that text is transferred to another context, either because the text is thus integrated into discourses where particular concepts have attained a different meaning or because recontextualised concepts are hybridized as they enter into new relationships of complementarity and/or oppositionality (Wodak and Fairclough 2010). Thus, the identification and examination of concepts in this study involved reference to how their meaning was defined by different levels of context. Specifically, the meaningfulness of concepts was examined with regard to (a) intra-textual relations, (b) inter-textual (discursive) relations, (c) features of the situational context in which the text was produced, i.e. the fields of (English) language policy in both Thailand and Malaysia, and (d) the broader socio-political and historical context that determines relations of power in those fields (Wodak 2008). Given the fact that the analysis compared policy environments with often differing features and contrasting histories, such a context-aware approach was vital to avoid oversimplifications or overgeneralisations, with the awareness that the meaningfulness of a concept may vary according to the context in which it is used.

Recontextualisations of CEFR in Thai and Malaysian language policy

Contextualizing CEFR: the bilingual agenda

An important early step in analysing how a globalized language policy is recontextualised into a particular local context is to examine how present-day language policies in that context are embedded into a broader historical trajectory. In the case of Thailand and Malaysia, such an examination foregrounds both points of comparison and difference with regard to how policies have related to the language ecology in either context. A key parallel is the countries' linguistic diversity, with both being highly linguistically complex with regard to the number of distinct languages spoken and in terms of the variation that may be found within larger languages. In Thailand, for instance, a plethora of indigenous and immigrant languages (e.g. Chinese dialects, Pattani Malay, Burmese, Khmer, Hmong, etc.) is spoken alongside Thai, which is itself highly differentiated, with the standardised Central variety, which bears great influence as a result of its symbolic association with the state, being distinct from the regional languages (dialects) spoken in the periphery (Kosonen 2017; Premsrirat 2011; Smalley 1994). A perhaps even greater level of variation may be found in Malaysia, where indigenous languages (such as Sama, Murut and Thai) are outnumbered by languages linked to historical immigration (Chinese dialects, Tamil). At the same time, the national language, Bahasa Malaysia (below: BM), was only standardised in the aftermath of independence in 1957, with its power being relativized by its perceived lack of neutrality from the perspective of non-Malays, in particular the economically powerful Chinese and Tamil communities (Coluzzi 2017). An additional challenge to the dominance of BM is the continued prominence of English (Gill 2014), both in its standard and indigenized variety (known as Colloquial Malaysian English or 'Manglish').

The differing status of the national language and English in the two nations is a consequence of their recent histories and in particular their contrasting positions *vis-à-vis* the British Empire. Malaysia was gradually colonized by the British from the seventeenth century onward and was therefore governed by an English-speaking elite until relatively recently (Pennycook 1994), which meant that there is no long-standing tradition of systematic use of local languages for official purposes (Gill 2014). Thailand, on the other hand, retained political independence throughout this period, being the only nation in the region not to come under either direct British, French or Dutch rule, a fact exploited by local elites for systematic promotion of centralised 'national' authority (Anderson 1998). As this was built on a single national language, little space was left for English to have any official role. However, while this distinction appears clear-cut with regard to Kachru's (1985) outer and expanding circles, it is in fact highly problematic. Anderson (1998) for instance argues that Thailand was de facto colonised from an economic perspective if not politically. Indeed, English has long been part of the language repertoires of the Thai elite despite its unofficial status (Diller 1988; Sukamolson 1998; Wongsothorn et al. 2002). Conversely, while Malaysia would

nominally be classified as an 'ESL' nation in Kachru's model, the diffusion of English among its population was historically low, being limited mainly to locals working in the colonial administration (Pennycook 1994), with little or no proficiency in the language continuing to be typical outside major urban centres (Gill 2014).

Such historical contrasts and similarities provide a broad frame of reference for the interpretation of present-day policies in either context. In Malaysia, the lack of diffusion of English, coupled with tensions between indigenous Malays and immigrant Chinese and Tamils, has produced a number of policy shifts. After independence in 1957, BM was made the national language and gradually replaced English in official functions. While ethnic nationalism thus led to English diminishing in status, recent decades have seen attempts at reversing this policy. In 2002, it was announced that English would replace BM as the medium of instruction in mathematics and science classes at all levels of education, a policy later re-reversed in 2012 after a series of issues hampered its implementation (Gill 2014). This did not, however, signal a shift away from English, as the adoption of a new language policy agenda, named MBMMBI (BM: *Memartabatkan Bahasa Malaysia dan Memperkukuhkan Bahasa Inggeris*, English: 'To Uphold Bahasa Malaysia and to Strengthen English'), was announced in the same period. As suggested by its title, this policy sets as its objective the establishment of a bilingual society, the rationale for which also can be seen in the policies analysed as part of this study:

In general, the Ministry has three goals for the learning of languages:

- Fostering a unique shared identity between Malaysians anchored in the ability to be proficient in the use of a common national language, Bahasa Malaysia;
- Developing individuals that are equipped to work in a globalised economy where the English language is the international language of communication; and
- Providing opportunities to learn an additional language. (*M-Blueprint*, pp. 4–10)

These points are indicative of the language hierarchy constructed by *M-Blueprint*, a broad government strategy published in 2013. Reflective of historical tendencies (see above), BM was positioned at the top, being exclusively granted the label of 'national language', and was associated with concepts like 'identity' and 'uniqueness', as well as 'unity' and 'nation-building' elsewhere in the document. The political and cultural nature of how BM is constructed through these concepts stands in contrast to the way in which English, as the second language in the hierarchy, was positioned. English was nearly exclusively constructed as a language of international (rather than national) communication and a (utilitarian) means of achieving economic success, one which bears no significance to identity or culture. At the bottom of the hierarchy is the open-ended category of 'additional language', which in fact consisted of a mix of different languages. In some parts of the text, the label was used to refer to the languages of other

ethnic groups, in particular Chinese and Tamil. Elsewhere, however, this term was also used to refer to the learning of other world languages such as Spanish, French and Japanese, while also being used to refer to languages of neighbouring countries. The vague nature of this term reflects the extent to which languages other than BM and English are backgrounded throughout the document, a position compounded by references to 'bilingual proficiency' as an objective in which no space at all is left for other languages.

Just as in Malaysia, the position of English in language policy in Thailand can be seen to have undergone a series of shifts. While English has been present in Thai education in various ways since the nineteenth century, it has drifted in and out of favour since the 1950s, being at times a required subject for all and at others an elective alongside other European languages, such as French or German, or Asian languages, such as Chinese and Japanese, with Arabic and Pali also offered (Sukamolson 1998). Broadly, however, policy in recent decades have seen the importance of English increase to the extent that it is now de facto the only foreign language most students learn (Baker and Jarunthawatchai 2017). The policies examined in this study reflected this orientation, with English education being extensively referenced as a key objective:

> The governmental organizations and related sectors must therefore help develop frameworks and direct the production and development of human resources in different fields so that the country will have the right people for the right job in the market for national development. The curricula for different levels that can give learners skills needed for the 21st century world, especially English, science, and digital skills, should be improved. (*T-Plan*, p. 100)

> Support people of all ages to be able to read and write Thai, their indigenous languages and languages of neighbouring countries. (*T-Plan*, p. 98)

These extracts illustrate the association between languages and concepts in *T-Plan*, an educational strategy published by the Thai government in 2017. The pattern most evident is the relatively intensive conceptualization of English, which featured prominently throughout the document, with conceptual associations often mirroring those observed in *M-Blueprint*. In particular, English was again exclusively constructed through an economic prism, being positioned in this case as one of the 'twenty-first century skills' central to the 'development of human resources' and being associated with the highly influential concept of 'national development' (Hill and Fujita 2012) and elsewhere in the text with the concept of 'Thailand 4.0', central to the economic policy of the government under which the policy had been created (Jones and Pimdee 2017). This stands in contrast to the few references made to other languages, which also saw few conceptual associations. Thai saw little attention in the policy, a clear contrast from the intensive conceptualization of BM in *M-Blueprint*, though this disparity can be seen as a reflection of the unchallenged status of Thai as the national language when compared to BM, which is not only a relatively newly standardised variety but also a code embedded in ethnic struggle (see above). However, the lack of attention to other languages used in and around Thailand, referred to

generically throughout the policy, is a reflection of their generally low status (Premsrirat 2011), indicating the existence of an implicit bilingual agenda similar to that explicitly adopted in Malaysia.

Recontextualizing CEFR: a selective endeavour

It is as an integral part of the 'national language plus English' bilingual agenda that CEFR was recontextualised into both the Thai and Malaysian context. In both sets of documents, CEFR was primarily positioned as an instrument of English language teaching and learning and was generally not associated with learning other languages. As outlined in the previous section, the focus on positioning English as the dominant second language to be taught in schools is an established part of Thai and Malaysian language policy and also broadly resonates with tendencies observed in the region (e.g. Baldauf et al. 2011; Kirkpatrick 2010, 2017). Comparing this orientation to CEFR, however, indicates a significant reconceptualization, since not only is the framework not intended to be tied to a particular language, it includes a remark that a way of achieving multilingualism might involve "reducing the dominant position of English in international communication" (Council of Europe 2001, p. 4).

Though such a tight conceptual nexus between CEFR and English may appear potentially contradictory when considering the plurilingual/pluricultural agenda promoted by the Council of Europe, both Thai and Malaysian documents avoided such incompatibility by selectively recontextualizing parts of the framework. As discussed above, CEFR has been critiqued for, among other reasons, failing to completely integrate the two key concepts of plurilingualism and pluriculturalism into its different elements. Indeed, among its chapters, only the introduction (pp. 1–8) and those outlining the framework's broad vision of language learning and teaching (pp. 131–156) and the development of curricula (pp. 168–176) consistently reference plurilingualism/pluriculturalism, whereas the concepts rarely, if at all, appear in the chapters where reference levels are described (pp. 43–130) and where task-based teaching (pp. 157–167) and assessment (pp. 177–196) are described. The latter elements, in particular the reference levels, are instead based on a monolingual construct of language proficiency (Shohamy 2011).

When considering the recontextualization of CEFR in terms of how these different elements were foregrounded and backgrounded as part of its transfer into the Thai and Malaysian contexts, a clear pattern emerges. In the policies examined, the CEFR reference levels were often the only element referred to explicitly while other elements of the framework were generally backgrounded, with plurilingualism and pluriculturalism receiving no mention in any of the documents examined. The Malaysian English language education strategy *M-Agenda*, for instance, made no reference to the Council of Europe nor to its language policy, with CEFR instead typically positioned as simply 'international':

> What we need to create is a programme that is simultaneously international, because it is aligned to international standards as represented by the CEFR,

> and national, because it is carefully tailored to the specific needs of Malaysia. (*M-Agenda*, p. 14)

This quote is significant because it illustrates the extent to which the foregrounding of the levels was naturalized in these documents, with the meaning of 'the CEFR' here implicitly being narrowed down only to its reference levels, since it is to these that a test or curriculum may be seen to be 'aligned'. Such a focus on the CEFR levels was also evident elsewhere in the Malaysian data: In *M-Blueprint*, for instance, parts of CEFR pertaining to the reference levels were wholly recontextualised despite the relatively broad nature of the document, with CEFR-specific terminology ('operational' and 'independent proficiency') used when setting objectives. Such references to CEFR-specific language were also common, with both M-Agenda and M-Roadmap including numerous references to the 'can do' approach of the framework. In documents more specifically aimed at implementation at the local scale, namely the two sets of teaching manuals, CEFR was referred to in more detail yet again narrowed in scope to only its six reference levels. Thus, Thai teaching manuals (*T-PrimMan* and *T-SecMan*) included copious reference to both the global level descriptions of CEFR and the more specific skill-by-skill 'illustrative' descriptions, which were presented in translation according to the level dictated by policy as being relevant (e.g. A1 for final year of primary school). Similar patterns of recontextualising CEFR may also be found in Malaysian documents aimed at teachers (*M-PrimMan* and *M-SecMan*).

The focus on levels is also evident when considering *FRELE-TH*, a unique text in this dataset in that it does not constitute a policy drawing on CEFR but is instead a localised version of the framework produced for use in Thailand. Thus, its main equivalent is the only other 'national' CEFR adaptation at the time, the Japanese CEFR-J, to which *FRELE-TH* also bears greatest similarity in the sense that its key feature is a vertical expansion of the levels. In CEFR-J, this expansion created 12-levels in place of the original six, which was achieved by creating a new bottom level (Pre-A1) and splitting remaining levels (A1 into three sub-levels and A2, B1 and B2 into two sub-levels each). A similar approach was taken in *FRELE-TH*, in which the original six levels have been expanded to ten with the addition of four so-called 'plus-levels' (i.e., A1+, A2+, B1+, B2+).[3] *FRELE-TH* thus features revised versions of 33 reference level descriptions—all those in the original CEFR and one additional ('Reading Literature') drawn from the EAQUALS Bank of Descriptors. It also includes a level-by-level description of language and content topics appropriate for each level (developed on the basis of the Core Inventory for General English) and a vocabulary database (developed from the Word Family Framework). However, while all these resources mark it as a significant adaptation, it is notable that they are all tied to the levels, which were in the case of *FRELE-TH* presented largely in isolation, with the brief accompanying text again making no reference to plurilingualism or pluriculturalism.

[3] The reason for such an expansion was a perception that the bottom range of the original CEFR, where most Thai speakers of English are seen to be concentrated, was insufficiently detailed to provide useful background information (for a presentation of this adapted version by its developers, see Hiranburana et al. 2018), a motivation similar to that referred to by the authors of CEFR-J (Negishi 2012).

Table 2 Juxtaposition of CEFR and PISA (*T-Plan*, pp. 83–84)

Indicators	At present	Year 1–5	Year 6–10	Year 11–15	Year 16–20
Quality					
(4) Average score on the Program for International Student Assessment (PISA) for 15-year-old students in reading, mathematics and science	421/409/415	500	510	520	530
(5) English proficiency of those who finish school at each level is higher after they are assessed by standardized test (CEFR) (junior high school level/high school level/undergraduate level)	A1/A2/B2	A1/A2/B2	A2/B1/B2^{+}	B1/B1^{+}/C1	B2/B2/C1^{+}

A key feature of how CEFR was integrated into the examined policies was thus the decontextualisation of its six levels, with key accompanying concepts seeing little uptake among policymakers. At the same time, however, CEFR was also placed into new conceptual relationships. Above, I outlined the broad language policy agenda into which they CEFR was integrated, one which sets societal bilingualism in the national language and English as its main objective. However, alongside its contextualization in a discourse about language policy, CEFR was also embedded into a broader discourse about education, one not necessarily centred on language. The key characteristic of this discourse, evident in *T-Plan* and *M-Blueprint* as the two overarching strategies, was their reference to instruments of transnational governance in education like the Programme for International Student Assessment (PISA), administered by the Organization for Economic Cooperation and Development (OECD). The policy impact such instruments, PISA in particular, have at the national level is significant (Grek 2009) and it is thus unsurprising that league tables based on them featured prominently in both *T-Plan* and *M-Blueprint*. In both cases, aspirational goals were also set, supported not by context-specific analyses but largely on transnational comparisons. Thus, *M-Blueprint* highlighted traditionally high-performing nations like Singapore and South Korea, describing them as those which 'Malaysia seeks to compete against in today's knowledge economy' (pp. 3–9).

Such comparisons also played a key role in *T-Plan* and it is here that an extreme example of how CEFR fits into this PISA-centric discourse may be found. Table 2 presents part of a section in which key indicators of 'Quality' were presented in reference to the 20-year plan outlined by the document. Here, future policy targets were set according to 5-year periods, with expectations regarding students' English ability expected to progressively rise until the target proficiencies were set at B2 for both junior and senior high school students and C1+ for university students.[4] What is most significant about

[4] This reference is likely an error since CEFR does not include a level C1+. Such 'plus levels' are used to represent half-way points between levels of proficiency (i.e. B1+ is an intermediate level between B1 and B2) but have only been described below the two highest levels (C1 and C2). For more information, see North (2014).

this table, however, is the juxtaposition of CEFR and PISA, which is an indicator of how CEFR levels may be interpreted when combined with elements of a global discourse about education. When juxtaposed with PISA, CEFR is positioned as an objective international standard which drives policy change by forcing comparisons between 'high-' and 'low-performing' educational systems on the basis of the CEFR levels they set. While this is an extreme example unique to *T-Plan*, it should be remarked that the consistent positioning of CEFR as an 'international standard' to which an educational system must be 'aligned' (exemplified by the above extract from *M-Agenda* but ubiquitous across the data-set) can also be broadly seen as reflective of such a perception.

CEFR in the hands of global and local agendas: ideological struggle or symbiosis?

As indicated in the previous sections, the recontextualization of CEFR in Thai and Malaysian language policy often followed parallel trajectories. Such textual trajectories are, as pointed out, determined by a variety of factors, such as the balance of power between actors in a particular nexus of practices or by the ideological struggles that characterise many discourses. Such struggles could also be observed when interrogating the ideologies underlying the interpretation of CEFR in Thailand and Malaysia, with two ideologies in particular vying for dominance.

Through its association with English, CEFR could be seen as a vehicle of neoliberal ideology in language policy, one which views language teaching and learning, and indeed all education, as an instrument of workforce production and profit (Ng 2018). In such a neoliberal imaginary, individuals positioned as agents whose language learning goals are driven by their quest for achievement and their need to remain competitive in a job market defined by flexibility and mobility (Block et al. 2012; Flores 2014; Kubota 2014, 2016). This orientation can be seen with regard to how English was positioned in the Thai and Malaysian policies above, where it is consistently presented in close association with concepts from the economic field such as 'employability' and 'skills', being in this way clearly differentiated from other languages in its ecology. The construction of such language hierarchies is a key part of such a neoliberal agenda in language policy, with languages with perceived high value for employability being prioritised over those which are deemed to have low value (Lorente and Tupas 2013). Most often, the 'high value' language favoured by such an agenda is English, with local languages in contrast facing exclusion (e.g. Kirkpatrick 2017).

A further characteristic of neoliberalism in language policy is its focus on the transnational scale, particularly on the establishment of transnational regimes of governance (De Costa et al. 2019). Here, CEFR plays a key role, since it does not only mediate local developmental agendas but has also been appropriated by powerful global institutions which seek to profit from its implementation. Here, I refer to institutions like Cambridge English and the British Council, which are seen by Phillipson (2010) as agents of 'linguistic imperialism' (see also Pennycook 1994) and which Block et al. (2012) collectively refer to as the 'ELT industry'. As outlined above, CEFR has seen increasing use by the ELT industry in the production and

marketing of tests and textbooks, and it is in part through these products that the conceptual nexus between CEFR and English has been reinforced. Indeed, the close association between the ELT industry and CEFR has had practical implications for how the framework has been implemented at the local level in Thailand and Malaysia. In the former, the British Council has run workshops in support of the rollout of the framework since 2014. In the case of the latter, the implementation of CEFR has had an even more pronounced effect, with locally-developed materials being replaced in 2017 by CEFR-aligned global textbooks—produced by MacMillan and Cambridge, the latter of whom had previously conducted an influential CEFR-based assessment of the Malaysian educational system (Cambridge English 2013).

While CEFR may thus be seen purely as an instrument of imperialism, a conclusion potentially reinforced by the way in which the policies examined in this research juxtapose the framework to other instruments of transnational governance like PISA, such a purely structural view only tells part of the story. In particular, it ignores the way in which CEFR also perpetuates local nationalist agendas by solidifying the position of English as a second language and, by extension, pushing codes other than the national language (Thai or BM) further down the pecking order. As argued by Kirkpatrick (2017), it is this rearrangement of the linguistic hierarchy in favour of the national language and to the detriment of local languages that is a key consequence of the growing focus on English in South East Asia. This indicates that while the transnational focus of neoliberalism means that it is often pitted in opposition with nationalism, being seen as endangering a community's homogeneity or indigeneity, the agendas of both ideologies may also co-exist and develop synergies.

The ability of CEFR to facilitate such a synergy in the Thai and Malaysian contexts may be seen as a logical extension of the framework's previous trajectory. While CEFR may have been produced under the auspices of the Council of Europe, thus representing the humanist agenda of that organization, its relatively rapid adoption by the European Union meant that it was soon embedded into a much more complex discourse, one in which the cultural agenda in language policy is often marginalized in the face of neoliberal arguments focussed on promoting multilingualism as a means of economic mobility (Krzyzanowski and Wodak 2011). At the same time, the 'European' legacy of CEFR is also rooted in local concerns over the influx of migrants, with the motivation underlying the development of its earliest precursor in the 1970s being directly linked to nationalist calls for the linguistic integration of migrants (Tabouret-Keller 1991). It is notable that the framework has until today also continued to mediate nationalist agendas in the European context, being regularly used as a means of gatekeeping in citizenship testing, for example (Extra et al. 2009).

In conclusion, what does this trajectory tell us about the nature of CEFR and other globalized language policies? Perhaps their defining characteristic is that they are either by nature able to mediate multiple agendas at different scales or are sufficiently open-ended that actors are able to twist them according to their needs. In this respect, globalized language policies come close to what Holland terms foundational documents, texts which 'are essential in defining larger religious, political and social agendas' (2014, p. 386). By using the example of *Quotations from Chairman Mao Tsetung*, also known as *The Little Red Book*, Holland shows how

the broad social impact of such documents is, while intertwined with their contents, also heavily dependent upon the ways in which particular fragments are recontextualised. This study suggests that a similar approach may be taken when examining globalized language policies, with the contents of the text playing as crucial a role in determining their meaningfulness as the ways in which the text (or fragments of it) are recontextualised and incorporated into new texts and discourses.

Acknowledgements Funding was provided by Thailand Research Fund (Grant No. MRG6080168). I would also like to express my gratitude to Professor Ruth Wodak and Dr Chonlada Laohawiriyanon for the advice they provided during the course of this research. Any errors or inaccuracies are, of course, my own responsibility.

References

Anderson, B. (1998). *The spectre of comparisons: Nationalism, Southeast Asia and the World*. London: Verso.

Appadurai, A. (1990). Disjuncture and difference in the global cultural economy. In M. Featherstone (Ed.), *Global culture: Nationalism, globalization and modernity* (pp. 295–310). London: Sage.

Arnott, S., Brogden, L. M., Faez, F., Péguret, M., Piccardo, E., Rehner, K., et al. (2017). The Common European Framework of Reference (CEFR) in Canada: A research agenda. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée, 20*(1), 31–54.

Baker, W., & Jarunthawatchai, K. (2017). English language policy in Thailand. *European Journal of Language Policy, 9*(1), 27–44.

Baldauf, R., Kaplan, R., Kamwangamalu, N., & Bryant, P. (2011). Success or failure of primary second/foreign language programmes in Asia: What do the data tell us*?*. *Current Issues in Language Planning, 12*(2), 309–323.

Barakos, E., & Unger, J. W. (Eds.). (2016). *Discursive approaches to language policy*. Basingstoke: Palgrave.

Block, D., Gray, J., & Holborow, M. (2012). *Neoliberalism and applied linguistics*. London: Routledge.

Blommaert, J. (2007). Sociolinguistic scales. *Intercultural Pragmatics, 4*(1), 1–19.

Byram, M., & Parmenter, L. (Eds.). (2012). *The Common European Framework of Reference: The globalization of language education policy*. Bristol: Multilingual Matters.

Cambridge English. (2013). *Cambridge baseline 2013: English language in Malaysian schools: Results report*.

Canagarajah, S. (2018). The unit and focus of analysis in lingua franca English interactions: In search of a method. *International Journal of Bilingual Education and Bilingualism, 21*(7), 805–824.

Coluzzi, P. (2017). Language planning for Malay in Malaysia: A case of failure or success? *International Journal of the Sociology of Language, 244,* 17–38.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Council of Europe. (2018). *Common European framework of reference for languages: Learning, teaching, assessment. Companion volume with new descriptors*. Strasbourg: Council of Europe.

De Costa, P. I., Park, J., & Wee, L. (2019). Linguistic entrepreneurship as affective regime: Organizations, audit culture, and second/foreign language education policy. *Language Policy, 18*(3), 387–406.

Diller, A. (1988). Thai syntax and "National Grammar". *Language Sciences, 10*(2), 273–312.

European Council. (2002). *Presidency conclusions: Barcelona European Council, 15 and 16 March 2002*. SN 100/1/02 REV1.

Extra, G., Spotti, M., & Van Avermaet, P. (2009). Testing regimes for newcomers. In G. Extra, M. Spotti, & P. Van Avermaet (Eds.), *Language testing, migration and citizenship: Cross-national perspectives on integration regimes* (pp. 1–34). London: Continuum.

Flores, N. (2014). The unexamined relationship between neoliberalism and plurilingualism: A cautionary tale. *TESOL Quarterly, 47*(3), 500–521.

García, O. (2009). Education, multilingualism and translanguaging in the 21st century. In A. Mohanty, M. Panda, R. Phillipson, & T. Skutnabb-Kangas (Eds.), *Multilingual education for social justice: Globalising the local* (pp. 128–145). Orient Blackswan: New Delhi.

Gill, S. K. (2014). *Language policy challenges in multi-ethnic Malaysia*. Amsterdam: Springer.

Grek, S. (2009). Governing by numbers: The PISA 'effect' in Europe. *Journal of Education Policy, 24*(1), 23–37.

Hamid, M. O. (2014). World Englishes in international proficiency tests. *World Englishes, 33*(2), 263–277.

Heller, M. (2011). *Paths to post-nationalism: A critical ethnography of language and identity*. New York: Oxford University Press.

Hill, R. C., & Fujita, K. (2012). "Detroit of the East": A multiscalar case study of regional development policy in Thailand. In B.-G. Park, R. C. Hill, & A. Saito (Eds.), *Locating neoliberalism in East Asia: Neoliberalizing spaces in developmental states* (pp. 257–293). Maiden, MA: Wiley-Blackwell.

Hiranburana, K., Subphadoongchone, P., Tangkiengsirisin, S., Phoocharoensil, S., Gainey, J., Thogsngsri, J., et al. (2018). A framework of reference for English language education in Thailand (FRELE-TH)—Based on the CEFR. The Thai experience. *LEARN Journal, 10*(2), 90–119.

Holland, J. (2014). Narrative fidelity to the Little Red Book in the framing efforts of the Red Guard Movement: A theoretical model for foundational documents. *Discourse & Society, 25*(3), 383–401.

Holzinger, K., & Knill, C. (2005). Causes and conditions of cross-national policy convergence. *Journal of European Public Policy, 12*(5), 775–796.

Hornberger, N. H. (2005). Opening and filling up implementational and ideological spaces in heritage language education. *Modern Language Journal, 89*(4), 605–609.

Hult, F. (2010). Analysis of language policy discourses across the scales of space and time. *International Journal of the Sociology of Language, 202,* 7–24.

Jessop, B. (2007). *State power: A strategic-relational approach*. Cambridge: Polity Press.

Jin, Y., Wu, Z., Alderson, J. C., & Song, W. (2017). Developing the China Standards of English: Challenges at macropolitical and micropolitical levels. *Language Testing in Asia, 7*(12), 1.

Johnson, D. C. (2013). Positioning the language policy arbiter: Governmentality and footing in the School District of Philadelphia. In J. W. Tollefson (Ed.), *Language policies in education: Critical issues* (2nd ed., pp. 116–136). London: Routledge.

Jones, C., & Pimdee, P. (2017). Innovative ideas: Thailand 4.0 and the fourth industrial revolution. *Asian International Journal of Social Sciences, 17*(1), 4–35.

Jones, N., & Savile, N. (2009). European language policy: Assessment, learning, and the CEFR. *Annual Review of Applied Linguistics, 29,* 51–63.

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: the English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), *English in the world: Teaching and learning the language and literatures*. Cambridge: Cambridge University Press.

Kirkpatrick, A. (2010). *English as a Lingua Franca in ASEAN: A multilingual model*. Hong Kong: Hong Kong University Press.

Kirkpatrick, A. (2017). Language education policy among the Association of Southeast Asian Nations (ASEAN). *European Journal of Language Policy, 9*(1), 7–25.

Koselleck, R. (1982). Begriffsgeschichte and social history. *Economy and Society, 11,* 409–427.

Koselleck, R. (2002). *The practice of conceptual history. Timing history, spacing concepts*. Palo Alto, CA: Stanford University Press.

Koselleck, R. (2004). *Futures past*. New York: Columbia University Press.

Kosonen, K. (2017). *Language of instruction in Southeast Asia: Background paper prepared for the 2017–18 Global Education Monitoring Report*.

Krzyzanowski, M. (2016). Recontextualisation of neoliberalism and the increasingly conceptual nature of discourse: Challenges for critical discourse studies. *Discourse & Society, 27*(3), 308–321.

Krzyzanowski, M., & Wodak, R. (2009). Theorising and analysing social change in Central and Eastern Europe: The contribution of critical discourse analysis. In A. Galasinska & M. Krzyzanowski (Eds.), *Discourse and transformation in Central and Eastern Europe* (pp. 17–40). Basingstoke: Palgrave.

Krzyżanowski, M., & Wodak, R. (2011). Political strategies and language policies: The European Union Lisbon strategy and its implications for the EU's language and multilingualism policy. *Language Policy, 10*(2), 115–136.

Kubota, R. (2014). The multi/plural turn, postcolonial theory, and neoliberal multiculturalism: Complicities and implications for applied linguistics. *Applied Linguistics, 37*(4), 474–494.

Kubota, R. (2016). Neoliberal paradoxes of language learning: Xenophobia and international communication. *Journal of Multilingual and Multicultural Development, 37*(5), 467–480.

Kulsiri, S. (2006). *A critical analysis of the 2001 national foreign language standards-based curriculum in the Thai school system*. Unpublished Ph.D. thesis. Canberra: University of Canberra.
Littlejohn, A. (2012). Language teaching materials and the (very) big picture. *Electronic Journal of Foreign Language Teaching, 9*(1), 283–297.
Lorente, B., & Tupas, R. (2013). (Un)emancipatory hybridity: Selling English in an unequal world. In R. Rubdy & L. Alsagoff (Eds.), *The global-local interface and hybridity: Exploring language and identity* (pp. 66–82). Bristol: Multilingual Matters.
Martyniuk, W. (Ed.). (2010). *Aligning Tests with the CEFR: Reflections on using the Council of Europe's draft Manual*. Cambridge: Cambridge University Press.
Maybin, J. (2017). Textual trajectories: Theoretical roots and institutional consequences. *Text & Talk, 37*(4), 415–435.
Negishi, M. (2012). *The development of the CEFR-J: Where we are, where we are going. Grant-in-aid for Scientific Research Report*. Retrieved from www.tufs.ac.jp/common/fs/ilr/EU_kaken/_userdata/negishi2.pdf. Accessed 18 Oct 2019.
Ng, C. J. W. (2018). Skilling the nation, empowering the citizen: Neoliberal instantiations in Singapore's lifelong learning policy. *Journal of Language and Politics, 17*(1), 118–140.
North, B. (Ed.). (1992). *Transparency and coherence in language learning in Europe: Objectives, evaluation, certification*. Strasbourg: Council for Cultural Cooperation.
North, B. (2014). *The CEFR in practice*. Cambridge: Cambridge University Press.
Pennycook, A. (1994). *The cultural politics of English as an international language*. New York: Longman.
Phillipson, R. (2010). *Linguistic imperialism continued*. London: Routledge.
Piccardo, E. (2010). From communicative to action-oriented: New perspectives for a new millennium. *CONTACT TESL, 36*(2), 20–35.
Piccardo, E. (2013). Plurilingualism and curriculum design: Toward a synergic vision. *TESOL Quarterly, 47*(3), 600–613.
Piccardo, E. (2014). The impact of the CEFR on Canada's linguistic plurality: A space for heritage languages? In P. Trifonas & T. Aravossitas (Eds.), *Rethinking heritage language education* (pp. 45–65). Cambridge: Cambridge University Press.
Pietikäinen, A. (2010). Sami language mobility: Scales and discourses of multilingualism in a polycentric environment. *International Journal of the Sociology of Language, 202,* 79–101.
Pilkinton-Pihko, D. (2013). *English-medium instruction: Seeking assessment criteria for spoken professional English*. Unpublished Ph.D. thesis. Helsinki: University of Helsinki.
Premsrirat, S. (2011). Redefining "Thainess": Embracing diversity, preserving unity. *The Journal of the Royal Institute of Thailand, 3,* 54–75.
Reisigl, M., & Wodak, R. (2015). The discourse-historical approach (DHA). In R. Wodak & M. Meyer (Eds.), *Methods of critical discourse studies* (pp. 23–61). London: SAGE.
Roger, C., & Dauvergne, P. (2016). The rise of transnational governance as a field of study. *International Studies Review, 18*(3), 415–437.
Rutkowski, D. (2007). Converging us softly: How intergovernmental organizations promote neoliberal educational policy. *Critical Studies in Education, 48*(2), 229–247.
Savski, K. (2016). State language policy in time and space: Meaning, transformation, recontextualisation. In E. Barakos & J. W. Unger (Eds.), *Discursive approaches to language policy* (pp. 51–70). Basingstoke: Palgrave.
Savski, K. (2018). The roles of field and capital in negotiating language policy in the Slovene parliament. *Journal of Language and Politics, 17*(1), 24–45.
Savski, K. (2019). Putting the plurilingual/pluricultural back into CEFR: Reflecting on policy reform in Thailand and Malaysia. *The Journal of Asia TEFL, 16*(2), 644–652.
Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *Modern Language Journal, 95*(3), 418–429.
Smalley, W. (1994). *Linguistic diversity and national unity: Language ecology in Thailand*. Chicago: University of Chicago Press.
Sukamolson, S. (1998). English language education policy in Thailand. *Asian Englishes, 1*(1), 68–91.
Tabouret-Keller, A. (1991). Factors of constraints and freedom in setting a language policy for the European Community: A sociolinguistic approach. In F. Coulmas (Ed.), *A language policy for the European Community: Prospects and quandaries* (pp. 45–58). Berlin: De Gruyter.
Trim, J. L. (2012). The Common European framework of reference for languages and its background: A case study of cultural politics and educational influences. In M. Byram & L. Parmenter (Eds.), *The Common*

European framework of reference for languages: The globalisation of language education policy (pp. 14–35). Bristol: Multilingual Matters.

Van Ek, J. A. (1975). *The threshold level for modern language learning in schools*. Strasbourg: The Council of Europe.

Wodak, R. (2008). Introduction: Discourse studies—Important concepts and terms. In M. Krzyżanowski & R. Wodak (Eds.), *Qualitative discourse analysis in the social sciences* (pp. 1–29). Basingstoke: Palgrave Macmillan.

Wodak, R., & Fairclough, N. (2010). Recontextualizing European higher education policies: The cases of Austria and Romania. *Critical Discourse Studies, 7*(1), 19–40.

Wodak, R., & Savski, K. (2018). Critical discourse-ethnographic approaches to language policy. In J. Tollefson & M. Pérez-Milans (Eds.), *The Oxford handbook of language policy and planning* (pp. 93–112). Oxford: Oxford University Press.

Wongsothorn, A., Hiranburana, K., & Chinnawongs, S. (2002). English language teaching in Thailand today. *Asia Pacific Journal of Education, 22*(2), 107–116.

Kristof Savski is presently a Lecturer in English at Prince of Songkla University in Hat Yai, Thailand, and holds a PhD in Linguistics from Lancaster University in the UK. His research contributes to current debates in language policy, sociolinguistics, critical discourse studies and TESOL, and has been published in several major journals in those fields.

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Policy change implementation: the case of the CEFR in Malaysian ESL classrooms / C. Alih, N. A., Abdul Raof, A. H., & Md. Yusof, M. A.**
Source	***Journal of Nusantara Studies (JONUS)*** **Volume 6 Issue 2 (2021) Pages 296–317** **https://doi.org/10.24200/JONUS.VOL6ISS2PP296-317** **(Database: UniSZA Press)**

Journal of Nusantara Studies 2021, Vol 6(2) 296-317
Universiti Sultan Zainal Abidin
ISSN 0127-9386 (Online)
http://dx.doi.org/10.24200/jonus.vol6iss2pp296-317

POLICY CHANGE IMPLEMENTATION: THE CASE OF THE CEFR IN MALAYSIAN ESL CLASSROOMS

Nur Ashiquin C. Alih, Abdul Halim Abdul Raof & Masdinah Alauyah Md. Yusof

Language Academy, Faculty of Social Sciences and Humanities,
Universiti Teknologi Malaysia, 81300 Johor Bahru, Johor, Malaysia.
*Corresponding author: m-halim@utm.my

Received: 29.03.2021 Accepted: 01.06.2021

ABSTRACT

Background and purpose: With rapid global development happening in the world today, the field of education has been awash with various change forces. In an effort to ensure its system is globally competitive, Malaysia has taken a step to align its English language education system to the Common European Framework of Reference (CEFR). However, any change in policy is not without problems and challenges as studies have shown that challenges were inevitable and stakeholders' buy-in on the new change is necessary. Hence, this study was undertaken to investigate the challenges faced by teachers in implementing the CEFR in Malaysian ESL classrooms, their belief on this new reform and their readiness to implement the CEFR.

Methodology: A mixed-method design was utilized with the use of questionnaire and semi structured interview as means for data collection. 365 English language teachers responded to the questionnaire while 15 English language teachers participated in interview sessions.

Findings: The data revealed five challenges namely teachers' motivation, materials, time, students' proficiency level and facilities. Despite facing multiple challenges, all teachers exhibited positive belief towards the adoption of the CEFR. The study also showed that although teachers were emotionally ready to accept the change, their cognitive readiness for change however is dependent upon three important facets namely time, collective effort and adequate materials.

Contributions: The study provides insights into the change in policy involving the CEFR implementation in the Malaysian English language education system which contributes to the scarce literature on the implementation of CEFR.

Keywords: Policy implementation, challenges in policy change, teachers' belief, Common European Framework of Reference (CEFR), readiness for change.

Cite as: C. Alih, N. A., Abdul Raof, A. H., & Md. Yusof, M. A. (2021). Policy change implementation: The case of the CEFR in Malaysian ESL classrooms. *Journal of Nusantara Studies, 6*(2), 296-317. http://dx.doi.org/10.24200/jonus.vol6iss2pp296-317

1.0 INTRODUCTION

The rapid global development has made it almost impossible for many organizations including the educational sector not to change the way they operate (Armenakis & Harris, 2009). Changes in the structural-functional of the educational system have to be made to keep up with various change forces (Kondakci, Beycioglu, Sincar, & Ugurlu, 2016). Malaysia, like any other countries in the world has undertaken numerous efforts to reform its education system for better outcome. In the span of 30 years, Malaysia has undergone at least three major reforms in its English education system (Azman, 2016) with the Common European Framework of Reference (CEFR) as stated in the Roadmap 2015-2025 being one of the initiatives taken by the Malaysian Ministry of Education as a stepping stone to ensure that the English language education system is globally competitive (Ministry of Education Malaysia, 2015).

The Common European Framework of Reference (CEFR) was developed by the Council of Europe in 2001 to assist language practitioners including language learners in identifying the direction of language learning and provides a means for reflection of what learners have to achieve in terms of language outcome and how they intend to achieve it (Council of Europe, 2001). Additionally, the CEFR also provides a basis for language certification and assists in planning and executing language program as well as eases language learners in self-directed learning in terms of raising the learners' awareness on their present knowledge of the language, self-setting objectives, self-assessment and selection of materials for their individual learning purpose. The framework is well-known for its six-level descriptors which provide users with detailed statements of what learners can do at each level which are known as the CEFR "can do" statements for listening, speaking, reading and writing. The descriptors categorizes language learners into three main groups based on their language ability

with each group comprising two levels: Proficient users (levels C1 & C2), Independent users (levels B1 & B2) and Basic users (levels A1 & A2).

2.0 LITERATURE REVIEW

2.1 Policy Changes in Malaysia

English language has long been a salient facet in the education system in our country. Since its inclusion in Malaysian education, English language education has had three major reforms (Azman, 2016). After the National Education Policy was reviewed by the Ministry of Education, the first initiative was introduced in 1982 with the introduction of the Integrated English Language Syllabus for Primary Schools (KBSR) and the Integrated English Language Syllabus for Secondary Schools (KBSM) with the integration of Communicative Language Teaching (CLT) as opposed to the focus on grammatical knowledge. However, the impact of this reform was found to taper off when differing results were received in terms of teaching in communicative way, mismatch between the objectives of the syllabus and CLT principles with the actual classroom practices as well as language assessment (Che Musa, Lie, & Azman, 2012). The second initiative was introduced in 2002, the Standard English Language Curriculum for Primary School (KSSR), aiming to enhance the CLT through School-based Assessment (SBA). The teaching of Mathematics and Science in English language (PPSMI) also happened during this time. However, this reform also faded out due to similar rejection as the previous reform. The recent move introduced was the English Language Education Roadmap 2015-2025 with the major notion to bring the English language in Malaysian education on par with the international level, benchmarked against a standard used by many different countries around the world -the Common European Framework of Reference (CEFR).

2.2 The Utilization of the CEFR Around the World

There is a growing trend in using the CEFR in the world for various purposes (Van Huy & Hamid, 2015; Figueras, 2012). Because of this, researchers in different countries around the world have devoted much effort to examine the use of this framework. For instance, a study which investigated the use of the CEFR in the European education system in examination, curriculum development, school books and teacher training was carried out by Broek and Ende (2013). The study pointed out evidence of close reference between the CEFR and elements in the education system namely the general approach to language learning, materials used as well as teacher training. Despite this, it was revealed that the links between the framework and language assessment however appeared weak as compared to the other elements. In using the

CEFR for benchmarking purposes, a study by Buckland (2010) in the Wall Street Institute (WSI) of Spain pointed a close match between the WSI levels and the CEFR can-do statements with a recorded correlation value of 80% and concluded that alignment between the WSI and CEFR was permissible.

The CEFR has also been used for teaching and learning purposes. Maldina (2015) who investigated the CEFR role and Sociolinguistics and Pragmatics (SP) in foreign language teaching in a high school in Italy found that the CEFR had somehow indirectly shaped teachers' curriculum development and instructional practices. However, this was highly dependent on two elements namely the external language examinations and textbooks used in schools. The study also indicated a weak relationship between the CEFR and teachers' practice in school which is believed due to the lack of understanding about the CEFR.

In a study in Japan, the CEFR was used as a means to develop EFL learners' communicative competence through task completion (Nakatani, 2012). The findings indicated that there was a significant improvement in the learners' post-conversation result. Although Nakatani believed that this improvement might be the result of the strategy training that the learners had undertaken and were aware of, the CEFR could still be viewed as a significant medium in improving learners' communication, nonetheless.

Although the CEFR was developed to serve as a means for synchronizing the language teaching, learning and assessment (Fulcher, 2004), in a later study Fulcher (2010) revealed that using the CEFR merely for standard-based assessment had become a popular trend in the world English education system which has out shadowed the initial purpose of the CEFR. For instance, in using the CEFR for standardization purpose, Lowie, Hainesa, and Jansmaa (2010) undertook a study where a standardized procedure was embedded in writing assessment in the academic context. Their project has demonstrated a general agreement of the different components of writing and the CEFR which implies the feasibility of standardization procedures within the CEFR. Additionally, for the purpose of using the CEFR for validation of local rating scales, Harsch and Martin (2012) did a study aimed at examining whether alignment of a local rating scale to the CEFR was possible. The study indicated that although the approach was not economical in terms of time and resources it demanded, the adaptation of the CEFR descriptors for local context rating purpose was permissible. In another study, Bérešová (2011) investigated the possibility of linking the national examination in English to the CEFR in Slovakia and found that the linking helped initiate strategic actions in developing three major areas in education namely the standard, evaluation and professional. The CEFR is also used as a benchmark in the development of a local framework in Japan (Masashi, 2012) which

eventually led to the birth of the CEFR-J. In China, the use of the CEFR for standard-based assessment is somewhat still in the discovery phase. Zheng, Zhang, and Yan (2016) studied the possibility of application of the CEFR on the College English Test (CET) writing assessment in China and revealed that the use of the CEFR was permissible through proper execution which could be achieved when adequate training on familiarization of the CEFR is provided to teachers.

Previous studies have shown that the enthusiasm on implementing the CEFR to the English education system has spread worldwide. The Malaysian Ministry of Education too has undertaken similar initiative of adopting the CEFR into the country's English language education system to improve the English language proficiency of its students (Mohd Don, 2015). This effort is also a stepping stone to align its system with the international standards as well as to ensure that it is globally competitive (The Roadmap, 2015- 2025). Therefore, to gain a better insight into the CEFR implementation that takes place in the current English language education system, getting acquainted with the status of the CEFR in Malaysia is deemed necessary.

2.3 CEFR in Malaysia

The decision to embark on a project to implement the CEFR in Malaysia has been preceded by careful and thorough studies on other countries which have used the CEFR to learn from their experiences (Roadmap 2015-2025). From the studies done upfront, there are two rationales for adopting the CEFR. Firstly, it is to keep Malaysia in touch with the international standard. As English is the global language, transforming our English language education will grant us support to be on the international network. Hence, it is clear that moving towards international standard in language education is definitely an appropriate step to take now. Secondly, it is more economical to use a readily available framework than to develop a new local framework. Moreover, a locally produced framework has the risk of being irrelevant outside the home country. Therefore, the decision to adopt the CEFR is a wise choice to avoid the risk of being obsolete in the world that has steered towards the international standard.

The CEFR is implemented in Malaysia in three phases. The first phase happened in 2013 to 2015. This two-year span is accentuated on elevating the English language proficiency of school teachers. To realize this, teachers were sent out for training including Professional Up-Skilling of English Language Teachers (Pro-ELT), the Native Speaker programme, the Fulbright English Teaching Assistant programme and the Expanded Specialist Coach (SISC) role for English (Sani, 2016). Preparation for the development of the CEFR descriptors as well as target setting for each educational level also happened in this phase.

Following this is the second phase of the CEFR implementation (2016-2020) which is further broken down into two parts. The initial stage of this phase was dedicated to setting the appropriate CEFR levels against each educational level ranging from pre-school to teacher education. Additionally, alignment of School Based Assessment (SBA) syllabus and curricula with the CEFR descriptors as well as determining the CEFR-aligned textbook and materials that will be utilized also happened in this stage of the implementation. The second stage of this phase continued with the validation process of the CEFR levels set for each level of education as indicated earlier. Parallel to this is the implementation of the new CEFR aligned curricula which started in 2017 and continued to 2020. Synchronously, teachers were also sent out to attend trainings related to the CEFR to equip them with the necessities to implement it in schools.

Finally, the roadmap ends with the third phase where evaluation, review and revision will be done by the council on the implementation that has taken place in schools. The results obtained from the processes mentioned will provide a basis for the development of the CEFR-M, which is the focal point for this phase of the roadmap. In this early stage of the implementation, the success or failure of the integration of the CEFR into the country's English language education system is hardly predictable. Nevertheless, experiences of other countries implementing the CEFR into their education system have shown that issues and challenges are almost inevitable and mixed feedback from teachers as implementer is to be expected (Goullier, 2012; Komorowska, 2012; Zou, 2012). Hence, this study attempts to discover the current situation regarding the implementation of the CEFR in schools with focus on the challenges teachers faced as well as their belief of this change. Because teachers' readiness in accepting a new change has always been the number one concern in any reform implementation (Chin, Thien, & Chew, 2019; Kondakci et al., 2016) this study also aims to examine teachers' state of readiness in accepting the new change.

2.3 Research Questions

1. What are the challenges faced by English language teachers in implementing the CEFR in schools?
2. What are the belief of teachers regarding the CEFR implementation in the Malaysian English language education system?
3. What is the state of readiness of English language teachers in accepting the policy change?

3.0 METHODOLOGY

3.1 Data Collection

The study utilized a mixed-method design which entails the use of questionnaire and semi-structured interview. The questionnaire was adapted from Bouckenooghe and Devos (2009) with a few alterations made to the original OCQ-C, P, R to suit the educational setting in which this study was undertaken. The questionnaire consists of three sections with each section dedicated for different dimension of readiness for change (RFC). Section A is on Emotional RFC, Section B accentuated on Cognitive RFC while Section C is devoted for the Intentional RFC. For the interview, a self-constructed interview protocol was used as it allowed the researcher to use probes to elicit more information from the interviewees' responses in which structured and non-structured interview do not entail (Creswell, 2012). The interview protocol comprised a total of six main questions which would require participants to share their views pertaining to the challenges they faced throughout the implementation process of the CEFR in schools as well as their belief regarding the implementation of the CEFR.

3.2 Participants

To elicit data for the questionnaire, 365 English language teachers who were determined through the use of sampling table developed by Krejcie and Morgan (1970) were selected. The questionnaires were distributed in person during a two-day seminar attended by over 300 English teachers all over the state of Johor. From here, 164 questionnaires were completed and returned. The remaining responses were collected online through the use of various social media tools (*whatsapp, facebook messenger* and emails). The data collection for the questionnaires took about three months to complete.

For the interview, 15 English language teachers were chosen through purposive sampling strategy, as suggested by Creswell (2012), with consideration of three main attributes. Firstly, the participants must be English language teachers. Secondly, they must be those who have undergone a CEFR training. Thirdly, taking into account time and travelling factors, participants were only selected among teachers from schools in Johor Bahru district. The 15 participants for this study were 2 males and 13 females ranging from 26 years old to 45 years old who were teaching in either primary or secondary schools. Prior to the interview sessions, the researcher approached individual teachers through emails and personal telephone calls to request for their consent to be participants of the interview. After permission has been granted from each participant, a face-to-face interview session was scheduled outside school time and setting, as requested by the participants. All interview sessions were audio taped with the

consent of the participants with each session lasting between 35 and 50 minutes.

3.3 Data Analysis

The questionnaire data were analyzed using the Statistical Package for the Social Science (SPSS) version 25 where descriptive statistics was performed. Because the questionnaire employed Likert scale, frequency for each response was recorded and data were presented in percentage form. On the other hand, the interview recordings were transcribed using the Microsoft Word and were analyzed using thematic analysis. Although there are many ways to approach a thematic analysis, the present study used the 6-step framework proposed by Braun and Clarke (2006) as it offers a clear and usable approach to thematic analysis as shown in Figure 1 below.

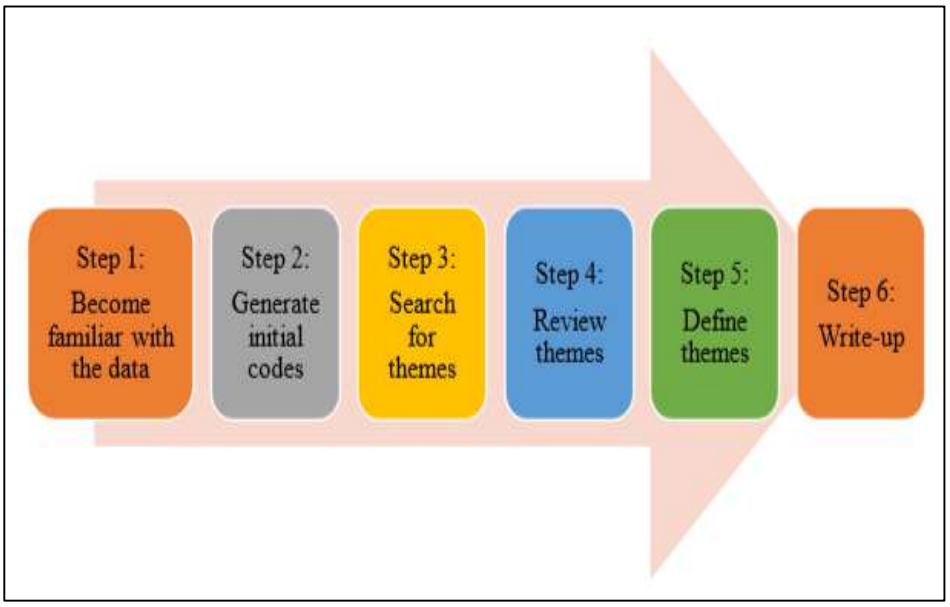


Figure 1: Braun & Clarke's Six-Step framework for doing Thematic Analysis

In step 1 of analysing the interview data, the transcribed data were read and re-read several times so as to be familiar with the data gathered. Significant parts of the interview were highlighted. This was followed by assigning initial codes to the identified parts in step 2. Some of the codes used were *'motivation', 'materials' and 'feel stress out'*. While coding the data, explicit mentions of factors related to the challenges were counted to identify factors that were frequently mentioned by the participants. In step 3, the codes that were identified in step 2 were reviewed and possible themes were identified. The themes that were identified in step 3 were reviewed in step 4 to ensure that all significant parts of the interview have been included. In step 5, the themes were defined. Concurrently, the interview transcriptions were broken down

into sections and grouped under their respective themes to ensure that related excerpts to support claims when presenting the findings are easily identifiable. The final step of the analysis was writing up the findings which is presented in the following section.

4.0 ANALYSIS AND DISCUSSION

4.1 Challenges in CEFR Implementation

Salient findings obtained from the semi-structured interview and questionnaire are sectioned into three parts, addressing the three research questions of the study. The analysis of the interviews held with 15 teachers revealed five challenges in implementing the CEFR which are teachers' motivation, materials, time, students' proficiency level and facilities.

4.1.1 Teachers' Motivation

The most recurring response from the interviews in relation to the challenges is motivation where fourteen out of fifteen participants testified that the greatest challenge in the implementation of the new reform was within their own self. This is aptly stated by one of the participants in the interview *"for me is more on my motivation, it's within me like I don't feel excited with this CEFR"* (R7). The analysis shows that participants' lack of motivation in implementing the new change seems to be invoked by several factors. Firstly, the feeling of anxiety and stress in implementing the CEFR as mentioned by five of the participants. They felt anxious and stressed out because they thought the change would entail a series of change including the way they teach in the classroom and the manner in which assessment is done. *"It's a stressful thing when you have to face change because many things will change as well.…when the system changed, the way you teach will have to change, how to do this and that, assessment and all will change as well. So many uncertainties"* (R1). This is supported by R4 *"even when they announced the change I was already stressed out...new change make me feel anxious because many things will change. So I am already demotivated from the start".* R7, R13 and R14 had similar opinion with R1 and R4. Having these thoughts had made these participants feel demotivated to implement the change from the start. The undesirable feelings as mentioned by the participants above are warranted as there is a mounting evidence in the education system in the world nowadays that due to a change in the nature of teaching, there is a great deal of uncertainties and identity crisis confronting teachers (Day, Elliot, & Kington, 2005).

The second reason mentioned causing the participants to feel demotivated was the limited knowledge they had about the CEFR. Not knowing what to do and how to do it made

the participants felt the pressure to try to make sense of the CEFR. As R3 stated that *"I feel like a newbie in my own field, I don't know very much about it"* and R13 who said that *"we teachers also can be demotivated especially when we don't know what we are doing"*. Not knowing whether what they were doing in the classroom was right or wrong was the thing that had resorted to their lack of motivation. R7 also added by saying *"I'm still grasping the whole CEFR thingy, I'm having one problem after another in trying to understand the whole thing, it kills your spirit to move forward"*. This situation is not only faced by the participants in this study but indeed has confronted majority of teachers alike to whom a new reform is being introduced. As Mohd Dzaquan (2020) revealed that teachers were constantly worrying about their inability to teach CEFR aligned syllabus to students. The situation was no difference in China where Zheng et al. (2016) exposed that the lack of knowledge has resorted in uncertainty and confusion among teachers and this has brought to the declining level of teachers' motivation to implement the change in the classroom.

Another reason that contributed to the participants feeling demotivated is the many changes that happened in the English language education system over the years. This is apparent in the sentiments articulated by R9, R11 and R15. R9 stated that *"we get tired also...I still feel numerous changes is not healthy not only for us teachers, I'm feeling slightly demotivated with all these changes"*. R11 voiced out her disappointment by saying "s*o we teachers we get fed up also, tired lah you change one then another one...I just don't understand why they keep changing the system"*. While R15 remarked that the whole process of change was tiring and too much for teachers to bear *"when everything changes again, we also have to change. Tiring even to think about it. You get tired of the whole process of change, too much for us to bear"*. Teachers' expressions in the excerpts above such as '*we get tired', 'we get fed up'* and *'too much for us to bear'* may signal that the change was pursued by teachers with much undesirable emotions. Such emotional instability in fact has been pointed out by a study by Vallax (2011) about change in school system being one of the important contributors of teachers feeling stressed out in schools. A great deal of pressure comes with a change as Lines (2005) pointed out that change implementation may bring with it many undesirable feelings of anxious, daunt and confusion.

4.1.2 Materials

Another challenge that was frequently mentioned by the participants was the limited available materials to support the implementation of the CEFR in the classroom. This challenge was articulated by thirteen out of the fifteen participants interviewed. The challenges related to the

lack of materials can be classified into two categories which are i) the lack of materials to support teaching and learning in the classroom and ii) the lack of materials in assisting teachers to try understand the CEFR as a whole.

In relation to the materials to support teaching and learning, thirteen participants testified that the materials provided by the ministry was only limited to the textbook where all teachings had to be based on the contents in the textbook. The opinion of the thirteen participants is highlighted by R2 and R6 as follows *"...the lack of supporting materials for students...like workbook, etc. that are interesting to supplement the use of textbook"* and *"...not having enough materials to teach students with, textbook alone is mmm not enough, but to find other supplementary materials that are CEFR-aligned, this is important, not many CEFR-aligned materials out there"*.

Another aspect of the scarcity of materials is in terms of the lack of materials to assist teachers in comprehending the whole concept of the CEFR. Ten participants argued that they needed extra materials that would enable them to have a comprehensive understanding of what the CEFR is all about and that would allow them to better implement the CEFR in the classroom. For instance, R3, R4, and R8 claimed that there were not many materials available for teachers to support the implementation of the CEFR. R3 testified by saying *"for me personally I try to understand the CEFR but problem is not many materials are available out there for us teachers"* and R4 said *"we were not given enough materials for ourselves"*. R8 echoed the above point when she mentioned that "*perhaps materials for me, for teachers such as reading or videos to demonstrate how the CEFR should be implemented"*. As far as this issue is concerned, R10 further testified that although there were materials given to them during the training they attended, it was far from sufficient to understand the new change "*they gave us during the training, in-house training, but it is not really all that we need to understand, not enough"*. This issue has been emphasized by Ghazali (2016), Kenayathulla and Ibrahim (2016) and Abdul Aziz, Abd. Rashid, and Zainudin (2018) where teachers were confronted with the lack of supporting materials in terms of handbook and module as well as curriculum documents to support them in the change implementation.

This issue of inadequacy of materials is somewhat universal in almost all change implementations in the field of education. For instance, Badugela (2012) has shown that in almost all change effort in South Africa, educators often face challenges such as inadequate sources to support the change implementation. This case is also true in Indonesia where a Winardi and Priyanto (2016) revealed that the inadequacies of materials was one of the issues that confronted teachers. This is unfortunate according to Abdul Aziz et al. (2018) because in

a case study of curriculum innovation in Malaysia, for the curriculum to be effectively implemented, the materials have to be of high quality and meet the needs of both teachers and students, have to be available in adequate numbers and at the appropriate time. This opinion is concurred by Berlinski and Busso (2017) when they proposed that materials support is paramount in building teachers' readiness in implementing a new change.

4.1.3 Time

Time is another challenge that was highlighted in the interview. According to thirteen of the participants, time was an essence to any change implementation but they were not given ample time to learn about the CEFR before the implementation commenced. R2, R8 and R11 made clear statements that teachers, students alike, were not given enough time to get familiarized with the CEFR before they were asked to implement it in schools. This view is shown in the following quotes from the interview *"for me time is always the challenge…there was not enough time given to get to know the CEFR first before we implement it in school, they announce it, and we implement straight away...I think it is a hasty decision"* (R2), *"I think because we were not given much time to be familiar with the CEFR first before it is being rolled out by the authority...that is the challenge"* (R8), *"we are not given enough time to familiarize ourselves with the CEFR"* (R11).

Another remark made by the participants was that time was never enough for them as there were other responsibilities they had to bear such as keeping record of students' result, attending school meetings, planning and running school events, disciplining students and many others that they had to complete apart from their core business of teaching. Hence, learning about the CEFR would take up some of their time and this would be another task they had to shoulder. Some of the responses that represent this view are as articulated by R13 and R14 where they claimed that *"time alone is already a challenge if you want to see it from the challenges we face...because not only we have to teach, think about our lesson plan, how to teach better in class, disciplining is also our responsibility, where when our students do something wrong, we have to be responsible for it, and plus, we have to plan school meetings, attend the meetings also, many other things la…and now adding the CEFR in the list…pretty intense"* (R13), *"Time is an issue also, there are a lot of things to do in school, not just teach, we also have to go for meetings, planning for school events, teachers' day celebration, and we also have to check students' assignments"* (R14). In this vein, both participants believed that they needed the extra time to get acquainted with the CEFR before would they be able to implement it properly. Time being one of the challenges in the CEFR implementation has

indeed been pointed out by Nurul Farehah and Mohd Sallehhudin (2018) in their study where they highlighted teachers had to be given more time to get familiar with the framework.

4.1.4 Students' Proficiency Level

Another challenge that seems apparent from the interview responses was students' proficiency level. Seven out of 15 participants mentioned that their students' proficiency level was one of the limiting factors to a successful change implementation in the classroom. R2 and R14 testified that some of their students could not understand basic instructions even though they were given examples of such instructions like *'please submit your work on my desk'* and *'don't forget to refer to the example given'*. According to R2 and R14, the two instructions they gave were simple yet their students could not comprehend the instructions. R2 ended her remark by saying that *'how are these students going to cope with CEFR?'* and R14 also voiced her frustration that the CEFR is way beyond her students' level which to her is disheartening. Similarly, R11 mentioned about her students being very 'slow' in learning the English language as she stated in the interview *"I tried to do speaking activity with them, asked them to talk about independence day celebration at school, after forcing them to speak you know only two out of my thirty-five students responded"* (R11).

Additionally, R5, R12 and R14 called attention to the issue with students' low English language proficiency level was not uncommon to teachers as R5 said *'the issue of students' low proficiency is not new to us'* and has always been the challenge as far as teaching English language is concerned even before the implementation of the CEFR. However, participants viewed this problem as even more worrying at the present time where the CEFR is brought to the classroom because of the internationalized contents of the textbook they used in lesson delivery in the classroom. As aptly described by two of the participants "*now is more challenging because we need to teach based on the standard set for international level"* (R12), *"now is more difficult because the textbook we used has contents that are based on international standard, no longer based on local context"* (R14). In addition, R5 and R14 voiced out their concern regarding the learning opportunity students could get when the contents were too hard for them and learning seemed to be impossible especially when students could not understand most of the contents of the lesson. This issue has also been pointed out by Mohd Dzaquan (2020) where teachers were always concerned about the lesson based on CEFR standard being too difficult for students to grasp.

In discussing the issue of internationalized contents of the textbook, eight out of 15 participants mentioned that because the contents of the textbook were not based on the local

context making it difficult for students to understand and comprehend because it is not within their background knowledge and not something that they were familiar with. To illustrate this view, R1 opined that *"the textbook is not based on Malaysian context...is more like it's suitable for the other like overseas kids"*. R4 on the other hand, believed that *"if the content is based on local context, maybe it is less difficult for them because they might have experienced it before"*. R9 added by saying that *"I think if the content is something that students are familiar with like Hari Raya celebration, maybe learning can happen"*. Hence, it can be inferred that when students learn about something that is not within the local context and something which is not familiar to them, teaching and learning can be very challenging as Sumaryono and Ortiz (2004) stressed that English language learners could be disconnected from the learning process if the teachers do not display sensitivity towards their cultural identity. When this happens, the content that is internationalized not only made it difficult for students to understand but also difficult for teachers to make the lesson comprehensible for them. This is pointed out by R6 *"they learn about thanksgiving, they don't know this. We teachers tried hard to make them understand, so difficult you know"*. The concern regarding the internationalized contents of the textbook being hardly comprehensible to students has been voiced out, among others by Monihuldin (2018) and Star (2018). Hence, in dealing with the issue mentioned, participants believed that localization of the material could be one of the initiatives to counter this issue. This is supported by a recent study by Deswila et al. (2020) who believe that cultures should be injected in the learning materials. This is warranted as the significance of localized materials has been proven to be effective in influencing participants' comprehension in a reading lesson (Mahabadi, 2012) where it was revealed that students' scores were higher when using a localized content material due to the familiarity of contents.

4.1.5 Facilities

Another challenge that was brought up by the participants is related to the teaching and learning facilities needed to implement the CEFR in the classroom. R2 had this to say when she did the speaking and listening activities "*because we're lacking in the facilities itself... especially when I do the speaking and listening part cause for speaking and listening we have to carry our own speakers to the class and all those things"*. This view is concurred by R7 where he mentioned about having to prepare all the needed facilities like speakers and CDs *"we have shortage of support for facilities needed...we have to provide all, especially when you want to do activities say listening for example, we teachers have to bring our own speaker, CD and all...all these have to be operated within the limited teaching and learning period…it takes up most of the*

times already". A similar view was raised by R13 where the lack of facilities like speakers and radio would be a challenge for teaches when they wanted to do listening activities in the classroom.

On the other hand, R7 claimed that his school had a media room where all the facilities needed such as radio, huge white screen and speakers were provided, however, this room had to be shared with all other teachers in the school. Therefore, teachers had to queue to use the room as it was only available on first come first served basis. According to this teacher, waiting to use the room to be available was time consuming and they had to book the room at least a week before the time of use. Moreover, having the media room far from the students' classroom was also a challenge as commuting to the media room would take up some of the lesson's time as R7 puts it *"And plus when students are asked to use the media room, almost 20 minutes will be wasted just for them to get there. So time consuming"*.

4.2 Teachers' Belief

Despite the challenges confronting the teachers, they were still able to look at the change as a positive move by the ministry. All 15 participants interviewed believe that the CEFR implementation could benefit the country's English language education system in the future. In order to demonstrate her positive view of the change, R2 articulated that *"I know it will work well through time, and of course it will be able to improve the system"*. R4 also seemed to be in agreement when she mentioned that *"I do think the whole process can work, but for better outcome, it needs time, we need time"*. R7 demonstrated his belief by saying that *"of course we change for better education, so does the CEFR and when all of us join together to make it work"*. Other participants also concurred in this view; *"I believe all changes are for better education system so does the CEFR and I am confident it will make a difference in the future"* (R9), *"I think the CEFR would benefit everyone through a proper implementation...it is good effort"* (R11). From the remarks, it is apparent that participants viewed the change positively. However, all the remarks were accompanied by certain conditions for instance, R2 stated *"through time"*, R4 mentioned *"it needs time"*, R7 highlighted "*when all of us join together"*, R9 said "*in the future*" while R11 pointed out *"through a proper implementation"*. From the participants' point of view, the change would be beneficial for the system only when certain conditions are met namely through appropriate timing, collective effort and proper implementation process.

It seems apparent from the findings above that despite the challenges confronting teachers in implementing the CEFR in schools, they still viewed the reform as a positive effort

from the ministry to uplift the English language education system of the country which is parallel to the findings of Fatima (2019) and Faez, Taylor, Majhanovich, and Brown (2011) where all teachers exhibited positive belief on the incorporation of the CEFR in ESL classroom. Through the findings, it can be said that this positive outlook on the change might have eventually influenced teachers' willingness to make necessary contributions to the implementation of the change, which is most closely associated with their readiness of accepting the change. The next section will discuss this.

4.3 Teachers' Readiness for Change (RFC)

Data from the questionnaire provide answers to the third research question of this study which revealed teachers' state of readiness in implementing the CEFR in schools. This would offer a much more comprehensive understanding on the teachers' actual acceptance in implementing the new change. Data were analyzed using SPSS and the frequency for each response is recorded and presented in the form of percentage (see Table 1 below).

Table 1: Teachers' readiness for change (RFC)

	Emotional RFC	**SA**	**A**	**N**	**D**	**SD**
1	I have a good feeling about the change.	7.0	51.9	24.9	16.2	-
2	I experience the change as a positive process.	7.6	70.8	18.1	3.5	-
3	I find the change refreshing.	11.1	63.2	18.6	7.0	-
4	I am completely ready for the change.	3.5	24.6	28.6	43.2	-
5	I am ready to accommodate and incorporate changes into my teaching.	7.0	43.8	27.3	21.9	-
	Cognitive RFC					
1	Most changes that are supposed to solve problems in English education are working effectively well.	-	20.8	48.1	27.8	3.2
2	I think the new change will be successfully implemented by teachers.	7.6	55.9	33.0	3.5	-
3	I believe that the change will improve my teaching.	10.5	33.5	38.6	17.3	-
4	I believe that the change will simplify work.	3.5	17.8	27.3	34.6	16.8
	Intentional RFC					
1	I want to devote myself to the process of implementing the change.	7.0	79.2	10.5	3.2	-
2	I am willing to make a significant contribution to the change.	11.1	78.4	10.5	-	-
3	I am willing to devote my energy into the process of change.	7.0	47.3	35.1	7.0	3.5

Table 1 shows the findings for all dimensions of readiness for change namely emotional RFC, cognitive RFC and intentional RFC. Under the dimension of emotional RFC, statement 2 - I

experience the change as a positive process, recorded the highest percentage with a total of 78.4% of the participants responded 'agree' and 'strongly agree' while none of them strongly disagreed with the statement. For cognitive RFC, statement 2 - I think the new change will be successfully implemented by teachers, recorded the highest mode percentage with about two-thirds (63.5%) of the participants agreed and strongly agreed with the statement while only 3.5% of them disagreed. Under the dimension of intentional RFC, statement 2 - I am willing to make a significant contribution to the change and statement 1 - I want to devote myself to the process of implementing the change, recorded high percentages of participants agreeing and strongly agreeing, 89.5% and 86.2% respectively. Comparatively, these two statements 1 and 2 under the intentional RFC recorded significantly high percentages as compared to the other statements in all three dimensions.

The findings from the analysis above seem to be in tandem with the findings from the interview. For instance, the findings under the dimension of emotional RFC indicated that majority of the participants believed that the change is a positive effort from the ministry. This positive view was also noted in the interview as explained under teachers' belief.

Additionally, from the data on cognitive RFC, majority of the participants agreed that they can implement the change successfully. The interview data seem to show parallelism to this stance. When asked about their opinion whether participants think they are capable of implementing the change in the classroom, 13 out of the total 15 participants claimed that they believed they are capable of implementing the change in schools. For instance, R3 stated that *"I believe we can implement it successfully in school with sufficient support of course, from the authorities...like enough materials, and trainings maybe do workshops for us"*, R7 also pointed out his view by saying *"I think we can do it...we can implement changes, but just give us time to be familiar with the system first, maybe let us learn what CEFR is, then we can implement it better"*, R8 demonstrated her thoughts through her sentiment *"if we have enough support we need from the government...sufficient materials and training, together we teachers also can support each other throughout the process, I am confident we can make it work" while* R11 also mentioned that *"if we are given more time to study and learn what CEFR is, I believe we can implement it successfully"*.

However, quite interestingly, despite being optimistic about their ability to carry out the change successfully, participants based this view on a certain condition. As can be seen in the excerpts above, the use of phrases like '*with sufficient support*', '*but just give us time', 'if we have enough' and 'if we are given more time'* were used by these teachers, indicating that their ability to implement the CEFR is dependent upon a certain requirement namely sufficient

time to get acquainted with the change, support from the authorities in terms of adequate materials and training, and collective effort from colleagues. However, despite being pointed out as determinants of influencing factors in shaping teachers' ability to successfully implement the change, insufficient time and inadequate materials unfortunately have been pointed out as issues in implementing policy change in school (Chin et al., 2019; Abdul Aziz et al., 2018; Othman, Md Saleh, & Mohd Nooraini, 2013). Abdul Aziz et al. (2018) pointed out that curriculum documents such as Curriculum Standard were not yet fully ready when teachers attended courses and this had caused problems throughout the course while Kenayathulla and Ibrahim (2016) mentioned the insufficient implementation modules during the time of change.

In addition, from the point of view of the intentional RFC, the finding dictates that majority of the teachers agreed that they were willing to devote themselves and willing to make a significant contribution to the change. These notions have also been pointed out by all participants in the interview when prompted about their willingness to pursue the change in classroom as follows *"I will, I am willing to change in any way especially my teaching as parallel to the CEFR standard"* (R1), *"since the change is happening now, I am and I know all are willing to contribute to the change"* (R3), *"of course I am willing to contribute to the change"* (R11). Similar views were also noted in all other responses in the interview. Hence, it can be opined that all of the participants were willing to contribute to the change implementation.

From the discussion, it can be postulated that despite challenges confronting teachers in the implementation of the new change, positive belief about the CEFR implementation is evident and this shows that teachers are emotionally ready to accept the change. However, important facets namely time, collective effort and sufficient materials are the determinants that would shape teachers' cognitive readiness for change. This study points to the direction that only when these supports are provided will teachers be able to successfully implement the CEFR in schools.

5.0 CONCLUSION

Given that the implementation of the CEFR in the English language education system is still at the stage of infancy, it is presumed that challenges revealed by this study namely teachers' motivation, materials, time, students' proficiency level and facilities are inevitable. Improvements would still be needed in the implementation of the CEFR. Nonetheless, constant support in terms of adequate time, sufficient materials and training from the authorities are deemed necessary in order to intensify teachers' readiness to implement the change because

when teachers are ready to accept the change, alteration of actions will happen accordingly (Armenakis, Harris, & Mossholder, 1993).

Change is not always easy however, if change is necessary all will have to come together and do their part to support it. The implementation of the CEFR is seen as a promising reform in the English language education system and a step in the right direction to ensure the standard of English language in our country is enhanced.

REFERENCES

Abdul Aziz, A. H., Abd. Rashid, R., & Zainudin, W. Z. W. (2018). The enactment of the Malaysian Common European Framework of Reference (CEFR): National master trainer's reflection. *Indonesia Journal of Applied Linguistics, 8*(2), 409-417.

Armenakis, A., Harris, S. G., & Mossholder, K. W. (1993). Creating readiness for organizational change. *Human Relation, 46*(6), 681-703.

Armenakis, A., & Harris, S. (2009). Reflections: Our journey in organizational change research and practice. *Journal of Change Management, 9*(2), 127-142.

Azman, H. (2016). Implementation and challenges of English language education reform in Malaysian primary schools. *3L: The Southeast Asian Journal of English Language Studies, 22*(3), 65–78.

Badugela, T. M. (2012). *Problems facing educators in implementing the national curriculum statement: The case of Tshifhena secondary school, Vhembe District, Limpopo Province, South Africa.* (Unpublished master thesis). University of South Africa.

Bérešová, J. (2011). The impact of the common European framework of reference on teaching and testing in central and Eastern European context. *Synergies Europe, 6*(2011), 177-190.

Berlinski, S., & Busso, M. (2017). Challenges in educational reform: An experiment on active learning in mathematics. *Economics Letters, 156*(1), 172-175.

Bouckenooghe, D., & Devos, G. (2009). Organizational change questionnaire – Climate of change, processes, and readiness: Development of a new instrument. *The Journal of Psychology, 143*(6), 559–599.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(1), 77-101.

Broek, S., & Ende, I. V. D. (2013). *The implementation of the common European framework for languages in European education systems.* Directorate-General for Internal Policies, European Parliament.

Buckland, S. (2010). *Using the CEFR to benchmark learning outcomes: A case study.* Wall Street Institute, Spain.

Chin, H., Thien, L. M., & Chew, C. M. (2019). The reforms of national assessments in Malaysian education system. *Journal of Nusantara Studies, 4*(1), 93-111.

Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment.* Cambridge University Press.

Creswell, J. W. (2012). *Educational research. Planning, conducting and evaluating quantitative and qualitative research* (4th ed.). Pearson.

Day, C., Elliot, B., & Kington, A. (2005). Reform, standards and teacher identity: Challenges of sustaining commitment. *Teaching and Teacher Education, 21*(1), 563–577.

Deswila, N., Kustati, M., Yusuf, Y. Q., Raja Nor Safinas, R. H., Besral, & Rehani. (2020). Cultural contents in the ELT textbook prepared by the Indonesian ministry of education for secondary schools. *Journal of Nusantara Studies, 6*(1), 222-241.

Faez, F., Taylor, S. K., Majhanovich, S., & Brown, P. (2011). Teacher reactions to CEFR's task-based approach for FSL classrooms. *Synergies Europe, 6*(1), 109-120.

Fatima, S. (2019). Perceived view of teachers towards Pentaksiran Tingkatan Tiga (PT3) (Form three assessment) English language: A case study. *Asian Journal of University Education, 15*(3), 34-44.

Figueras, N. (2012). The impact of the CEFR. *ELT Journal, 66*(4), 477–485.

Fulcher, G. (2004). *Testing second language speaking.* Routledge.

Fulcher, G. (2010). The reification of the Common European Framework of Reference (CEFR) and effect-driven testing. In A. Psyaltou-Joycey & M. Mattaioudakis (Eds), *Advances in research in language acquisition and teaching* (pp. 15-26). GALA.

Ghazali, N. H. C. M. (2016). An evaluation of teachers' perception in implementing school based assessment. *Jurnal Pendidikan Sains dan Matematik Malaysia, 6*(1), 24-40.

Goullier, F. (2012). Policy perspectives from France. In M. Byram & L. Parmenter (Eds.), *The common European framework of reference: The globalisation of language education policy* (pp. 37-44). Multilingual Matters.

Harsch, C., & Martin, G. (2012). Adapting CEF-Descriptors for rating purposes: Validation by a combined rater training and scale revision approach. *Assessing Writing, 17*(1), 228-250.

Kenayathulla, H. B., & Ibrahim, M. S. (2016). Analisis terhadap pelaksanaan dasar pentaksiran berasaskan sekolah di Malaysia: Persepsi guru [Analysis about the implementation of school-based assessment in Malaysia: Teachers' perception]. In H. Ahmad & M. S.

Ibrahim (Eds.), *Transformasi pendidikan nasional: Antara aspirasi dan anjakan* (pp. 395-423). Penerbit Universiti Malaya.

Komorowska, H. (2012). Academic perspectives from Poland. In M. Byram & L. Parmenter (Eds.), *The common European framework of reference: The globalisation of language education policy* (pp. 104-113). Multilingual Matters.

Kondakci, Y., Beycioglu, K., Sincar, M., & Ugurlu, C. T. (2016). Readiness of teachers for change in schools. *International Journal of Leadership in Education, 20*(2), 176–197.

Krejcie, R. V., & Morgan, D. W. (1970). Determining sample size for research activities. *Educational and Psychological Measurement, 30*(3), 607-610.

Lines, R. (2005). The structure and function of attitudes toward organizational change. *Human Resource Development Review, 4*(1), 8–32.

Lowie, W. M., Hainesa, K. B. J., & Jansmaa, P. N. (2010). Embedding the CEFR in the academic domain: Assessment of language tasks. *Procedia - Social and Behavioral Sciences, 3*(1), 152–161.

Mahabadi, S. (2012). The role of localized materials in learning of FFL students. *Journal of Languages and Culture, 4*(5), 65-74.

Maldina, E. (2015). *The impact of the common European framework of reference on foreign language instruction: The case of sociolinguistic and pragmatic competence*. (Published master thesis). University of Toronto.

Masashi, N. (2012). The development of the CEFR-J: Where we are, where we are going. http://www.tufs.ac.jp/common/fs/ilr/EU_kaken/_userdata/negishi2.pdf

Ministry of Education Malaysia. (2015). *English language education reform in Malaysia: The roadmap 2015-2025.* English Language Standards and Quality Council, Ministry of Education Malaysia.

Mohd Don, Z. (2015). English language proficiency, graduate employability and the role of CEFR. http://repository.um.edu.my/98438/1/Asean%20Seminar%202015%20English%20language.pdf

Mohd Dzaquan, I. M. A. (2020). *Practice of teachers' teaching common European Framework of reference aligned writing syllabus for young learners.* (Unpublished master thesis). Universiti Teknologi Malaysia.

Monihuldin, M. (2018). Discussion on use of foreign English textbooks in schools sparks debate. http://news.iium.edu.my/2018/01/14/discussion-on-use-of-foreign-english-textbooks-in-schools-sparks-debate/

Nakatani, Y. (2012). Exploring the implementation of the CEFR in Asian contexts: Focus on communication strategies. *Procedia - Social and Behavioral Sciences*, *46*(1), 771-775.

Che Musa, N., Lie, K. Y., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA Online Journal of Language Studies, 12*(1), 35–51.

Nurul Farehah, M. U., & Mohd Sallehhudin, A. A. (2018). Implementation of CEFR in Malaysia: Teachers' awareness and the challenges. *3L: The Southeast Asian Journal of English Language Studies, 24*(3), 168-183.

Othman, I., Md Saleh, N., & Mohd Nooraini, N. A. (2013). The implementation of school based assessment in primary school standard curriculum. *International Journal of Education and Research, 1*(7), 1-10.

Sani, R. (2016, July 18). Journey to master English. *The New Straits Times Online.* https://www.nst.com.my/news/2017/03/159164/journey-master-english

Star, T. (2018, February 4). Let's have our own textbooks. *The Star Education.* https://www.thestar.com.my/news/education/2018/02/04/lets-have-our-own-textbooks/

Sumaryono, K., & Ortiz, F. W. (2004). Preserving the cultural identity of English language learner. *Voices from the Middle, 11*(1), 16-19.

Vallax, P. (2011). *The common European framework of reference for languages: A critical analysis of its impact on a sample of teachers and curricula within and beyond Europe.* (Unpublished PhD thesis). The University of Waikato, New Zealand.

Van Huy, N., & Hamid, O. M. (2015). Educational policy borrowing in a globalized world: A case study of common European framework of reference for languages in a Vietnamese University. *English Teaching: Practice & Critique, 14*(1), 60–74.

Winardi, & Prianto, A. (2016). Various determinants of individual readiness to change and their effects on the teachers' performance (A study on certified teachers in Jombang Regency East Java, Indonesia). *IOSR Journal of Business and Management, 18*(2), 22-32.

Zheng, Y., Zhang, Y., & Yan, Y. (2016). *Investigating the practice of the common European framework of reference for languages (CEFR) outside Europe: A case study on the assessment of writing in English in China.* British Council.

Zou, W. (2012). Perspectives from China. In M. Byram & L. Parmenter (Eds.), *The common European framework of reference: The globalisation of language education policy* (pp. 183-196). Multilingual Matters.

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**Supporting English as a second language (ESL) teachers' professional development through multi-platforms' online affinity space / Aziz, A. H. A. A., Zainudin, W. Z. B. W., & Rashid, R. A.**
Source	***International Journal of Interactive Mobile Technologies*** **Volume 14 Issue 9 (2020) Pages 258–264** **https://doi.org/10.3991/ijim.v14i09.11158** **(Database: ResearchGate)**

Supporting English As a Second Language (ESL) Teachers' Professional Development Through Multi-Platforms' Online Affinity Space

https://doi.org/10.3991/ijim.v14i09.11158

Abdul Hakim Ali Abdul Aziz (✉)
Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia
aliex87@gmail.com

Wan Zhafirah Binti Wan Zainudin
Sekolah Kebangsaan Sultan Ibrahim Satu, , Kelantan, Malaysia

Radzuwan Ab. Rashid
Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia

Abstract—This paper narrates the use of multi-platforms' online affinity space called Teacherfiera.com to support ESL teachers' professional development. Teacherfiera.com utilizes three online platforms which are BlogSpot, Facebook Group and Telegram Group that works in parallel with each other. Each platform is readily available to be used for free by the public and provides contextualized, personal as well as group interactions. These platforms which have been integrated into Teacherfiera.com also work as a resource bank for English Language Teaching materials which are accessible 24/7 and allows users to respond to specific materials or engage in a general topic of discussion. The creation of Teacherfiera.com started as a response towards the need for teachers' support in the Malaysia's nationwide dissemination and implementation of the new Common European Framework of References for Languages (CEFR)-aligned English Language Curriculum.

Keywords—Teachers' professional development; peer mentoring; curriculum innovation; affinity space

1 Introduction

This study aims to narrate the use of multi-platform online applications by Teacherfiera.com as the medium for affinity space among ESL teachers, as part of their professional development. Teacherfiera.com consists of a BlogSpot website, a Facebook group and a Telegram group that combined to be a unified platform that can be considered as an affinity space, a concept built by [1]. This concept explains how people interact within a space [1]. The administrator chose these social media platforms as it is the current trend and holds huge ability to accommodate the objective of Teacherfiera.com [2]. Currently, there are more than 6.8 million pageviews on their

BlogSpot, 31000 members on their Facebook group and 33000 subscribers on their Telegram group.

This platform is developed and maintained by two local ESL teachers and Master Trainers, Mdm. Wan Zhafirah and Mr. Hakim together with a Senior Lecturer from University Sultan Zainal Abidin, Associate Professor Dr. Radzuwan, in response to the need of Malaysian Primary School English Teachers to get support and guidance in implementing the new CEFR-aligned curriculum starting from January 2018. The training provided by the ministry for the new curriculum was thought to be inadequate [3] as it was evidently had the common problems of cascade training programme such as transmissive training instead of reflective, inadequate time, and most importantly, no follow up or after-training support provided [4].

The administrator of teacherfiera.com provides in-house made materials, such as printable picture cards, books, modules, audio clips as well as PowerPoint slides that are fully compatible with the new curriculum, while at the same time hosting a platform for discussion on the materials as well as other topics related to ESL teaching in Malaysia. More importantly, the administrator together with other members also provides online training for ESL teachers using videos and online chat, on how to conduct lessons and the use of different types of materials.

Despite being led by Master Trainers, Teacherfiera.com is considered as informal learning for the ESL teachers. This is because it existed in the online social network environment [5] and there is no fixed curriculum or programme to be followed [6], but rather the sharing of knowledge and experience are made personally by users while being moderated by the administrators. The different platforms within teacherfiera.com are interrelated and provide different functions that will be explained in the following sections.

2 Platforms Within Teacherfiera.com

There are three main components or platforms within teacherfiera.com which are the Blogspot, Facebook Group as well as Telegram Group.

2.1 Teacherfiera.com blogspot

Teacherfiera.com Blogspot serves as the main 'face' for Teacherfiera.com. It is used as the principal database where the users use to access all Teacherfiera.com's in house-made resources that includes printable materials, powerpoint slides, tutorial videos as well as audio clips.

It also provides tutorial videos in making teaching aids, teaching techniques and other related videos in the form of Youtube extension. The use of Youtube extension is regarded as very effective as it fast and posed very little limitations [7]. As of 4th October 2019, there are more than 800 posts made to this site.

The layout of the web page is arranged within three main columns and labelled for each one of the posts, pages or links. When a visitor reaches the main page, they will have the option to either open the posts and download the materials provided within the

post or they may also click on the links that will bring them to Teacherfiera.com Facebook Group or Telegram Group. The transition between the three platforms is seamless, as the web page is designed to integrate with the other platforms and are accessible by using either computer, smartphones or tablets.

Based on the Google Analytics report since 1st December 2017 until 4th October 2019 (Refer Appendix A), Teacherfiera.com BlogSpot received more than 6.8 million page views, and served more than 600 000 users mainly from Malaysia. As most of the materials are made to cater to the Malaysian ESL Curriculum, it is natural for most visitors to come from Malaysia. Still, since the materials are also compatible with the CEFR, which is adopted by many countries [8], [11], the website also receives visitors from other countries as mentioned above.

2.2 Teacherfiera.com Facebook group

The materials that were published in the Teacherfiera.com website will be showcased in the Teacherfiera.com's Facebook Group whereby each individual material set is being uploaded as one post. As Facebook is a very popular platform among teachers in Malaysia, it is a great platform that provides an easy way to reach more users using it's Share button as well as to receive comments and questions on the specific materials on each post made. Teachers will comment with their experience of using the materials or ask questions on how to adapt the materials to their students' or classroom's specific needs. They will also share the posts to their colleague who might benefit from the posts. The discussions are always joined by administrators together with other members. As of April 2019, there are more than 500 posts and more than 23000 active members (Refer Appendix B) in this Facebook Group. [5] notes that the learning through Facebook is considered as an informal one but are part teachers' professional development.

2.3 Teacherfiera.com Telegram group

While the Facebook group is the place to discuss each material set specifically, as it appears in individual posts, there is a need for a platform to discuss and share general knowledge in teaching such as issues of pedagogy or methodology within the context of English Language Teaching as well as education as a whole. To do this, an official Teacherfiera.com Telegram Group was made to address related issues in the context mentioned before. It is not only being appreciated by English teachers, but also by parents and anyone who are interested in English Language teaching to discuss a broad area of concern, such as problematic classroom management, getting children to be interested with English Language or simple things like how to use certain grammar rules correctly.

As of May 2019, there are more than 33000 members in this group, which consist of mainly primary school English Language teachers, as well as parents, education officers and lecturers with different educational background and experience. The diversity in the members' background and experience allows for deep discussion and

complex sharing of ideas. Every question will be answered base upon each members' different opinions, leading to different ways of doing things.

3 Teacherfiera.com Users Flow

Based on the previous explanation of different platforms that are interrelated with each other, the flow of members or users of Teacherfiera.com can be visualized as Diagram 1:

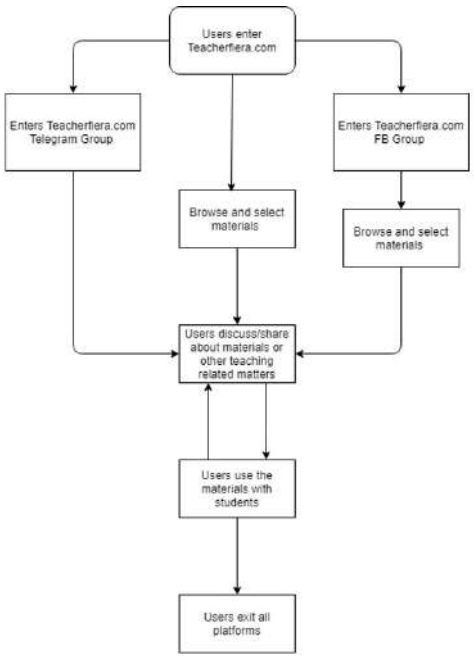


Fig. 1. Teacherfiera.com Users' Flow Chart

4 Context of Innovation

Teacherfiera.com aims mainly to assist the Malaysian English Language teachers by providing wide range of materials that cater broad range of students' and teachers' need. But, as the implementation of the new CEFR-aligned curriculum is still on-going, and the administrators are the Master Trainers for the new curriculum, there is an inclination to focus more on supporting teachers in adapting to the new curriculum. As of the year 2019, the new curriculum is being used by Standard 1, Standard 2 and Standard 3, and will continue to be used by Standard 4 in 2020 and so on. Not just that, anyone who are interested in English Language Teaching, such as parents who are interested in teaching their children at home also benefited from Teacherfiera.com. This innovative practise is not restricted to education sectors only but can also be applied to provide support for

any types of training in any field, either before training, during training or after training support.

5 Importance for the Malaysian ESL Community

Since the inception of the new CEFR-aligned curriculum in January 2018, many teachers, especially the non ESL majors, were struggling in its implementation with many still confused of how they should do things and why [3]. Teacherfiera.com solves the confusion and answer the questions by furnishing accurate information. At the same time, the platform also encourages dialogic reflection among members which is a very important part in Teachers Professional Development (TPD) [12]. It is evident that Teacherfiera.com plays an important role in helping the implementation of CEFR in Malaysia as it is not only used by more than 40000 Malaysian English Language Teachers, but also Education Officers and trainers within the Ministry of Education Malaysia. Materials from Teacherfiera.com are used by these officers and trainers during their training sessions nationwide.

6 The Future of Teacherfiera.com

In its current form, Teacherfiera.com has won several awards at international and national innovation competition. This includes a gold medal at International University Carnival on E-Learning (IUCEL) 2018, Educational Project of Innovation Competition (EPIC) 2018 and UniSZA's *Minggu Penyelidikan dan Inovasi* (Innovation and Research Competition) 2019. Even though these competitions were held in Malaysia, it is considered as an international competition as it were participated by innovators from other countries including Thailand, Singapore and Indonesia.

Based on the feedbacks given by the judges, Teacherfiera.com will be further enhanced. It is in the process of to be transformed into a single integrated and connected application that will be available to all types of mobile devices as well as computers and laptops.

7 References

[1] J. P. Gee, "Semiotic social spaces and affinity spaces: from the age of mythology to today's schools.," in Beyond communities of practice: language, power and social context, B. Dave and K. Tusting, Eds. Cambridge: Cambridge University Press, 2005, pp. 214–232.https://doi.org/10.1017/cbo9780511610554.012

[2] Y. Aleksandrova and S. Parusheva, "Social media usage patterns in higher education institutions - An empirical study," Int. J. Emerg. Technol. Learn., vol. 14, no. 5, pp. 108–121, 2019. https://doi.org/10.3991/ijet.v14i05.9720

[3] A. H. A. B. A. Aziz, A. R. Radzuwan, and W. Z. B. W. Zainudin, "The enactment of the Malaysian common European framework of reference (CEFR): National master trainer's reflection," Indones. J. Appl. Linguist., vol. 8, no. 2, pp. 409–417, 2018. https://doi.org/10.17509/ijal.v8i2.13307

[4] H. K. Bett, "The cascade model of teachers' continuing professional development in Kenya: A time for change?" Cogent Educ., vol. 3, no. 1, pp. 1–9, 2016. https://doi.org/10.1080/2331186x.2016.1139439
[5] R. Rashid, M. F. Yahaya, M. F. Rahman, and K. Yunus, "Teachers' informal learning via social networking technology," Int. J. Emerg. Technol. Learn., vol. 11, no. 10, p. 76, 2016. https://doi.org/10.3991/ijet.v11i10.5908
[6] D. Livingstone, Adults' informal learning: definitions, findings, gaps and future research. Toronto: Ontario Institute for Studies in Education, 2001.
[7] W. M. Jackman, "YouTube Usage in the University Classroom: An Argument for its Pedagogical Benefits," Int. J. Emerg. Technol. Learn., vol. 14, no. 9, 2019. https://doi.org/10.3991/ijet.v14i09.10475
[8] J. C. Richards, "Teaching English through English: Proficiency, Pedagogy and Performance," RELC J., vol. 48, no. 1, pp. 7–30, 2017. https://doi.org/10.1177/0033688217690059
[9] H. T. M. Nguyen and A. Burns, Teacher Language Proficiency and Reform of English Language Education in Vietnam, 2008-2020. Phnom Penh: IDP Education (Cambodia) Ltd, 2017.
[10] X. M. Ngo, "Diffusion of the CEFR among Vietnamese teachers: A mixed methods investigation," Asian EFL J., vol. 19, no. 1, pp. 7–32, 2017.
[11] J. Read, "The Influence of the Common European Framework of Reference (CEFR) in the Asia-Pacific Region," Learn J. Lang. Educ. Acquis. Res. Netw., vol. 0, no. 0, pp. 33–39, 2014.
[12] R. A. Rashid, "Dialogic reflection for professional development through conversations on a social networking site," Reflective Pract., vol. 19, pp. 105–117, 2018. https://doi.org/10.1080/14623943.2017.1379385

8 Authors

Abdul Hakim Ali Abdul Aziz is a PhD (TESL) Researcher at Faculty of Languages and Communication, Universiti Sultan Zainal Abidin, Malaysia. Email: aliex87@gmail.com

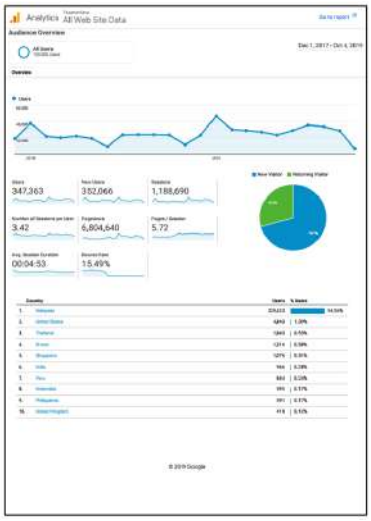
Wan Zhafirah Wan Zainudin is an Excellent Teacher (English Language) from Sekolah Kebangsaan Sultan Ibrahim Satu, Pasir Mas, Kelantan.

Dr Radzuwan Ab Rashid is the Deputy Dean of Research and Innovation at Faculty of Languages and Communication, Universiti Sultan Zainal Abidin, Malaysia.

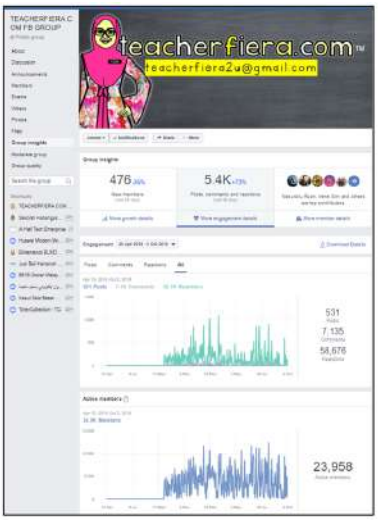
Article submitted 2019-06-30. Resubmitted 2019-10-04. Final acceptance 2019-10-07. Final version published as submitted by the authors.

9 Appendices

Appendix A



Appendix B



ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**TESL in-service teachers' experiences when evaluating an oral history workbook / Aziz, A. A., & Makhtar, R.**
Source	***3L: Language, Linguistics, Literature*** **Volume 27 Issue 3 (2021) Pages 56–72** **https://doi.org/10.17576/3L-2021-2703-04** **(Database: UKM Press)**

3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies
Vol 27(3), September 2021 http://doi.org/10.17576/3L-2021-2703-04

TESL In-Service Teachers' Experiences When Evaluating an Oral History Workbook

AZLINA ABDUL AZIZ
Faculty of Education,
The National University of Malaysia, Malaysia
azlina1@ukm.edu.my

ROSMIZAL MAKHTAR
Faculty of Education,
The National University of Malaysia, Malaysia

ABSTRACT

The implementation of the Common European Framework of Reference (CEFR) in English curriculum has led to the use of CEFR aligned foreign textbooks in Malaysian schools. However, the application of CEFR English textbooks from the United Kingdom has encountered a few oppositions as the advocates of Malaysian-based textbooks believe that English should be written contextually by emphasizing the local cultures and histories. In lieu of this, a group of in-service teachers in a Master's course had developed and evaluated each other's Oral History materials. The research objectives are to find out what are the material evaluation criteria frequently attended to by in-service TESL teachers when evaluating an oral history workbook and to what extent the in-service teachers have gained from evaluating oral history workbook developed by their peers. A mixed methodology research approach using basic frequency count, percentage value and qualitative data was employed in this study. Data was gathered from 109 in-service TESL teachers via their Personal Reflective Journal (PRJ), which they had written after evaluating an oral history workbook based on Tomlinson's and Mukundan's evaluation checklists. The findings identified the frequently attended criteria by the in-service teachers and that there are new criteria that can be added to the evaluation checklists. It is anticipated that this study could encourage educators to develop their own classroom materials, and material evaluators to consider the use of locally based English textbook and a revision to the material evaluation checklists to reflect current 21st century pedagogy.

Keywords: Common European Framework Reference; Material Evaluation Checklist; TESL; In-Service teachers

INTRODUCTION

The Common European Framework of Reference (CEFR) is currently recognized globally as the standard language proficiency framework to be utilized. Therefore, the Ministry of Education (MOE) of Malaysia has taken the initiative to utilize CEFR aligned foreign English textbook from the United Kingdom for English language instruction for both primary and secondary schools. According to Abdul Aziz et al. (2019), the Standard-Based English Language Curriculum for Secondary (SBELC) curriculum would utilize foreign English textbooks as the main source to teach the syllabus. The MOE aspires to produce students who can attain international standards of proficiency level; hence, the transformative initiative was taken.

However, the transformation has been widely discussed among educationalist, scholars and policymakers as it has disclosed mixed responses (Johar & Abdul Aziz, 2019). As the learners' intercultural skills have been embedded in the English curriculum, the main issue that lingers around is whether the foreign English textbook is able to fulfill the need of local learners (Abdul Rahim & Jalalian Daghigh, 2019). Moreover, in a study exploring the experiences of student-teachers in developing oral history texts by Abdul Aziz et al. (2019), they expressed the view that there is a hidden message that English texts produced in the UK by its native speakers are unmistakably, progressively superior in terms of language and content, and that

they are the only main experts of the English language. Additionally, the content of foreign English textbook does not emphasize the voice of the local people and culture. The absence of local cultural awareness in foreign English textbook should be further discussed and considered by policy makers. Therefore, the absence of learners' own culture in foreign English textbook has driven the researchers to explore the experiences of TESL in-service teachers when evaluating oral history workbook, which has been developed contextually, whereby it is based on local context and the Malaysian culture. In this study, the in-service teachers produced an oral history workbook, evaluated each other's work based on Tomlinson's and Mukundan's material evaluation checklists, and wrote reflections on their experiences evaluating the materials.

The workbook, which consists of a written Oral History text, lesson plans and activities, is used to develop the four skills i.e. reading, writing, listening and speaking. However, in this paper, we are only reporting on the in-service teachers' experiences when evaluating the oral history workbook rather than their experiences developing the materials as this study is an extension of a previous research conducted by Abdul Aziz et al. (2019). This research seeks to answer the following questions: 1) What are the material evaluation criteria frequently attended to by in-service TESL teachers when evaluating an oral history workbook? and 2) What are the experiences of in-service TESL teachers when evaluating an oral history workbook?

LITERATURE REVIEW

CONCEPTUAL FRAMEWORK

There are three theories underpinning this study. They are Material Evaluation Principles, Contextualized Learning Materials and Oral History as Part of Cultural Awareness.

MATERIAL EVALUATION PRINCIPLES

Tomlinson (2013) defined material evaluation as a process that involves determining the value of a set of instructional materials. He further explained that this process involves making decisions about the impact of the materials on users of instructional materials and this process attempts to measure some criteria by using material evaluation checklists. Tomlinson (2013, p. 2) has provided a set of material evaluation checklists which include:

1) the appeal of the materials to the learners,
2) the credibility of materials to learners, teachers, and administrators,
3) the validity of the materials,
4) the reliability of the materials,
5) the ability of the materials to interest the learners and the teachers, and
6) the ability of the materials to motivate the learners.

The criteria listed indicate that the needs of learners should be identified before designing and developing material evaluation checklist.

Researchers believed that the validity and reliability of the material evaluation checklist should be verified by experts before it is able to be used by material evaluators (Ali, 2018). Based on the criteria provided by Tomlinson, teachers or instructional material developers can adapt and adopt it when developing their instructional materials according to their students' language needs and proficiency levels. Tomlinson also suggested that material evaluation checklist should be established before materials are created, and it is used to make decisions about the approach, procedures and activities to be adopted as well as to observe the material development and its consequent use [as cited in Maroko, 2013).

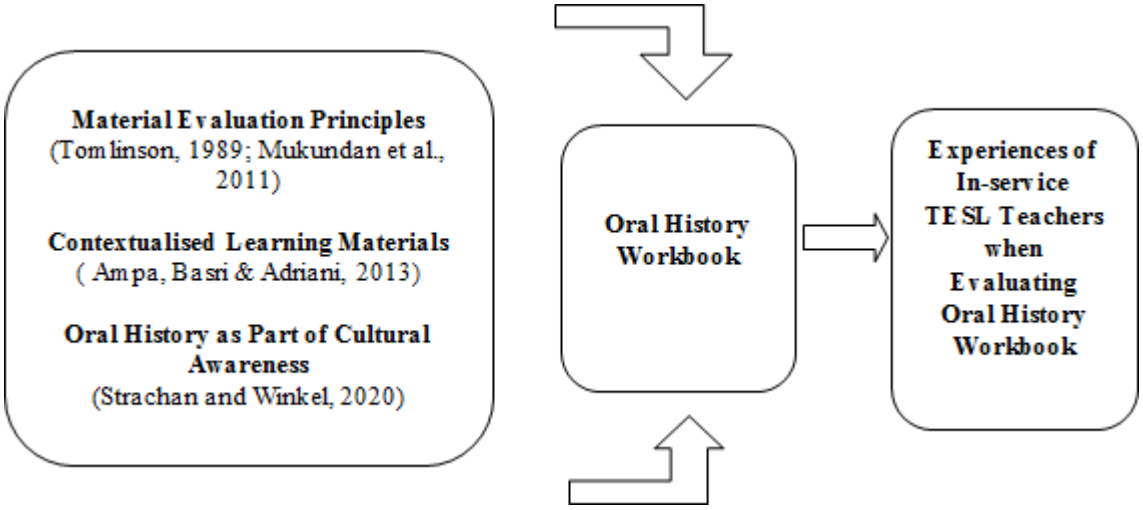


FIGURE 1. Conceptual Framework

The principles of material evaluation were then further expanded by Mukundan et al. (2011). Mukundan et al. (2011) examined the principles of material evaluation by developing an English language textbook evaluation checklist. The purpose of the study was to improve on Tomlinson and Mukundan et al's evaluation checklist. The results of the study are anticipated to be beneficial for English as a second language (ESL) teachers, curriculum designers, and instructional material developers and evaluators. Mukundan and Nimehchisalem (2012) believed that the assessment is made simpler, more objective and effective when it depends on a substantiated instrument. This indicates that teachers as instructional material developers and consumers can adopt and adapt the material evaluation checklist from Tomlinson and Mukundan et al. to cater to students' language needs as well as to achieve their teaching and learning objectives. In this research, we had utilized both Tomlinson's (Appendix 1) and Mukundan's checklists (Appendix 2) to evaluate Oral History materials by a group of TESL in-service teachers.

CONTEXTUALIZED LEARNING MATERIALS IN ENGLISH CLASSROOM

Learning materials are commonly the most significant part of language pedagogy considering they provide language input and exercise necessary for language learning to occur (Ampa et al., 2013). Meanwhile, contextualization can be characterized as an effort to create a real-life situation based on students' experiences of their real lives circumstance (Rohayati, 2013). Ampa et al. (2013) defined contextualized learning material as materials that allow learners to process new data and knowledge on their own based on previous reference. This suggests that learners naturally seek meaning in context that corresponds to their schemata. Furthermore, contextualized material is believed to be able to offer a promising method in assisting learners to learn more efficiently (Bakeret al., 2009). This is because contextualized materials are able to help students to relate the content of what they are learning to the authentic life content that they have experienced.

A good learning material can guarantee certain desirable results such as in encouraging learners to become more independent, even in circumstances where educators may not be prepared to teach effectively (Ahmed, 2017). Besides, the first thing that should be considered when developing contextualized texts is the inclusion of recognizable situations whereby the content includes situations that are realistic to learners (Saqlain et al., 2014). The main aim in utilising contextualised materials is to offer learners such models that were straightforward and have a place with their own local context. Language input in learning materials also cannot merely be

understandable but also memorable to the point that language is contextualized by utilizing germane themes for students (Bourke, 2006).

A research conducted by Rohayati (2013) revealed that the application of contextualization was effective in enhancing students' vocabulary mastery. This showed that contextualization does play a significant role in improving students' language learning. Moreover, Rohayati (2013) also suggested the teachers who intend to carry out contextualization need to choose a proper context based on students' experiences and real live situations and to create a lively classroom atmosphere as well. Therefore, the suitability and effectiveness of language learning materials deserve critical considerations since they are such a key element of language classroom (McGrath, 2013).

Howard and Major (2004) also mentioned that there are a few factors that may influence English language teachers to develop their own teaching materials, regardless of the accessibility of commercially produced materials. Factors that affect teachers include contextualization, individual needs, personalization and timeliness (Howard & Major, 2004). Contextualization is a necessity when it comes to teacher-designed materials because commercially produced materials do not aim at any target group of students, nor particular cultural or informative context. This is clearly stated by Howard and Major (2004) that the current learning materials are loaded with discourse acts and events of which most English language learners will never experience.

ORAL HISTORY AS PART OF CULTURAL AWARENESS

Oral history is the progression of recording, safeguarding, and spreading our understanding of the past through life stories. It also stimulates the idea of learning from another person's experiences and enables students to link the learning content to their experiences in real life where it would match the interest in reading among the students (Abdul Aziz et al., 2019). Therefore, the use of oral history workbook in English language pedagogy can be a part of cultural awareness among second language learners in Malaysia. The voice of local people and culture shared in the oral history workbook can encourage learners to explore and gain a deeper understanding of other people's cultures. Malaysia, which is a home to multi-ethnic groups such as Malay, Chinese, Indian, Iban, Bidayuh and Kadazan can offer many stories based on its multi-ethnic cultures. Additionally, the utilization of oral history workbook in the English classroom can also bridge the gap between different cultures and, in a way, challenge the underlying discrimination which is not commonly brought up and debated in professional settings. Cultural appreciation can also be imparted among language learners through oral history workbook as learners gain general knowledge of other cultures.

According to Strachan and Winkel (2020), oral history has become a vital instrument in academic circles in collecting first-hand experience. Anthropologists and historiographers, specifically, have utilized oral narratives for a considerable length of time as an instrument for examination and information assortment whereby first-individual records widen the extent of a request as they cast a focus on an individual or lived experience, providing nuanced understandings of lifeways and occasions over a significant time span. Additionally, a new viewpoint will be offered to learners as the culture and identity of others were integrated into oral history texts (Burgo, 2016). This implies that oral history can inculcate culture and identity awareness among students as they get to be exposed to other people's first-hand experiences. This claim is supported by Tseng (2002) as he believed that culture could change individuals' perceptions and it is important to broaden individuals' perceptions of the world (as cited in Burgo, 2016).

METHODOLOGY

RESEARCH DESIGN

This research employed a mixed methodology approach using basic frequency count, percentage value and qualitative data. The research design is beneficial in obtaining a deep understanding of the relationship between quantitative outcomes and qualitative discoveries of the in-service teachers' experiences in evaluating an oral history workbook.

RESEARCH SITE AND PARTICIPANTS

The research was carried out in one of the public universities in Malaysia. The participants of this research were 109 TESL in-service teachers, who had enrolled in the course, 'Current Trends in Language Teaching Research' at the Faculty of Education, Universiti Kebangsaan Malaysia. In the course, they were required to create an oral history workbook based on the themes in the Standard-Based English Language Curriculum for Secondary Schools (SBELC). The Oral History workbook could serve as an additional workbook for the current English textbook. As they worked in groups, they were required to interview a member of the community based on the themes 'People and Culture' and 'Health and Environment', which could be found in the lower secondary form curriculum. Stories from local community would provide sources of authentic and contextualized narratives and thus help to raise cultural awareness amongst their pupils. Once they have transcribed the interview, the data was rewritten as an Oral History text, in narrative form, from the interviewee's perspective. Once they have created this Oral History text, they were asked to develop four lesson plans based on reading, writing, listening and speaking skills. The Oral History workbook they developed were exchanged between groups and evaluated using Tomlinson's and Mukundan et al.'s evaluation checklists. They were required to write a written reflection of their experiences in evaluating oral history workbook based on evaluation checklist adapted from Tomlinson (2008) and Mukundan et al. (2011). Before the researchers conducted the research, a consent letter was given to each in-service teacher to gain their consent to be the participants of the study. In addition, pseudonyms were also used in order to protect the participants' identities.

DATA COLLECTION AND DATA ANALYSIS

A duration of four weeks was given to the participants to come out with a written reflection of the evaluation of each other group's Oral History material. The participants were required to share their experiences in evaluating oral history workbook based on Tomlinson's (2008) and Mukundan et al.'s (2011) material evaluation checklists. The participants did not necessarily attend to all the criteria in the checklists when evaluating the workbook as there are many items. Instead, they attended to criteria that they found to be important, and they were also encouraged to include other criteria that were not mentioned in the checklists. This process revealed the practicality of the checklists as the participants attached more significance to some criteria over the others when they reported their material evaluation experiences. The data were collected from 109 Personal Reflective Journal (PRJ) by the end of the 'Current Trends in Language Teaching Research' course. Each participant was asked to write a PRJ discussing the material evaluation criteria they had attended to when evaluating the oral history workbook. The researchers analyzed the data thematically by reading it line by line, which is then followed by creating codes that helped the researchers to determine the occurrence of themes related to the study. This thematic analysis approach searches and extracts the common patterns through multiple data readings (Yukhymenko et al., 2014). The occurrence of themes was recorded using basic frequency count to show the themes distribution patterns. The frequency count

would reveal which criteria of the material evaluation checklist were given greater importance by the in-service teachers as well as addition of new criteria that was not present in the material evaluation checklist. In total, 109 PRJs were collected, coded and analysed thematically following the material evaluation checklists as a guide.

FINDINGS

This study aims to explore the experiences of in-service teachers when evaluating oral history workbook based on Tomlinson's and Mukundan et al.'s material evaluation checklists. Based on the checklists, there are a few themes that can be derived, namely suitability for learners, teaching and learning content, and development of English language. This was developed after the data were rigorously analysed and the similarities in the criteria listed by Tomlinson (2008) and Mukundan et al. (2011) were identified. Based on the findings, the researchers also added three new criteria that were not mentioned in both material evaluation checklists. In the teaching and learning content theme, the 2 new criteria are i) To what extent do the activities/ lesson plans adhere to the CEFR? and ii) To what extent do the materials/ activities/ lesson plans utilize technology and 21stcentury learning? In the development of the English language theme, the new criterion is iii) 'To what extent do the materials/ activities enrich students' vocabulary'. The tables below (Table 1 – 3) shows the frequency count of each criterion mentioned in the 109 PRJs based on the respective main three themes.

SUITABILITY FOR LEARNERS

The first theme, 'Suitability for learners' was identified by grouping a few material evaluation criteria that fit within this main theme. Within this main theme, the in-service teachers are asked to examine how the material they are evaluating would be suitable for students in terms of its engagement, meaningfulness, and how interesting and challenging the materials are. Based on Table 1, the majority of the participants gave priority to 'To what extent are the texts likely to interest the learners' when evaluating the oral history workbook with the highest percentage of 54.1%. This shows that reading texts that would interest the learners are given priority by the participants when evaluating the workbook. In order to enhance students' reading skills, reading texts should be able to entice students to read more and better as well as to engage students with more relatively ubiquitous topics. The more students are able to improve their reading skills, the more knowledge students are able to grasp (Akbaşlı et al., 2016).

TABLE 1. Suitability for Learners criteria

Suitability for Learners	N	Percentages (%)
To what extent is the exposure to English in use likely to be meaningful to the target learners?	26	23.9
To what extent are the texts likely to interest the learners?	59	54.1
To what extent are the activities likely to engage the target learners cognitively?	44	40.4
To what extent are the activities likely to engage the target learners affectively?	16	14.7
To what extent are the activities likely to provide achievable challenges to the learners?	50	45.9

Meanwhile, the second most frequently cited criterion was 'To what extent are the activities likely to provide achievable challenge to the learners' with 45.9% occurrence. The participants considered this criterion as important because in order to advance language learners to the next level, challenges should be part of learning. Providing achievable challenges in language learning will also encourage learners to stay motivated to learn and improve their language skills.

The third frequently cited criterion that was given attention to by the TESL in-service teachers was, 'To what extent are the activities likely to engage the target learners cognitively'. This criterion comprises 40.4% occurrence. Language learners who are able to relate with the activities cognitively are usually exposed to meaningful learning. It is essential for learners to perceive activities as being meaningful because it will help the learners to retain the knowledge they gain for a long time. A study has shown that if learners do not deem a learning activity deserving of their time and effort, they probably would not take part in a satisfactory manner or may even withdraw from joining the activity altogether (Fredricks et al., 2004).

Next, the fourth frequently cited criterion used by the participants was 'To what extent is the exposure to English in use likely to be meaningful to the target learners' with a frequency of 23.9%. Al-Zoubi (2018), emphasized that the exposure to English among second language learners should be done and encouraged by educators in order to help learners improve their English language as it can affect learners' self-ability and apprehension of the target language. Therefore, the participants of this study have also utilized this criterion when assessing the materials.

The least attended criterion in this theme was 'To what extent are the activities likely to engage the target learners affectively' with 14.7% occurrence. This suggests that affective strategy is given less emphasis by the participants when evaluating the suitability of materials to teach English.

TEACHING AND LEARNING CONTENT

TABLE 2. Teaching and Learning Content criteria

Teaching & Learning content	N	Percentages (%)
To what extent do the activities/ lesson plans adhere to the CEFR?	54	49.5
To what extent do the materials/ activities/ lesson plans utilize technology and 21stcentury learning?	56	51.4
To what extent do the materials provide opportunities for cultural awareness?	32	29.4
To what extent are the materials likely to sustain positive impact?	18	16.5

Table 2 shows that the criterion that was most frequently used by the participants when evaluating the oral history workbook was 'To what extent do the materials/ activities/ lesson plans utilize technology and 21st century learning' with 51.4% occurrence. This is a new criterion that was identified by the researchers when analyzing the data and it is not listed in Tomlinson's and Mukundan et al.'s material evaluation checklists. The criterion was added in the checklist because the participants repeatedly mentioned it in their written reflections, and this shows that the participants also considered other criterion when evaluating the materials. As the world is gearing up towards technology-based education, the implementation of digital tools in the teaching materials, activities and lesson plans has been emphasized as it is the current need of the learners. The utilization of technology in English language learning is believed to be able to help the learners to learn faster and easier as everything is accessible and learnable (Sarica & Cavus, 2009). To add, in order to optimize students' learning through technology, teachers play a significant role in using technology in the classroom (Mofareh, 2019).

The second commonly cited criterion that was given attention to by the participants was 'To what extent do the activities/ lesson plans adhere to the CEFR' with 49.5% response. This new criterion is also included by the researchers when analyzing the data. As the MOE decided to implement CEFR in 2013 as an initiative to reform English language education, the participants of this study also considered this new criterion as important when assessing materials. The availability of content or learning standards in the oral history workbook are

also vital because the materials were developed with an aim to meet students' needs.

Meanwhile, the third regularly cited criterion that was used by the participants was 'To what extent do the materials that provide opportunities for cultural awareness' (29.4%). Most of the oral history texts developed by the participants undeniably evoke and reiterate the cultures of other people. Hence, this criterion is significantly important even though the participants did not mention this criterion as frequently as the previous ones. Oral history texts also allow the readers to explore and learn about other people's stories which are not mentioned in the historical records (Walbert, n.d.). Therefore, oral history texts, especially texts that are developed locally need to be given greater recognition as it offers more interesting local stories than the ones in the foreign English textbook.

The criterion that was the least frequently attended to by the participants was 'To what extent are the materials likely to sustain positive impact' where only 16.5% of responses was recorded. It can be concluded that this criterion is the least preferred criterion in this theme because the participants put more emphasis on cultural awareness, which is more relevant to the oral history workbook.

DEVELOPMENT OF ENGLISH LANGUAGE

TABLE 3. Development of English Language criteria

Development of English language	N	Percentages (%)
To what extent do the materials provide exposure to English in authentic use?	17	15.6
To what extent do the materials provide opportunities for learners to gain feedback on their effectiveness of their use of English?	10	9.2
To what extent do the activities provide opportunities for meaningful use of English?	42	38.5
To what extent do the activities provide opportunities for learners to make discoveries about how English is used?	5	4.6
To what extent do the materials/ activities plans enrich students' vocabulary?	42	38.5

Two criteria that scored the highest value in this theme were: 1) 'To what extent do the activities provide opportunities for meaningful use of English', and 2) 'To what extent do the materials/ activities enrich students' vocabulary'. Both have a frequency of 38.5% respectively. The latter is a new criterion that was added by the researchers as the participants repetitively mentioned the criterion in their written reflections.

The second most frequently cited criterion that was used by the participants in evaluating oral history workbook was 'To what extent do the materials provide exposure to English in authentic use' (15.6%). This shows that a great emphasis should be placed on the use of authentic materials. The use of authentic materials in English classrooms can make the learning process significantly more engaging, imaginative and stimulating for learners (Zazulak, 2017). Similarly, authentic materials can be practical for the teachers to elicit authentic reactions from learners.

Meanwhile, the third most frequently cited criterion that was given attention to by the in-service teachers was 'To what extent do the materials provide opportunities for learners to gain feedback on their effectiveness of their use of English' (9.2%). This criterion highlights the responses that the learners can get after they have used the materials. However, based on the finding, it shows that the participants did not put much emphasis on this criterion. Klimova (2015) stated that feedback can be given formally or informally in English classrooms and the feedback-giving sessions should be established in an improvement-oriented manner. Teachers also should give feedback in a neutral and positive way to provide a good impact on students (Reynolds, 2013).

The least employed criterion in this theme was 'To what extent do the activities provide opportunities for learners to make discoveries about how English is used' with a frequency of

4.6%. Among all the criteria mentioned, this criterion is the least preferred among the participants when evaluating the materials.

DISCUSSION

The findings above found that the theme 'suitability for learners' was the most important criteria used by the TESL in-service teachers to evaluate the oral history workbook. Within this main theme, the criterion that was frequently mentioned by the participants in their PRJs was 'To what extent are the texts likely to interest the learners'. Meanwhile, the least attended theme in the material evaluation checklist was the development of English language and the subtheme that was least mentioned was 'To what extent do the activities provide opportunities for learners to make discoveries about how English is used'. In this section, we are only discussing the most attended criteria listed in each theme.

SUITABILITY FOR LEARNERS – THE VALUE OF ORAL HISTORY TEXTS

The findings of this study indicate that 54.1% of the respondents found the criterion 'how interesting the oral history text is' played a significant role during the evaluation process. A study by Horn (2014, p. 78) showed sufficient evidence that the utilization of oral history in English classroom demonstrates many benefits to learners which include "developing historical skills, content knowledge and historical context as it relates to significance and meaning of historical events". When evaluating the oral history workbook provided by the other participants, they examined the value of Oral History in improving pupils' knowledge of local culture and history and how these knowledge are suitable for the level of Form 2 pupils.

The majority of participants considered that the oral history texts developed by others were interesting and appropriate to students' level of knowledge. This aspect is evident in Akmal's written reflection. He stated,

> "Overall, the oral history texts written by members of Group 4 are appropriate for the intended learners. The authors have provided a clear storyline, and both appear appealing and informative to the readers".

One of the participants also mentioned that the oral history texts developed by their classmates were authentic as they represented real narratives and this can be observed from Nabila's written reflection,

> "Based on our evaluation for Group 9's work, I could say that both oral history texts were interesting to read and portrayed real stories from real people. It was not something that pupils can find in their textbook".

They also realized that the narrative form used to develop oral history texts makes the texts more interesting to pupils as Vienna mentioned,

> "The story telling technique used by the author was captivating and it made the readers hunger for more stories. Overall, the text managed to grasp the attention and interest, both at the same time with excellent word choices, suitable for Form 2 readers".

Oral History texts in the workbook, which are developed by the TESL in-service teachers, provide authentic real-life experiences of people within the community. Students will find the text relatable, informative, interesting and affirming. It challenges the stereotypes and the generic representations of communities often found within textbooks written for the

purpose of brevity and simplicity. The Oral History texts in the Oral workbook present a more nuanced and complex representations of individuals with the community written in a narrative, appealing format.

TEACHING AND LEARNING CONTENT – INTEGRATION OF TECHNOLOGY AND 21ST CENTURY LEARNING

The ever-evolving world has seen a constant progress in the education system along with the introduction of new theoretical approaches which are deemed fit in catering to the needs of second language learners. The 21st century classroom has introduced several sets of skills which can be implemented by teachers in order to improve their teaching as well as learners' experiences in the ESL classroom. The participants stated that language acquisition can be encouraged through the integration of the 21st century skills which comprises of the 4Cs, namely communication, collaboration, critical thinking, and creativity to connect the learners with the language content taught in the classroom. This can be seen from Ain's written reflection. She stated,

> "Next, the other group did include the four skills in the 21st century classroom which were communication, creativity, collaboration, and critical thinking. This had shown that they were following the latest format of the lesson plan. The skills were clearly shown in their gallery walk activities, brainstorming, and writing the ending of a story".

In order to ensure that the integration of technology and 21st century learning can be implemented efficiently, teachers should consider the availability of facilities in schools that can aid students' learning. School facilities such as the internet, LCD projectors, and laptops should be accessible for teachers and students to use for academic purpose. Badriah mentioned it in her written reflection,

> "…I find some of their lesson plans don't cater to the schools in the interiors. The activities they suggested require access to the internet and laptops. Being a teacher froma suburban area, this is still an issue, I wonder how the teachers in rural area will implement the activities in their classrooms. Plus, the group didn't provide any other activity for the lesson in case technology fails. They should have considered this before they come up with the lesson plan since this has been mentioned in the first lecture".

Richards and Rodgers (2001) pointed out that learners who are involved in cooperative context are encouraged to take charge of their own learning process throughout the phases of planning, monitoring as well as assessment. Besides, students of today are known as digital generation which suggests that they have advanced technological literacy skills that can be beneficial for their language development (Pazilah & Hashim, 2018). Therefore, they are very much drawn to the incorporation of technology in ESL classroom because the use of gizmos is very enticing to them [34]. Based on the participants' written reflections, they stated that the oral history workbook, which incorporated 21st century learning, allows learners to do self-learning and self-exploration in academic setting. Kasturi wrote that,

> "After reading this text, I am very sure that pupils who do not know much about this topic will definitely do some reading on this topic, which is a good thing. It helps pupils to do some self-learning. This will integrate the 21st century-based learning as pupils need to do things on their own and not solely depend on their teachers.",

Similarly, Harry stated that,

> "I agree with the evaluators as learner-centered teaching and learning strategies optimize pupils'

> opportunities to learn effectively in the class. Self-exploration during lesson promotes active learning and creative thinking. Henceforth, this must be taken into serious consideration as pupils must be allowed to learn and engage in a dynamic learning environment as a part of the requirement to develop effective and efficient 21st century learner".

Throughout the evaluation process, the participants of this study also learnt a lot from each other. They have gained invaluable experiences which can be beneficial for their future career prospect in terms of technology utilization in ESL classroom. They believed that the integration of technology will provide interesting activities to learners. One of the participants, Chua, expressed this aspect in his written reflection,

> "Moreover, I could also learn some of the creative yet innovative idea through evaluating lesson plans of another group. Some of the ideas are quite useful, fits in the learning objectives and on par with the 21st century learning. I would be able to refer some of the useful and interesting activities in my future lessons".

Meanwhile, Hanisah wrote in her written reflection regarding the usefulness of 21st century activities in education,

> "The lesson plans complied with almost all the criteria we looked through from the given list. We learnt a lot on how to go about planning for 21st century activities in the lessons. I learnt that a lot from evaluating other group's work. I even learn on how to set up activities using Quizlet, an online quiz. It was very easy and interesting to the teenagers".

Additionally, the integration of technology into literacy teaching promotes a more dynamic and interactive learning environment in terms of more collaboration of ideas based on peer feedback and positive interactions. This would drive students into deeper learning and exploration with the language. If done correctly, the chances of achieving active and meaningful learning will be likely and students will be able to apply the knowledge obtained to any content area of the tasks given at hand (Zakaria & Abdul Aziz, 2019). This is evident in Catriona's written reflections, as she wrote,

> "After my group reviewing the weaknesses in our lesson plans, it has prompted us to look into efforts to overcome them. We need to rethink some alternative teaching strategies which reflect 21st century learning as well as revise the instructions in the lesson plans to make them more explicit. Therefore, feedback by peers is invaluable as it helped my group to look at our own developed materials and lesson plans in order to discover our strengths and weaknesses and re-evaluate them".

The use of 21st learning and technology has become an important focus for teachers when creating teaching and learning materials since they are stated and emphasised in the Malaysian English curriculum. It can no longer be excluded from material development and evaluation checklists as they are shown to be given priorities by the in-service teachers even when these two criteria were not listed in the material evaluation checklists. Thus, the checklists could be updated to include 21st century learning and use of technology.

DEVELOPMENT OF ENGLISH LANGUAGE – MEANINGFUL LANGUAGE LEARNING ACTIVITIES

Many researchers believe that the ability to use English language effectively is an added value that will help students to move into their career phase where current employers are seeking potential employees who are proficient in English language (Md Yunus et al. 2013). Therefore, in order to expose students to meaningful use of English, the majority of participants agreed that language learning activities should be established according to students' language proficiency and preferences. One of the participants, Jane, agreed on this aspect as she wrote,

> "I also learnt that at times, even though the materials might be suitable to our own preferences, it might not be other people's cup of tea. Therefore, instead of creating materials based on our own liking, we should always prioritize our learners' needs in their learning process".

They also realized the importance of placing consideration for students with different abilities. Nasha wrote, "The next lesson that I obtained would be the importance of developing learning materials which are at par with pupils' proficiency levels by taking into consideration the intermediate and low proficiency pupils if it is a mixed-ability class".

Moreover, the participants also mentioned about the importance of employing student-centered learning in the English classroom where it allows students to be independent learners. As mentioned by Harry,

> "As an educator, it is necessary to have a progressive lesson with the incorporation of learner-centered teaching and learning strategies. From the given critique, these aspects are crucial to induce discovery, collaborative and even independent learning".

Other than that, teachers also play an important role in providing appropriate learning materials to students to ensure that they can experience engaging and meaningful learning during classroom activities. Amila addressed this view in her written reflection,

> "All in all, this project has given me an insightful experience as we are exposed to the secondary school syllabus which showed that the topics and themes are to be more serious and should not be taken lightly. From this, my mind is now widely opened and think that classroom activities are not merely taken as it is from the textbook. As teachers, we need to be wiser in providing the materials that make learning more meaningful and relevant for the pupils as it is crucial for the growth in their teen years".

One of the participants also stated that the teachers should be fully prepared in terms of lesson plans and activities. This is to ensure a smooth transition of the learning activities and that teachers can handle any circumstances that might interrupt the flow of the lesson. This is illustrated by Sofea,

> "Besides that, after evaluating other group's oral history texts, lesson plans and materials, I realized that a lesson plan is the basis of everything. A lesson plan ensures that pupils acquire knowledge through activities that have appropriate teaching and learning strategies. If a teacher teaches without preparing a lesson plan beforehand, a lesson might become havoc and he would not be prepared to overcome problems faced in a classroom".

STUDENTS'VOCABULARY ENRICHMENT THROUGH THE USE OF ORAL HISTORY WORKBOOK

Vocabulary is important in English language learning because it is the backbone of proper language command. Students who can master vocabulary surely have little to no problem in English reading proficiency. According to Huckin (1995), English language learners depend on vocabulary knowledge to a great extent and the scarcity of that knowledge is the greatest hurdle for learners to get over [as cited in Alqahtani, 2015). Therefore, to enrich students' vocabulary, the addition of a glossary at the back page of oral history texts should be encouraged. The participants were aware that glossary is important in reading materials as Suresh wrote in his written reflection,

> "The selection of vocabularies for instance wasn't on par with the glossary or list of words that students in lower secondary should acknowledge. For example, there were terminology that described the process of funeral and death occurrence in the Iban community. As readers, we might wonder what those terms mean. The writer should have listed these terms and defined them for references purposes".

Meanwhile one of the participants commented on how the absence of glossary may affect readers that come from different cultures. Agnes wrote,

> "The only thing they forgot to include is the glossary. Glossary is very important because they are using unfamiliar words in the text and students who are going to read the text is not just Iban. So, they need the glossary to find the meaning of the words in the text".

Aside from attempting to meet students' language proficiency, the use of complex words in reading activity can give more vocabulary exposure to the students. Quoting from one of the participants' reflections, Amy stated, "For the Reading Skill, we could also use simpler words intended for the pupils to understand easier but using complex words can actually expose pupils to more vocabularies". Interestingly, the participants also learnt a lot of new vocabularies from the evaluation process, as one of them mentioned;

> "Throughout the evaluation of their oral history texts, I learned quite a lot of new vocabularies and phrases from their texts. The biggest challenge for me is to watch for their errors especially in the choice of words and the sentence structures", Adlin

Moreover, a participant believed that teaching vocabulary should not be differentiated according to students' language proficiency level. Instead, teachers should increase the difficulty of vocabulary learning so that students can enhance their vocabulary expansion. Also, standardizing vocabulary learning can be implemented to help the underachieving students to feel motivated to learn and strive harder in English language classroom. Ravitha wrote this in her written reflections,

> "I do not believe in having an easier material or easier vocabulary to the weaker students, but the teacher needs to provide more guidance and motivate the pupils to try better in the class. I could relate this during one of my practicum lesson when I provided an easier version of the task sheet and when she compares it to her classmates who got a difficult one, she got demotivated as she felt that I am having a lower expectation on her compared other friends. This saddens her and she stopped attempting my task. This could be one of the reasons why I do not accept the suggestion and write only one set of vocabulary for all the pupils instead of each set for each level".

In contrast, one of the participants believed that teachers should not hasten students' language learning by presenting too many new vocabularies as it would have made low-proficiency students feel intimidated to learn the words. This is evident from Nasha's written reflection, "The other group has created a lengthy oral history text with difficult vocabularies which might be incomprehensible for under-achievers. In my opinion, a teacher should not rush pupils' learning by introducing many new vocabularies in one lesson as this would scare them".

CONCLUSION

The participants' PRJs revealed the criteria they frequently attended to as they evaluated the Oral History workbook using Tomlinson's and Mukundan et al.'s material evaluation checklists. The experiences of in-service TESL teachers when evaluating an oral history workbook disclosed 3 important criteria that can be added to Tomlinson's and Mukundan et al.'s material evaluation checklists. The three added criteria that were found in the findings were:

1) 'To what extent do the activities/ lesson plans adhere to the CEFR';
2) 'To what extent do the materials/ activities/ lesson plans utilize technology and 21st century learning' and;
3) 'To what extent do the materials/ activities enrich students' vocabulary'.

These criteria were not included in the two evaluation checklists, but were regarded as important criteria by the TESL in-service teachers. The criteria are relevant to the current education scenario as 21st century learning and the use of technology are emphasised in the Malaysian CEFR aligned English curriculum. Furthermore, the vocabulary in Oral History texts provide examples of socially and culturally specific words which could be used to describe the experiences of students. Thus, we would like to suggest that it can be included in the material evaluation checklist. Despite having different opinions in certain aspects of the checklists, the participants have gained invaluable experiences throughout the evaluation process. The input that the participants have obtained certainly will hone participants' skills in English language teaching and encourage them to become better teachers. The three themes derived from the findings can be a set of future references for English teachers and material evaluators to assist them to develop a good English workbook.

Based on the findings, it is hoped that this study will assist curriculum designers and English instructional material developers to incorporate local culture in their English instructional materials. It is also anticipated that this study can be beneficial for English policy makers to reconsider the choice of instructional materials used in local English classrooms and to evaluate them contextually. This is to ensure that ESL learners can benefit much more from the instructional materials which are based on local content that promote local cultures and histories. ESL teachers should also take part in designing instructional materials that promote local stories which are closely related to their students' lives. This will provide students with an engaging and meaningful English language learning while helping them to have better retention of the language learned.

ACKNOWLEDGEMENTS

The researchers would like to thank the National University of Malaysia (UKM) for its support for this research via its Grants GG-2019-072 Dana Penyelidikan FPEND, TAP-K017393 Tabung Agihan Penyelidikan (TAP) and GGPM-2017-126.

REFERENCES

Abdul Aziz, A., Narayanasamy, J., Mohamad, M., & Abdul Hameed, H. (2019). Student-teachers' experiences in creating oral history texts. *3L: The Southeast Asian Journal of English Language Studies*, *25*(3), 79-89.

Abdul Rahim, H. & Jalalian Daghigh, A. (2019). Locally developed versus global textbooks: an evaluation of cultural content in textbooks used in English language teaching in Malaysia. *Asian Englishes*, *22*(3), 317-331. doi:10.1080/13488678.2019.1669301

Ahmed, S. (2017). Authentic ELT materials in the language classroom: An overview. *Journal of Applied Linguistics and Language Research*, *4*(2), 181-202.

Akbaşlı, S., Şahin, M. & Yaykiran, Z. (2016). The effect of reading comprehension on the performance in science and mathematics. *Journal of Education and Practice*, *7*(16), 108-121.

Al-Zoubi, S. M. (2018). The impact of exposure to English language on language acquisition. Journal of *Applied Linguistics and Language Research*, *5*(4) 151-162.

Ali, I. (2018). ELT materials evaluation: A system and criteria. *Theory and Practice in Language Studies*, *8*(7), 797-812.

Alqahtani, M. (2015). The importance of vocabulary in language learning and how to be taught. *International Journal of Teaching and Education*, *3*(3), 21-34.

Ampa, A. T., Basri, M. & Adriani, A. A. (2013). The development of contextual learning materials for the English speaking skills. *International Journal of Education and Research*, *1*(9), 1-10.

Baker, E. D., Hope, L. & Karandjeff, K. (2009). *Contextualized teaching & learning: A promising approach for*

basic skills instruction. The Research & Planning Group for California Community Colleges.
Bourke, J. M. (2006). Designing a topic-based syllabus for young learners. *ELT Journal*, *60*(3), 279–286.
Burgo, C. (2016). Oral History as an innovative language teaching technique for Spanish heritage language learners. *International Journal of Teaching and Learning in Higher Education*, *28*(3), 451-459.
Fredricks J. A., Blumenfeld, P.C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1) 59-109.
Hajrulla, V. (2012). Bringing cultural content and authentic materials to enhance problem-based learning in EFL classes. *Lingua Mobilis*, *5*(38), 95-108.
Hashim, H., Md Yunus, M. & Embi, M. A. (2016). Pre-University English as second language (ESL) learners' attitude towards mobile learning. *Creative Education*, *7*(8), 1147-1153. http://dx.doi.org/10.4236/ce.2016.78119
Horn, K. (2014). Oral history in the classroom: Clarifying the context through historical understanding. *Yesterday & Today*, 11, 72-83.
Howard, J. & Major, J. (2004). Guidelines for designing effective English language teaching materials. *The TESOLANZ Journal*, 12, 50-58.
Johar, N. A. & Abdul Aziz, A. (2019). Teachers' Perceptions on Using the Pulse 2 Textbook. *Journal of Educational Research & Indigenous Studies*, 2(*1*), 1-15.
K. Walbert. Oral history in the classroom. (n.d.). http://www.aughty.org/pdf/oralhist_classroom.pdf
Klimova, B. (2015). The role of feedback in EFL classes. *Procedia - Social and Behavioral Sciences*, 199, 172-177.
Maroko, G. M. (2013). Development of language materials for national development: A language management perspective. *International Journal of Education and Research*, *1*(7), 1-14.
McGrath, I. (2013). *Teaching materials and the roles of EFL/ESL teachers: Practice and theory*. Bloomsbury.
Md Yunus, M., Nordin, N., Salehi, H., Sun, C. H., & Embi, M. A. (2013). Pros and cons of using ICT in teaching ESL reading and writing. *International Education Studies*, *6*(7), 119-130.
Mofareh, A. (2019). The use of technology in English language teaching. *Frontiers in Education Technology*, *2*(3),168-180.
Mukundan, J. & Nimehchisalem, V. (2012). Evaluative criteria of an English language textbook evaluation checklist. *Journal of Language Teaching and Research*, *3*(6), 1128-1134.
Mukundan, J., Nimehchisalem, V. & Reza, H. (2011). Developing an English language textbook evaluation checklist: A focus group study. *International Journal of Humanities and Social Science*, *1*(12) 100-106.
Pazilah, F. & Hashim, H. (2018). Using infographics as a technology-based tool to develop 21st century skills in an ESL context. *Journal of Educational and Learning Studies*, *1*(1), 35-38.
Reynolds, L. (2013, June 11). Giving student feedback: 20 tips to do it right. *informED*. https://www.opencolleges.edu.au/informed/features/giving-student-feedback/
Richards, J. C. & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge University Press. 10.1017/cbo9780511667305.021
Rohayati, T. (2013). The implementation of contextualization in teaching vocabulary to elementary students. *Journal of English and Education*, *1*(2), 115-123.
Saqlain, N. U., Saad, I. & Islam, R. U. (2014). Impact of contextualized text on students' learning of writing skills at tertiary level. *JISR-MSSE*, *12*(1), 73-92.
Sarica, G. N. & Cavus. N. (2009). New trends in 21st century English learning. *Procedia Social and Behavioral Sciences*, *1*, 439-445.
Strachan, L. M. & Winkel, C. (2020). Oral history: An unpredictably effective strategy for teaching ESL to Saudi Arabian students. *Academic Journal of Interdisciplinary Studies*, 9(*1*), 158-168.
Tomlinson, B. (2013). *Developing Materials for Language Teaching*. Bloomsbury Publishing.
Tomlinson, B. (Ed.). (2008). *English language learning materials: A critical review*. Continuum.
Yukhymenko, M. A., Brown, S. W., Lawless, K. A., Brodowinska, K. & Mullin, G. (2014). Thematic analysis of teacher instructional practices and student responses in middle school classrooms with problem-based learning environment. *Global Education Review*, *1*(3), 93-109.
Zakaria, M. A. & Abdul Aziz. A. (2019). The impact of digital storytelling on ESL narrative writing skill. *Arab World English Journal*, *5*, 319-332.
Zazulak. S. (2017, April 26). *Using authentic material from the real world to teach English*. https://www.english.com/blog/authentic-material/

APPENDIX 1

Material evaluation guidelines (Tomlinson, 2008)

No.	Evaluation guidelines
1.	To what extent do the materials provide exposure to English in authentic use?
2.	To what extent is the exposure to English in use likely to be meaningful to the target learners?
3.	To what extent are the texts likely to interest the learners?
4.	To what extent are the activities likely to provide achievable challenge to the learners?
5.	To what extent are the activities likely to engage the target learners affectively?
6.	To what extent are the activities likely to engage the target learners cognitively?
7.	To what extent do the activities provide opportunities for learners to make discoveries about how English is used?
8.	To what extent do the activities provide opportunities for meaningful use of English?
9.	To what extent do the materials provide opportunities for learners to gain feedback on their effectiveness of their use of English?
10.	To what extent are the materials likely to sustain positive impact?
11.	To what extent do the materials help the learners to make use of the English speaking environment outside the classroom?
12.	To what extent do the materials help the learners to operate effectively in the English speaking environment outside the classroom?
13.	To what extent do the materials treat English as an international language?
14.	To what extent do the materials provide opportunities for cultural awareness?

APPENDIX 2

Material development checklist (adapted from Mukundan & Nimehchisalem, 2012)

No	Development Guidelines
1.	**Suitability of learners**
a.	It is compatible to the age of the learners.
b.	It is compatible to the needs of the learners.
c.	It is compatible to the interests of the learners.
2.	**Learning teaching content**
a.	Most of the tasks are interesting.
b.	Tasks move from simple to complex.
c.	Tasks objectives are achievable.
d.	Cultural sensitivities have been considered.
e.	The language in the textbook is natural and real
3.	**Reading**
a.	Texts are interesting.
b.	Texts are culturally relevant.
c.	Texts are culturally appropriate.
4.	**Vocabulary**
a.	The load (number of new words in each text) is appropriate to the pupil's level.
b.	There is a good distribution (simple to complex) of vocabulary load across the texts.
5.	**Activities**
a.	They are adequate.
b.	They are learner friendly.
c.	They match the syllabus specifications.
d.	The activities can be exploited fully.
e.	Activities can work well with methodologies in ELT.
f.	They help students who are under/overachievers.
g.	Activities are developed to initiate meaningful communication.
h.	Activities have achievable goals and take into consideration learner capabilities.
i.	Activities are balanced between individual response, pair work and group work.

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**The application of language proficiency scales in education context: a systematic literature review / Zhu, A., Mofreh, S. A. M., & Salem, S.**
Source	***SAGE Open*** **Volume 13 Issue 3 (2023) Pages 1-19** **https://doi.org/10.1177/21582440231199692** **(Database: ResearchGate)**

Original Research

SAGE Open
July-September 2023: 1–19
© The Author(s) 2023
DOI: 10.1177/21582440231199692
journals.sagepub.com/home/sgo
Sage

The Application of Language Proficiency Scales in Education Context: A Systematic Literature Review

Aihua Zhu[1], Samah Ali Mohsen Mofreh[1], and Sultan Salem[2]

Abstract

The application of the language proficiency scales (LPS) in education validates its function, as it can explore the value in-depth. However, little systematic research on applying LPS has been conducted due to the complex intertwining of stakeholders and a lack of theoretical framework and practical approaches. Adopting the framework proposed by Y. Jin and Jie (2020), this study explored how the Common European Framework of Reference (CEFR) and China's Standards of English (CSE) were used and impacted various stakeholders in the education context. The literature search was taken from WoSCC, Scopus, and CNKI from 2018 to 2022. Qualitative content analysis was used for systematic review. Results showed that policymakers used LPS in education policy guidance; teachers applied them as the benchmark of diagnostic assessment to get accurate language profiles of students and create new approaches to teaching; students used them as goal-setting guidance and self- or peer assessment criteria to track progress; test developers aligned them with tests to obtain reliable results; curriculum designers tailored descriptors and scales from CEFR to develop new curricula, align, or revise the existing ones; researchers used LPS as references to develop new rubrics, frameworks and assessing models. This study could provide insight in scientific application of LPS. However, it focused mainly on the CEFR and CSE with a framework for exploring the impact of language testing. Studies containing more scales and theorizing the framework of aftereffects of LPS should be encouraged.

Keywords

language proficiency scales, systematic literature review, stakeholders, education context, application

Introduction

Language proficiency scales (LPS) are extensively used for different purposes, such as course, syllabus and materials design (Nikolaeva, 2019), language learning, teaching, and assessment (J. Liu & Yang, 2021). The application value of LPS is empowered when it is designed, whether it is learning-oriented, assessor-oriented, or constructor-oriented (Jones, 2014). It is significant to study the application of LPS in the educational context. Evidence collected within the first few years after its publication is of prime importance and a key decision-making basis for the further implementation of the scale (Zhu, 2016) because it can tap the value of LPS in-depth and discover its positive role and impact on different stakeholders.

Exploring the application of LPS is a validation of its usefulness. The feedback from the stakeholders can back up the scale's validity and provide evidence for its revisions or adaptation (Y. Jin & Jie, 2020). For example, the Common European Framework of Reference (CEFR) is the most widely used LPS for planning and evaluating curricula, certifications, examinations, and textbooks (Byram, 2020). After its publication, studies on its impacts amounted (Brunfaut & Harding, 2020; Byram, 2020; Green, 2018; Sahib & Stapa, 2022), revealing its strengths and weakness in education. The feedback contributed to the release of The Companion Volume (CV) in 2020 (Council of Europe, 2020). Exploring the application of LPS can also offer a better understanding of its usefulness in teaching, learning, testing, and curriculum

[1]School of Educational Studies, University Sains Malaysia, Penang, Malaysia
[2]Department of Economics, Birmingham Business School, College of Social Sciences, University of Birmingham, UK

Corresponding Author:
Samah Ali Mohsen Mofreh, School of Educational Studies, University Sains Malaysia, Block 45, USM Pulau Pinang 11800, Malaysia.
Email: samahmofreh@usm.my

design, providing enlightenment for teachers, students, testers, and policymakers, as the exploration can demonstrate how challenges are addressed and offer new perspectives on moving the field further (Harsch, 2014).

There are some influential language proficiency scales widely used for many years in international language education, such as ILR (*Interagency Language Roundtable*), ACTFL (*American Council on the Teaching of Foreign Languages*), CLB (*Canadian language benchmarks*), and CEFR (*Common European Framework of Reference*) (Zhou & Liu, 2021). CEFR is the most influential officially published scale worldwide and is widely used for curricula planning and evaluating, assessment, textbooks development, teaching, and learning.

China's Standards of English Ability (CSE) in 2018 is a new scale developed to coordinate teaching, learning, and assessment (J. Liu, 2017). It is the first officially published scale in China, which owns the world's largest population of English learners (Bolton & Bacon-Shone, 2020; H. Liu, 2016). After publication, CSE has been widely applied from primary schools to colleges (M. Liu & Huang, 2019; M. Liu & Liu, 2022; J. Liu & Yang, 2021; Peng & Liu, 2021; Xiong & Liu, 2020).

Considering their influence and number of users, this study selects CEFR and CSE as representatives of LPS to explore their application and impact on education. Research exploring the use of CEFR and CSE has included inclusive stakeholders and approaches. However, up to now, no systematic literature review has been conducted to collect evidence of their usefulness from different stakeholders. To bridge this gap, this study aims to tap the value of LPS in the educational domain by systematically reviewing related papers and providing a solid foundation for future academic research in education.

Framework of Systematic Review

It is challenging to learn how LPS is used in the educational domain because there is a lack of theoretical frameworks and practical approaches. The general method is adopted from language testing because they share much in common. Bailey (1996) proposed a basic model that identified participants, processes, and products (3Ps) which may influence or be influenced by washback in language testing. Based on this 3Ps theory, Y. Jin and Jie (2020) constructed a model to study the application and impact of the CSE speaking scale, as shown in Figure 1.

This model illustrates seven types of stakeholders and the impact of LPS from educational and social domains. The solid arrows (numbers 1, 2) represent the impact of scales on stakeholders and their education practice; the dotted ones (marked 3, 4) indicate the washback of applied research to the scale. While X and Y refer to impacts other than the education domain. According to this model, the impact of LPS in education can be explored from six different stakeholders, as shown in Figure 1: stakeholders 1 to 6 from top to bottom are in the education domain, and the seventh stakeholder is concerned with how societies select talent by using scales.

Although Jin's framework is designed for speaking scale, it contains the core elements of the mechanism of how the impacts are generated: how the stakeholders use LPS in teaching, learning, testing, and selecting talents, and the effects of the implementation. Hence, this study adopted this framework, and the research route is outlined in Table 1.

Since this study focused on implementing LPS in the educational domain, it will explore the application of LPS by stakeholders in education and their effects. The social effect was excluded. The questions guiding this study were:

1. How is LPS used in the educational domain?
2. What is the effect of using LPS?

Method

Resources and Database

To answer these questions, this study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses protocol (PRISMA), and three databases were selected: China Knowledge Network (CNKI) core journals, WOS (core collection), and Scopus. Since CSE is a scale applied in China, literature on this range would likely be better studied and indexed by this CNKI database. And to ensure the quality of articles, only core collections were selected. WoS, the majority of which is a core collection(Carloni et al., 2018), is one of the two important and most comprehensive sources of publications and impact indicators worldwide (Pranckutė, 2021). The other is Scopus. WoS CC and Scopus are trusted publisher-independent global high-citation databases (Baas et al., 2020; Birkle et al., 2020) containing many peer-reviewed, high-caliber academic journals published worldwide. These databases could provide useful tools for systematic literature review.

PRISMA

PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses Statements) is the most commonly used reporting guideline for systematic reviews (Page et al., 2021). The methods and results are reported in sufficient detail to enable users to evaluate the

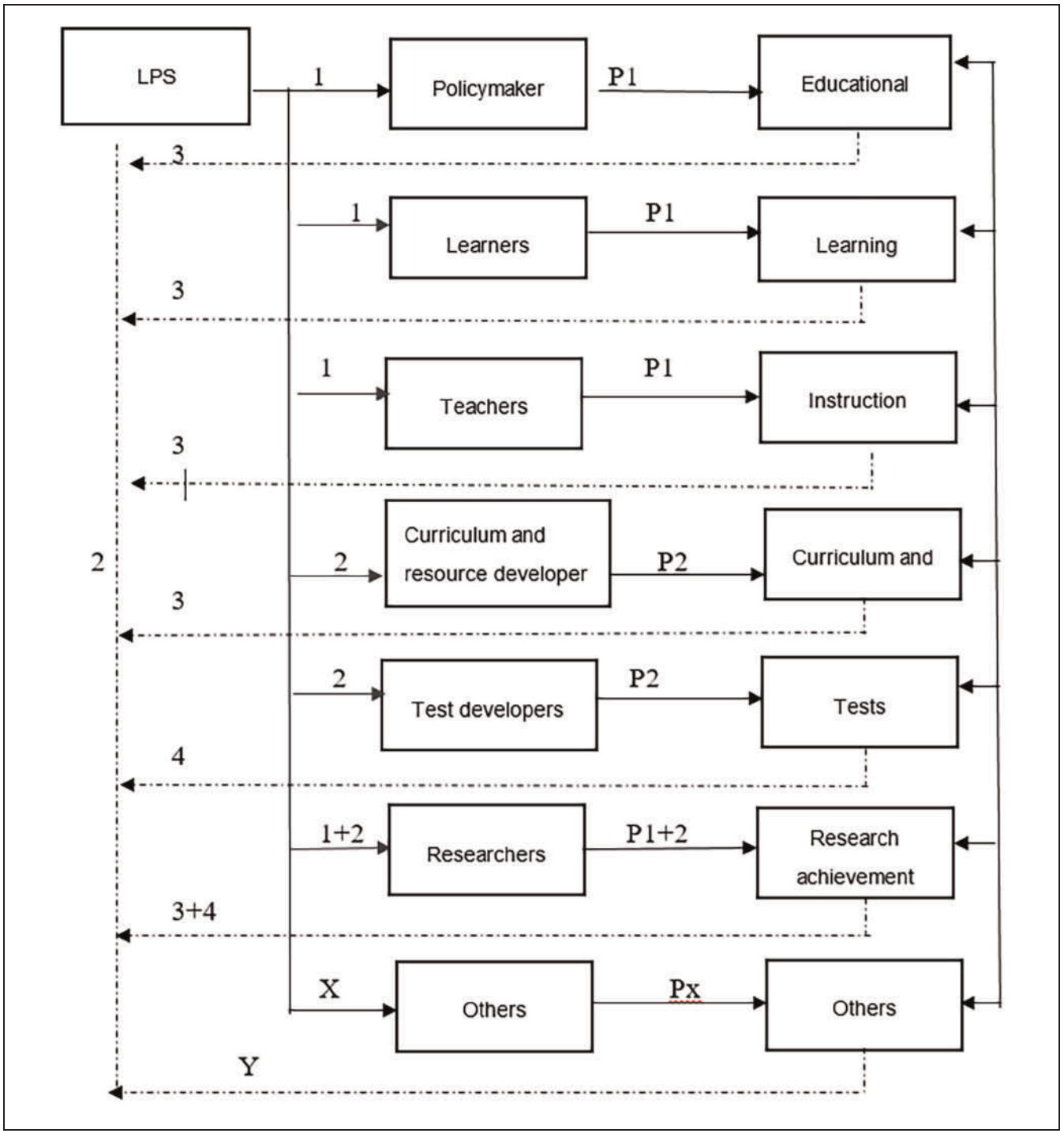


Figure 1. Framework of t exploring the application of LPS.
Note. This figure is adopted from Y. Jin and Jie (2020). Copyright 2020. Reprinted with permission.

applicability and credibility of the review findings. Besides, the PRISMA statement can make systematic review reporting more transparent, comprehensive, and accurate. Hence, it enables a thorough search for information and scientific techniques relevant to the use of LPS in education. The retrieval process is illustrated in Figure 2.

Systematic Review Process

Identification. This systematic review was initiated in 2022 and mainly involved choosing keywords for information search. In the CNKI database, "CSE," "China's Standard of English Proficiency Scales," and "application" were used as keywords. In WoSCC and SCOPUS,

Table 1. Method of Exploring the Impact of LPS.

Perspective	Stakeholders	Guidance of research
Application of LPS in Education	Policymakers	How does the use of LPS affect educational policies?
	Learners	How does the use of LPS affect students learning.
	Teachers	How does the use of LPS affect instruction.
	Curriculum designers and resource developers	How does the use of LPS affect curriculum design and resource development.
	Tester-developers	How does the use of LPS affect test development.
	Researchers	How do researchers improve or operationalize LPS.

Note. This table is adapted from Y. Jin and Jie (2020). Copyright 2020. Adapted with permission.

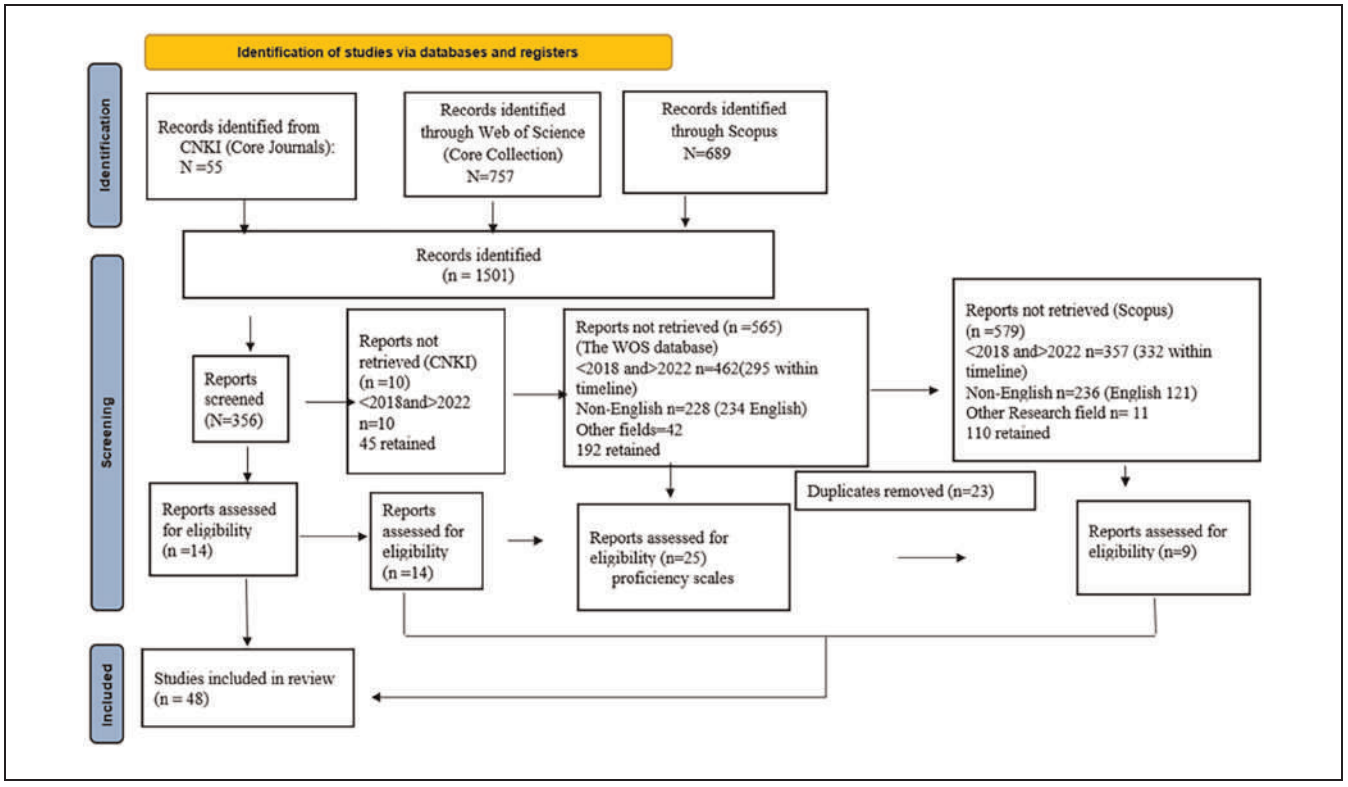


Figure 2. An overview of the search protocol based on the PRISMA statement.
Note. This figure is adapted from Page et al. (2021). CC BY 4.0

"Common European Framework of Reference for Languages," "application," or "CEFR," or "China's Standard of English Proficiency Scales," were used as keywords. For this review focused on the application of language proficiency scales in the education context, "validation" was excluded (See Table 2). As a result, 1,501 papers have been detected searching any studies in which the keywords were stated. Fifty-five papers were retrieved from CNKI, 689 from Scopus, and 757 from WOS, respectively, as seen in Figure 2.

Screening (Inclusion and Exclusion Criteria)

Following the inclusion and exclusion criteria in Table 3, the second stage of the systematic literature review involved screening. The first criterion was time, which was controlled in 5 years (from 2018 to April 2022) for the following considerations: limiting the research within 5 years can ensure the freshness of the literature; besides, CSE was released in 2018, and the application came after its publication; additionally, the finalized 'CEFR Companion Volume with New Descriptors' was also published this year (Eaquals, 2018).

The second criterion is language. Only English and Chinese publications in Scopus, WoSCC, and CNKI databases were included to overcome the distortion of meaning caused by the translation. Third, this research was refined in psychology, education, and linguistics. Other research fields were excluded, given the relevance of the articles published. As a result, 45 articles remained

Table 2. Keywords and Information Search Strategy.

Database	Keywords
CNKI (PKU core journals and CSSCI journals)	"Language proficiency scale" OR "CSE," OR "China's Standard of English Proficiency Scales," "application"
Web of science (core collection)	TS = (common European framework of reference for languages)) AND (application) OR TS = (China's Standard of English Proficiency Scales)) NOT TS = (validation))
Scopus	TITLE-ABS-KEY Common European Framework of Reference for Languages OR China's Standard of English Proficiency Scales AND NOT validation

Table 3. Inclusion and Exclusion Criteria.

Criterion	Included	Excluded
Timeline	2018–2022	<2018 and >2022
Language	English and Chinese	Other languages
Research field	Psychology, social science, education and educational research, linguistics	Other research fields

from CNKI core journal database, 110 from Scopus, and 192 from WOS, as seen in Figure 2.

Eligibility. Eligibility refers to the authors' manual inclusion or exclusion of literature considering criteria in line with the research question and the study objectives. Among these 356 items, 23 papers from Scopus duplicated with WoSCC were deleted. Therefore, 333 papers were retained for manual appraisal. In this process, the authors reviewed the abstracts and full text to confirm their relevance: all the research should focus on applying LPS in education. Studies focusing on the development, validation, perception, and semantic analysis of CEFR and CSE were excluded. Reviews were also removed. Finally, 48 papers (14 articles from CNKI core journals, 9 from Scopus, and 25 from WOS) were retained for review. Sixteen of them focused on the application of CSE, and 32 on CEFR.

Categorizations. The 48 articles that remained were categorized by the authors following the framework stated in Figure 1 and Table 1. By qualitative context analysis of the abstracts and full texts, these articles were categorized into six groups according to stakeholders: policymakers, teachers, learners, curriculum and resource developers, testers, and researchers.

Synthesis and Findings

This section aims to unveil how different stakeholders apply LPS. After synthesis, results indicated that the application for CEFR covers five groups of stakeholders, except students; CSE focused on four groups, excluding policymakers and curriculum and resource developers. Among the reviewed articles, 8 papers concentrate on the policymakers (CEFR only);12 on teachers (8 for CEFR, 4 for CSE); 17 on testers (11 on CEFR, 3 on CSE, and 3 on them both); 3 on curriculum and resource developer (CEFR), 3 on students (CSE), 5 on researchers (2 on CEFR and CSE respectively, 1 article covers both), as seen in Figure 3. The effects of their application are also described in the following paragraphs.

Policymakers

CEFR is claimed to provide a common basis for elaborating language syllabuses, curriculum guidelines, and teacher development. The studies in Table 4 showed the adaptation of CEFR in different countries. Nishimura-Sahi (2020) analyzed the educational trends and domestic needs for practical communicative proficiency in English to increase Japan's economic competitiveness on the global stage. The CEFR-Japan was developed and successfully implemented by assembling various actors-government officers, researchers, commercial actors, administrators, and teachers. The author suggested that CEFR should be borrowed selectively to serve as a viable solution to further long-term educational and political agendas. To ensure its viability, all actors—different stakeholders and publishing houses and materials (such as guidelines and books) should be brought together.

Savski (2019) justified how to use CEFR in Thai and Malaysia. The communicative orientation was unsuccessful in these two countries, and post-communicative philosophy should be advocated. He proposed that content and activities should be developed for learners to reflect

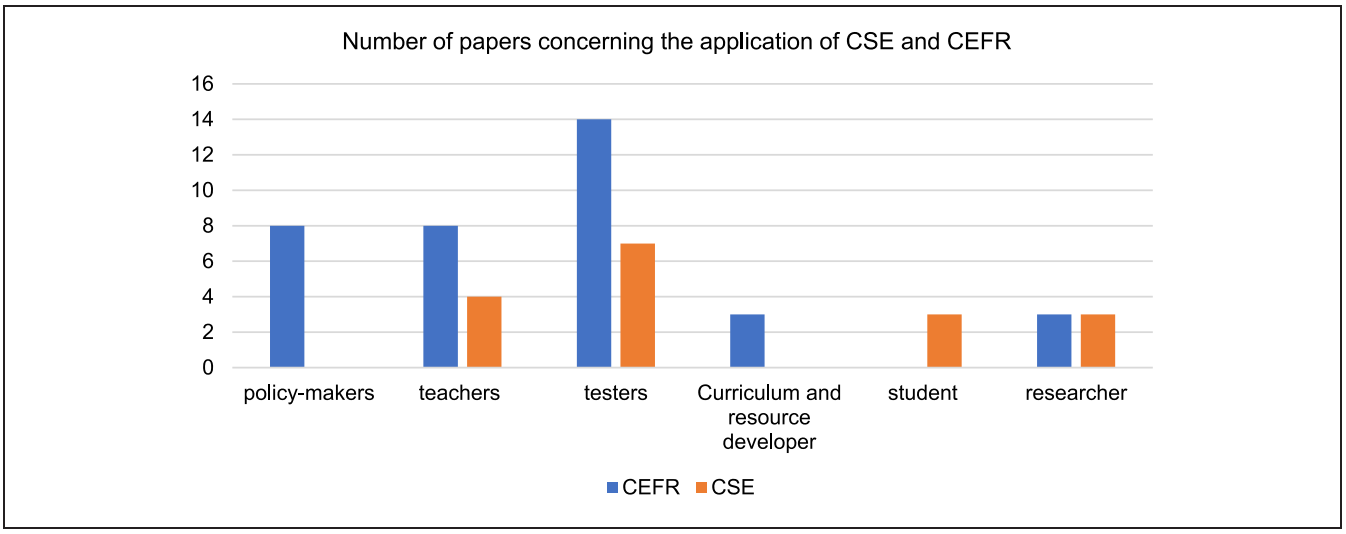


Figure 3. Number of grouped articles.

Table 4. LPS Were Used by Policymakers.

Authors	Focus	Methods	Results
Nishimura-Sahi (2022)	Analyze the "context-specific reasons" for CEFR borrowing in the Japanese context	Qualitative content analysis of the policy documents	The CEFR was borrowed selectively from different stakeholders. The CEFR served as the framework of the new curriculum for the course of study; besides, the CEFR reference levels were adapted to reform university entrance examinations.
Brunfaut and Harding (2020)	How Luxembourg's educational contexts may influence standard-setting practices using the CEFR	Thematic analysis	Four key sources of influence on the adoption of the CEFR: Luxembourg's distinct language ecology, streamed schooling, national curriculum, and ongoing exam reform project.
Savski (2019)	How the CEFR was used in Thai and Malaysian	Literature method	Implementing communicative language teaching (CLT) in Thailand and Malaysia has been unsuccessful; three other alternatives are outlined.
Nguyen and Hamid (2021)	Explore what factors induced the employment of the CEFR in Vietnam	Document analysis	The following conditions induced the employment of CEFR: English language policy changes, the need for economic and political innovations, the initiatives to reform higher education, and administrators' tendency to solve domestic issues by looking outward.
Piccardo et al. (2019)	Successful strategies for introducing CEFR in Canada and Switzerland	Mixed method: qualitative interview and quantitative	Teacher education and CEFR-based examinations are essential to present the CEFR project
Franz and Teo (2018)	Teacher's perception of the introduction of CEFR by moe of Thailand	Grounded Theory Methodology, qualitative analysis	CEFR was introduced as an assessment tool for teachers, where 94% failed to reach the targeted level of B2. Moreover, it was not applied to classroom teaching or learners' assessment.
Aziz et al. (2018)	Problems of implementing CEFR in pre-primary and secondary schools in Malaysia	Qualitative	Teacher training needs improvement. All stakeholders must be adequately synchronized, aware of their responsibilities, and updated on the most recent information. Superficial training should be complemented by more support from the government.
Deygers et al. (2018)	The impact of CEFR on European university entrance policies, tests, and testers	Qualitative	The B2 level is the most adopted, and CEFR levels are frequently abused for marketing purposes or to restrict university entrance.

on their identities as individuals and members of society. The content-based instruction (CBI) approach matched closely with the action-based concept of CEFR, thus having great potential as an alternative to CLT. Besides, policymakers should also consider how to empower students with CEFR criteria.

Nguyen and Hamid (2021) explained the historical and social context of adopting CEFR in Vietnam. They claimed that the following factors contributed to accelerating the adoption of CEFR in local milieus: English language policy changes, the need for economic and political innovations, the initiatives to reform higher education, and administrators' tendency to solve domestic issues by looking outward. Nguyen and Hamid's study demonstrated how the CEFR unfolded on the ground and interacted with the local educational context. It also highlighted the importance of global standards attached by educational actors at different levels. Piccardo et al. (2019) probed into the successful strategies of introducing CEFR in Canada and Switzerland. Results demonstrated that teacher education and CEFR-based examinations were the most important practices.

However, the nexus to CEFR is not always successful and enjoyable. Luxembourg is a critical case presenting the conflicts between international language proficiency standards and local realities (Brunfaut & Harding, 2020). The distinct language ecology, streamed schooling, national curriculum, and ongoing exam reform project limited the setting practice of using the CEFR. A dogmatic approach to CEFR as a common currency cost high in this country. Hence, a better way of theorizing should be proposed to incorporate local knowledge into the standard-setting process without compromising procedural validity when international standards collide with local educational cultures.

Another example is the introduction of CEFR in Thailand (Franz & Teo, 2018).

CEFR failed in its postulated aims of teaching in basic education and teachers' linguistic and instructional skills. Most instructors felt that CEFR was introduced as a tool for evaluating their proficiency scales, not in classroom teaching and assessment. Additionally, they claimed this tool was suitable for Europeans, not for them, as they failed to meet the targeted B2 levels, causing them to lose face.

Aziz et al. (2018) revealed problems with implementing CEFR in pre-primary and secondary schools in Malaysia, where teacher training still needed improvement. They stated that all stakeholders must be adequately coordinated, aware of their roles, and informed of recent developments. Superficial training should be complemented by more support from the government. Deygers et al. (2018) explored the impact of CEFR on European university admission exams. Their findings indicated that B2 is the most adopted level of university entrance. However, the CEFR levels are frequently abused for marketing purposes or to restrict university entrance.

Teachers

Teachers are the primary users of LPS, as seen in Table 5. Generally, they use LPS in assessment and teaching.

For Assessment. Mazlaveckienė (2018) used the CEFR grammar scales to assess Lithuanian English Philology students. Results indicated that these students had a limited repertoire of grammatical structures ranging from level B1 to B2. It shed light on important trends in developing English Philology students' foreign language competency in Lithuania. Zhao and Zhao (2023) explained how teachers and students in China co-constructed writing assessment criteria based on CEFR. The findings supported the efficacy and significance of developing these criteria for improving learners' cognitive and metacognitive knowledge of writing and assessing. They highlighted the importance of learners' competence in developing assessment criteria and implementing a future-drive self-assessment using the CEFR or LPS in local settings.

Shi and Zheng (2021) developed an intelligent diagnostic learning APP based on CSE, in which sports majors practise English adaptively. *T*-tests and questionnaires revealed their effectiveness in motivating and improving students' learning outcomes. He et al. (2021) used CSE-based Cognitive diagnosis models (CDM) to assess the writing abilities of Chinese undergraduates. The linear logistic model analysis demonstrated that diagnostic results could distinguish masters from non-master and facilitate learning by increasing students' competency through feedback and remedial activities. As the authors stated that using CSE for diagnostic purposes could provide methodological support for using a CDM-based approach in diagnostic assessment; it could also provide diagnostic feedback for L2 learners to improve learning.

For Teaching. The above articles concern how LPS were used in assessment, while the following articles focused on teaching practice. Zhong (2019) practised CSE in listening and speaking course in a vocational college by constructing a model combining self-assessment, peer assessment, AI assessment, and teachers' assessment. This model successfully enhanced students' sense of learning responsibility and produced customized learning objectives and strategies. It stressed the importance of applying CSE as guidance in teaching planning and

Table 5. Application of LPS by Teachers.

Authors	Focus	Method	Results
Mazlaveckienė (2018)	Assessing English grammar proficiency in terms of CEFR scales in a university in Lithuania	Qualitative	Lithuanian English Philology students often have a limited repertoire of grammatical structures ranging from level B1 to B2.
Zhao and Zhao (2023)	Teachers and learners co-constructed writing criteria based on CEFR to improve learning	Quantitative	The collaborative effort increased the viability and application of the ELP descriptors, and developed students' cognitive and metacognitive knowledge, and their skills on setting up assessment criteria and evaluating their performance against the criteria.
Shi and Zheng (2021)	Apply CSE-based intelligent autonomous diagnostic APP in English teaching in China	Quantitative	The teaching mode assisted by an adaptive learning system is conducive to implementing formative assessment, the effect of the mode is remarkable, and students have high satisfaction with the teaching mode.
He et al. (2021)	Cognitive diagnosis models (CDM) based on CSE to assess the writing abilities of Chinese undergraduates	Quantitative	Diagnostic results could distinguish masters from non-masters. Students in high proficiency group were higher than ow proficiency students for all attributes.
Zhong (2019)	Apply CSE in English listening and speaking course in China	Qualitative	Integrating CSE in teaching can enhance vocational college students' sense of learning responsibility and produce customized learning objectives and strategies.
Xiong and Liu (2020)	Use CSE in teaching adult English in open learning for ESP	Qualitative	The course tailed for ESP adult learners based on CSE proved effective in enhancing students' interests and learning outcomes.
Rehner et al. (2021)	How CEFR training impacts teachers' French Instruction	Quantitative: retrospective reports	Teachers shifted their planning priority and time, classroom delivery, and assessment practices after the CEFR-related training.
Choong et al. (2021)	Assessment of Grade 5 and 6 pupils before and after the introduction of CEFR amid COVID-19	Qualitative	Before the introduction of the CEFR, not all teachers conducted speaking assessments. However, their teaching and assessment changed as the CEFR emphasized the need for teaching and conducting speaking examinations.
Poonpon et al. (2022)	Develop a model named TIGA based on CEFR and Thailand's basic education and core curriculum for low English proficiency students in rural secondary schools	Quantitative	The results revealed a significant difference in the experimental and control groups' English abilities. The research revealed that the teaching strategy might encourage and engage low-ability students in improving their English proficiency.
Juan Muñoz Andrade	Use CEFR to facilitate learning in universities in Seville	Report	CEFR was used as a methodological and evaluative tool to chart students' progress and give feedback. Students' language proficiency was greatly improved, and they were more confident in speaking.
Sidhu et al. (2018)	The use of CEFR-aligned school-based assessment (SBA) in the Malaysian primary ESL classroom	Mixed method	SBA implementation was far from formative assessment; teachers held positive attitudes toward SBA but had limited comprehension of the CEFR-aligned ESL curriculum. They offer little or no feedback on tasks. Students were discouraged from reflecting on their work, and no self- and peer assessment was found.
Yüce and Mirici (2022)	Implementation of CEFR self-assessment in EFL classes in secondary education in Turkey	Qualitative method	The checklist of self-assessment based on CEFR was provided at the end of each unit in the textbooks; however, they were in low compatibility with CEFR and were not implemented by teachers.

Table 6. Application of LPS by Learners.

Study	Focus	Method	Results
Zhang and Wang (2022)	Scaffold CSE in college English writing and the effect of its application	Mixed method: correlational analysis of students' self-assessment and quantitative study of students' report	Students' self-assessment and writing abilities improved significantly; their learning confidence also improved.
Li (2022)	Use CSE-based peer assessment and task value in writing	Quantitative method: ANOVA	Assessment for learning based on CSE and task value significantly improved students' writing ability and enforced self-regulated learning.
He and Zhang (2021)	Incorporate CSE in self-diagnostic assessment, goal setting, and remedial instruction and learning	Quasi-experiment and qualitative (students report)	Students' listening skills were significantly improved, and they held positive attitudes about this approach to learning, especially the function of CSE in setting SMART goals

Instruction. Xiong and Liu (2020) emulated the reform of ESP teaching for open universities in China based on CSE. The adapted teaching content in CSE, with references to students' work backgrounds and assessment criteria based on CSE descriptors, made this course suitable for learners in distance education. And this method proved effective in motivating their interest and enhancing the learning outcome. Rehner et al. (2021) showed how K-12 teachers' planning, classroom delivery, and assessment practices change after CEFR-related professional learning. Teachers prioritized speaking and listening with less time allotted to writing and reading after learning CEFR; they also shifted their focus away from previous attention to language structure and error correction toward real-life situations; besides, they focused more on students' sociolinguistic and pragmatic competencies. As for assessment, teachers prioritised functional competence and pragmatic and sociolinguistic appropriateness, contrary to the initial focus on grammatical accuracy and orthographic control. These shifts in teachers' planning, classroom delivery, and assessment practices after CEFR-related training signaled an apparent change of their grammar-based model to an action-oriented approach in which language learning took place in genuine communication in an authentic everyday situation.

Likewise, the study of Choong et al. (2021) illustrated how CEFR affects primary school English teachers' behaviors in Japan. Before introducing the CEFR, not all teachers conducted speaking assessments and teaching. Their concepts of Instruction and evaluation changed with the incorporation of CEFR in the elementary English curricula. Thus, speaking was highlighted in teaching and assessment.

In Thailand, Poonpon et al. (2022) reported a model named TIGA based on CEFR and core curriculum for low English proficiency students in rural secondary schools. Results from their quasi-experiment found a significant improvement in the experimental group. This model emphasized the importance of teaching strategy in engaging low-proficient students, especially the authenticity of learning tasks.

Infante Mora et al. (2019) reported how the CEFR was used as a methodological and evaluative tool to chart students' progress and give feedback at a university from Seville. Students' language proficiency was greatly improved, and they were more confident in speaking. This report proved that feedback based on standards is crucial in learning, and teachers' role as facilitators should be amplified.

Like the nexus of CEFR with local context policies, not all cases are enjoyable. Despite the introduction of CEFR in education policy, some ESL teachers in Malaysia (Sidhu et al., 2018) and Turkey (Yüce & Mirici, 2022) had limited comprehension of CEFR-aligned curriculum and could not use CEFR properly. Moreover, students were discouraged from reflecting on their work based on the tasks. More work should be done to overcome ESL teachers' constraints and help them bridge their knowledge between policy and practice.

Learners

Table 6 reveals how students used CSE. Zhang and Wang (2022) explored the scaffolding role of the CSE writing scale in college students. Results showed that with the intervention of CSE, students' assessing ability, writing skills, and learning confidence were significantly improved. Li (2022) examined the effect of CSE-based peer assessment and task value on Chinese undergraduates' self-regulated learning (SRL). Results indicated that learners' SRL was significantly improved. Another study from He and Zhang (2021) incorporated the CSE in diagnostic assessment, goals setting, and remedial instruction to facilitate learning. The quasi-experiment indicated significant improvement in listening skills. Students' self-

Table 7. How Curriculum and Resource Developer Use LPS.

Author	Focus	Methods	Results
Mohamed (2021)	Compiling a list of salient features for curriculum development that would be a basis for designing a framework for a CEFR-aligned Arabic curriculum in UK universities	Inductive research approach using qualitative, interpretive methods.	The study described the context and technique for developing a CEFR-aligned Arabic curriculum framework using a collection of curriculum salient features from the CEFR.
Kalnberzina (2018)	Aligning intercultural components in the English curricula for secondary and tertiary education in Latvia with the CEFR	Documentary analysis.	Despite the differences in terminology, context, and level of impact, these documents were generally compatible.
Little (2018)	How the CEFR was applied in the design of a curriculum framework in Irish primary schools	Report	Borrowing part of scales and descriptors from CEFR to develop the Irish curriculum.

report demonstrated that they held positive attitudes about this approach to learning, especially the function of CSE in setting SMART (specific, measurable, attainable, realistic, and time-bounded) goals.

Course Designers and Resource Developers

Table 7 shows how curriculum and resource developers use LPS. The study by Mohamed (2021) offers a practical model for constructing a CEFR-aligned curriculum. First, it should be essentially action-oriented and concentrate on supporting students in putting their competence into practise. Second, branching each CEFR level into two sub-levels (e.g., A2 into A2.1 and A2.2) would be handy; courses designed to help learners track and monitor their progress could improve their sense of achievement and motivate them. Third, introducing different themes and integrating similar functions that produce equivalent results can be more successful. Forth, a grammar syllabus should be practical and accessible for classroom learning. Mohamed provided an example of options and modifications that teachers may need to consider in implementing CEFR in their contexts. Kalnberzina (2018) compared the intercultural component in secondary and tertiary education curricula. It revealed the compatibility of these documents despite their differences in terminology, context, and level of impact; Little (2018) explored how the CEFR was adopted in designing a curriculum framework for Irish immigrant primary schools. Part of the CEFR scales and descriptors were tailored to the Irish context. And the mediation skills in the CEFR can supplement the deficiency of analytical thinking and problem-solving abilities in secondary education.

Testers-Developers

Tester developers generally use LPS to align tests and different frameworks as a criterion for rating.

Alignment. Table 8 indicates that tester developers mainly use LPS to align tests like IETSL and TOEFL in the international arena, large-scale tests in specific contexts, and in-house tests. Also, alignments between LPS were conducted.

Alignment With International Tests. Fleckenstein et al. (2020) aligned the writing rubric of TOEFL with CEFR in a standard-setting methodology in Germany and Switzerland. Results indicated that the TOEFL test could be meaningfully expressed within the framework of the CEFR. However, the study by Green (2018) showed that IETSL, TOEFL, CAE, and PET-A test agencies made little use of CEFR categories to explain test content and arrived at conflicting conclusions about the test scores and CEFR levels. Among these tests, PTE-A was the only one that defined "at" a level in terms of success likelihood in relation to "Can Do" descriptors for users; others' band boundaries did not correlate directly to CEFR levels. He highlighted the importance of content and quality of assessment procedures. Hidri (2021) aligned the CEFR with International English Language Competency Assessment (IELCA) in listening, reading, speaking, and writing. He also demonstrated that the alignment of five major stages (familiarization, specification, standardization training, and benchmarking standard setting, and validation) could provide abundant evidence of dependable results and made the skills and items in test more specific to reflect the CEFR descriptors.

Hidri prioritized using CEFR to map tests by addressing different mapping stages. It could help teachers effectively use the CEFR descriptors to align IELCA tests and empower them to implement curriculum activities in class transparently and coherently.

Alignment With In-House Tests. Wang (2020) showed how the CSE levels 4 to 7 aligned with SJTU-EPT

Table 8. Application of LPS by Tester-Developers.

Authors	Focus	Methods	Results
Green (2018)	The general alignment of CEFR with IETSL, TOEFL, CAE, PET-A.	Qualitative analysis of IELTS, TOEFL, IBT PTE-A CAE documents	Testing agencies seldom used CEFR categories to interpret test content; they depicted the relationships between their tests and the CEFR in different terms and reached conflicting conclusions about the correlation between test scores and CEFR levels.
Fleckenstein et al. (2020)	Alignment of CEFR with TOEFL rubrics in upper secondary education in Germany and Switzerland	The standard-setting methodology was used to establish the linkages.	The TOEFL test results can be meaningfully expressed within the framework of the CEFR, which underlies educational standards in both countries.
Hidri (2021)	Alignment of CEFR with International English Language Competency Assessment (IELCA) in listening, reading, speaking, and writing	Familiarization, specification, standardization training, and benchmarking standard setting, and validation	The five linking stages explained that the IELCA suite examinations' four levels, B1, B2, C1, and C2, onto the CEFR, providing fair judgments and informed decisions about this mapping task's practical consequences and validity arguments.
Peng (2021)	Alignment of CSE with CEFR writing scales	Rasch model analysis	CSE levels 1 and 2 correlate primarily to CEFR levels A1 and A2, 3 to A2, 4 and 5 to B1, 6 to B2, 7 to C1, 8 to C1 and C2, and levels 9 to C2 and above.
Peng and Liu (2021)	Alignment of CSE with CEFR listening scales	Rasch model analysis	CSE level 1 matches mainly to the CEFR A1 level, level 2 to A2, level 3 to A2 and B1, level 4 to B1, level 5 to B1 and B2, level 6 to B2, level 7 to C1, and level 9 to C2.
Peng et al. (2022)	Aligning CSE with CEFR	Rasch model analysis	CSE level 1 corresponds primarily to the CEFR level below A1, level 2 to A1, level 3 to A2, level 4 and level 5 to B1, level 6 to B2, level 7 to B2 and C1, level 8 to C1 and C2, and level 9 to C2.
Wang (2020)	Alignment of CSE with in-house English proficiency tests in reading, listening, and writing skills	Correlation analysis of students' self-assessment and teachers' assessment based on the CSE description	The seven reported levels of the SJTU-EPT can be linked to the CSE levels four to eight.
Min and Jiang (2020)	Alignment of the listening subtest of an in-house English test and CSE	In the standard setting, the Modified Angoff Method, Contrasting Groups Method, and Multi-Facet Rasch Analysis	The listening subtest of the in-house tests aligns with level 5 of CSE; the two standard-setting approaches produce congruent results.
Harsch and Kanistra (2020)	Align the Integrated Skills of English (ISE) suite in Trinity College London to the CEFR	item-descriptor-matching (IDM) method, Cronbach alpha, and Multi-faceted Rasch Modeling analysis	High agreement for task judgments, acceptable reliabilities and consistency for examinee-centered ratings, and varying levels of agreement for descriptor choices.
Baharum et al. (2021)	Alignment of CEFR with English Language Competence Score Average (ELCSA) in a university in Malaysia	Quantitative method: correlative analysis	The results showed a significant positive correlation that varied in strength, with writing the strongest correlation.
Sufi et al. (2021)	Mapping English writing skills tests in English Proficiency Tests (EPT) with CEFR in International Islamic University of Malaysia (IIUM)	Quantitative method: correlation analysis	EPT writing bands correlated positively to scales of the CEFR.
Shak and Read (2021)	Aligning English for Occupational Purposes (EOP) meeting assessment in Malaysia to the CEFR level.	Qualitative: NVIVO-coding	A revised set of language assessment criteria was introduced; results showed how the scoring criteria could be aligned with the CEFR scale through a systematic comparison of language functions generated in the meeting task.

(continued)

Table 8. (continued)

Authors	Focus	Methods	Results
Shermis (2018)	Provide a crosstalk between CEFR and autorotated writing evaluation (AWE) system	Regression model approach	The CEFR traits and their machine scoring system aligned in fluency, coherence, and accuracy. While traits of range and interaction were less well aligned.
Al Habbash et al. (2021)	Alignment of CEFR standards with Emirates Standardized Test (EmSAT) in the United Arab Emirates and IELTS	Quantitative and qualitative	EmSAT and IELTS are not rigorously aligned with the CEFR standards. Furthermore, the EmSAT mostly aligned with the lower levels of the CEFR, whereas the IELTS mostly aligned with the higher levels of the CEFR.
Jie (2019)	Alignment of CET-CET 4 with speaking scales of CSE	Multi-facet Rasch model	Through the test task analysis, the panelists could select relevant descriptors. Following thorough training, they demonstrated good consistency and accuracy at each level of standard setting.
Holzknecht et al. (2018)	Raters from Finland and Austria use the CEFR-based rating scale to measure students' writing abilities	Rasch model analysis	Although the Austrian raters were marginally more lenient than the Finnish raters, the range of disagreement was tiny. Thus, these two teams mostly agreed upon the participants' CEFR levels.
Silveira and Martins (2020)	How experienced raters use CEFR holistic and analytic scales to assess oral proficiency progress in English as a second language	Quantitative method, correlation analysis of analytical and holistic tests	Significant positive correlations existed between holistic and analytic assessment, and raters rated consistently with analytical scales. A better speaker performance across time is detected in fluency, while pronunciation and grammar improvement was insignificant.

(Shanghai Jiaotong University English Proficiency Tests). The scores based on the descriptors of the CSE from the teachers and students claimed that the B and B + levels in SJTU-EPT corresponded with level 6 in CSE, and the C and C + aimed at level 5. Most of the A-level students corresponded with level 7, and D aimed at level 4 in CSE. Min and Jiang (2020) aligned an in-house English test subtest at Zhejiang University in China with CSE with Modified Angoff Method and the Contrasting Groups Method. The congruent results indicated that the tests align with level 5 of CSE.

Harsch and Kanistra (2020) aligned the ISE suite in Britain, Trinity College London, to the CEFR with an item-descriptor-matching (IDM) method and a complementary benchmarking approach. Results showed high agreement for task judgments, acceptable reliabilities and consistency for examinee-centered ratings, and varying levels of agreement for descriptor choices. In Malaysia, scholars aligned CEFR with Competence Score Average (ELCSA), English Proficiency Tests (EPT) in universities, and English for Occupational Purposes (EOP) meeting assessments (Baharum et al., 2021; Shak & Read, 2021; Sufi et al., 2021). Results showed an overall positive correlation between CEFR and these in-house tests, proving the acceptability and credibility of these tests. Shermis (2018) established a crosstalk between CEFR and one automated writing evaluation system (AWE) in America. The CEFR traits and their machine scoring system were clearly aligned in fluency, followed by coherence, and accuracy. While traits of range and interaction were less well aligned. The author highlighted that operationalizing "good writing" and advocating the traits in CEFR could help machine scoring accurately.

Alignment With Large-Scale National Tests. In China, CET-4 is a test with the largest population. Most college students attend this test. Jie (2019) aligned CET-4 with speaking scales of CSE, demonstrating good consistency and accuracy at each standard-setting level. Another alignment made by Al Habbash et al. (2021) revealed that the large-scale tests–Emirates Standardized Test (EmSAT). In United Arab Emirates was not rigorously aligned with the CEFR standards.

Alignment Between CEFR and CSE. A series of alignments between these scales were made by the same author, proving that CSE level 1 corresponds primarily to the CEFR level below A1, level 2 to A1, level 3 to A2, level 4 and level 5 to B1, level 6 to B2, level 7 to B2 and C1, level 8 to C1 and C2, and level 9 to C2 (Peng, 2021; Peng & Liu, 2021; Peng et al., 2022). As the author stated: these studies contributed to the internationalization of the Chinese assessment system and provided references

Table 9. Application of LPS by Researchers.

Authors	Focus	Methods	Results
Ma and Chen (2021)	Construct a pragmatic competence assessment model and standard based on the CSE	Mathematical modeling, Delphi method, quantitative method	A college English pragmatic competence assessment model was developed and tested in practice. This model proved to be complementary to current assessment forms.
Y. Liu et al. (2021)	Developing a framework of English proficiency standards at the open university of China based on CSE and CEFR	Delphi method, Rasch analysis	A five-level scale was developed, and the policies and procedures for accrediting adult learners' English learning outcomes were formulated.
Yang et al. (2021)	Developing Typical Interpreting Activity Scales (TIAS) based on CSE	Quantitative analysis	Interpreting activities were categorized into eight groups; the "Can do" descriptors present interpreting performance at different topics, interpreting models, and skills in ascending levels.
Yannakoudakis et al. (2018)	Developing an Automated Writing Placement System for ESL Learners based on CEFR full scales	Quantitative method	The system is developed to assess learners' proficiency levels on the CEFR scale. This model was incorporated into Cambridge English Write & Improve system to offer diagnostic feedback for learners.
Schmidt et al. (2019)	Developing a guidebook and tools to implement the CEFR for course design	Report	The CEFR-related resources were thematically rearranged based on the following function: curriculum and course design, assessment, and learner autonomy as a guidebook.

for the alignment of language standards and language education in China.

Rating. Another type of application concerns how raters use CEFR scales to assess learners' proficiency levels. Silveira and Martins (2020) explored how experienced raters used CEFR holistic and four analytic scales (vocabulary, grammar, fluency, and pronunciation) to measure students' oral proficiency progress. The results demonstrated that the five scales were positively correlated, and raters were consistent in using these criteria. However, subscales detected significant progress in fluency only. Grammar and pronunciation improvement was hardly seen. The authors claimed that even in a communicative teaching context, grammar and pronunciation should be emphasized to coordinate the development of subcomponents of oral proficiency. The research from HoIzknecht et al. (2018) also showed that raters in Finland and Austria might differ in leniency; they agreed to a large extent on the CEFR levels of the participants if they are trained and experienced in the CEFR-based rubric.

Researchers

Researcher concerns with the development and improvement of scales (Y. Jin & Jie, 2020). As seen in Table 9, a Study by Y. Liu et al.(2021) demonstrated how the *English proficiency standard for adult learners* in open universities (OUSE) is developed based on CES and CEFR and how this scale is used to certify their learning outcomes. The authors reported the specific steps of designing and applying this scale, showing that the OUSE provided a benchmark for assessing adult learners in open universities. Ma and Chen (2021) constructed a pragmatic competence assessment model and standard based on the CSE, complementing the current assessment. Yang et al. (2021) developed Typical Interpreting Activity Scales (TIAS) based on CSE.

Yannakoudakis et al. (2018) developed an Automated Writing Placement System for ESL Learners based on CEFR full scales. This model was incorporated into Cambridge English Write & Improve system to offer diagnostic feedback for learners, facilitating self-assessment, tutoring, and improvement in learning. Schmidt et al. (2019) developed a guidebook and tools to implement the CEFR for course design, simplifying the implementation of CEFR, and fostering the novice use of it.

These studies operationalized standards in LPS and guided language teaching and assessment from theoretical to practical levels.

Discussion

This paper reviewed and summarized the application of LPS in education, focusing on the CEFR and CSE. After eligibility, 48 articles met our inclusion criteria; 16 studies were about CSE, and 32 were on the CEFR. These studies showed how LPS was applied by policymakers, curriculum designers, researchers, test developers, teachers, and students.

Policymakers

Mohamed (2021) stated that the CEFR had been used more frequently at the macro level, that is, for policy-making. The results of this review, as seen in Figure 3, confirmed this finding. The CEFR is adopted globally for its open-mindedness and vagueness, which scholars often criticize. However, this quality made the CEFR flexible for local contexts (Savski, 2019). The reviewed papers demonstrated how the CEFR was adopted selectively in different backgrounds. The interaction of CEFR with the local setting can be positively and negatively influenced by local realities. Social needs necessitated the adoption of global standards, and the government or policymaker selectively borrowed criteria from the CEFR (Nguyen & Hamid, 2021; Nishimura-Sahi, 2022).

However, the introduction of foreign LPS should be highly cautious. They should be tailored to the specific context. Otherwise, conflicts may arise. The study of Brunfaut and Harding (2020) served as an extreme case of the tension between the CEFR and the local realities. The standard setting process of Luxembourg Épreuve Commune for English was highly influenced by the local realities, such as multilingual learning ecology, streamed schooling system, national curriculum, and exam reform. The lessons of contextualization of CEFR shed some light on policy-making in other countries. When global criteria and local context collide, a better way of theorizing how to integrate local knowledge and international standard should be proposed without compromising standard-setting procedures (Brunfaut & Harding, 2020).

Besides, the introduction of LPS in policy should be accompanied by updating teaching philosophy. As Savski (2019) claimed, the incompetence of the old version of communicative teaching cannot match the concepts advocated by the CEFR in Thailand and Malaysia. New agendas for policymakers to adapt to the practicing of CEFR were proposed: alter teaching philosophy to a post-communicative concept, devitalize the teaching process, and empower students with criteria.

Furthermore, introducing foreign standards should also prioritize the training of teachers. The unsuccessful adoption and implementation of CEFR in Thailand, Malaysia, and European universities (Aziz et al., 2018; Deygers et al., 2018; Franz & Teo, 2018) prioritized the significance of teacher training and synchronization of all stakeholders. Otherwise, the CEFR would be misused or abused.

Teachers

Teachers adopted LPS as an assessment for learning. LPS are benchmarks of assessments to evaluate students' language proficiency more accurately and pinpoint essential trends in developing students' language competency (Mazlaveckienė, 2018). They could also be references to diagnostic assessments, from which the feedback could inform students what remedial works should be performed to improve their language proficiency. This type of application highlighted the merits of advanced psychometric techniques to provide diagnostic feedback for L2 learners (He et al., 2021; Shi & Zheng, 2021).

The application of LPS could also lead to new approaches to teaching. Zhong (2019) constructed a model combing self-assessment, peer assessment, AI assessment, and teachers' assessment to facilitate learning. Zhong's study exemplified the function of CSE as goal-setting in the teaching process. Shi and Zheng (2021) designed an adapted smart testing system based on CSE to meet the objective and subjective needs for the practice of a "learning, teaching, and testing" integrated teaching model. This model highlighted how diagnostic assessment could be used to facilitate learning. Xiong and Liu (2020) explored using rubrics and contents adapted from CSE in assessing and improving English proficiency in open universities. The study of Zhao and Zhao (2023) demonstrated that the collaborative process improved the feasibility and usefulness of the CEFR descriptors and developed students' cognitive and metacognitive knowledge and skills for setting up assessment criteria. These studies proved the effectiveness of LPS as a tool in improving learning outcomes and activating learning interests.

Teachers practised LPS creatively to facilitate students' learning, and in turn, their concepts and teaching philosophies were affected by LPS. As CEFR promoted an action-oriented approach in language teaching, it changed teachers' concepts from grammar-oriented learning to authentic-task-based learning, input-focused to out-put-focused teaching (Poonpon et al., 2022; Rehner et al., 2021). It also gave birth to the student-centered concept, emphasising the collaborative learning process of teachers and students (Zhao & Zhao, 2023).

Students

Generally, students used LPS as self- or peer-assessment tools and the goal-setting benchmark. Self-assessment based on LPS plays a scaffolding role in learning, and improved students' language proficiency, assessment literacy, and confidence (Zhang & Wang, 2022). When integrated into learning, LPS improves the efficacy of self- and peer-evaluation, enhances students' self-regulation, and boosts the value of assessment for learning (Li, 2022). Moreover, LPS could guide goal-setting and offer students a benchmark to analyze and reflect on their learning progress critically and actively and remedy their learning (He & Zhang, 2021). In the long term, LPS

could be an essential learning guidance and assessment tool to cultivate independent lifelong learners.

Curriculum and Resource Developers

Adopting descriptors from LPS scales to local context and adding them to an existing curriculum document should be encouraged. Compiling descriptors from the CEFR, Mohamed (2021) developed a generic concise Arabic curriculum of salient features and aligned it with CEFR. This curriculum conformed to the CEFR's philosophy, that is, transparent, coherent, and flexible. Kalnberzina (2018) added some CEFR cultural components and standards into the secondary school curriculum to develop learners' intercultural decision-making abilities. These two studies set good examples of complementing existing curricula by aligning CEFR standards. Both studies highlighted the importance of adapting to local context and content alignment. However, the number of retained papers indicates that studies on this type need to be fleshed out by further explorations.

Testers-Developers

Test scores alone are insufficient to support administrators or teachers in making meaningful decisions, nor can the test takers be well-informed of their proficiency levels. Aligning the scores with the "can do" descriptors in LPS is significant in teaching, learning, and assessment (Wang, 2020).

The alignment of tests to LPS levels could provide learners with a valuable sense of their current language ability and a more detailed and comprehensive view of students' linguistic profiles (Sufi et al., 2021). With the help of a competent teacher—this alignment might form the basis for further study or remedial learning (Fleckenstein et al., 2020). It could also reflect how to use unified standards to interpret students' authentic language proficiency, provide feedback for teaching, and back up the learning plan and objectives (Wang, 2020).

Alignment with tests also accentuates the importance of using LPS to ensure the credibility and reliability of testing results (Hidri, 2021), especially for in-house test, which differs from school to school. The same score in different schools does not claim the same level of proficiency. Aligning in-house tests with LPS can bridge this gap by measuring students' ability more accurately with a common benchmark and promote the accreditation of academic scores in different schools (Min & Jiang, 2020); it could also provide evidence for further improvement of language tests (Baharum et al., 2021).

Studies on alignment between different LPS are also crucial. As alignments could promote the recognition of standards from other areas and cultures, highlighting the significance of language scales in use and construction, helping the mutual-recognition of different standards (Peng & Liu, 2021; Peng et al., 2022).

Furthermore, LPS is a benchmark or a tool for language assessment. Slightly modified LPS can be a valid and reliable tool for assessing language proficiency. However, training LPS standards and the raters' language proficiency levels should be reinforced (HoIzknecht et al., 2018; Silveira & Martins, 2020). Under these premises, using the LPS descriptors for rating can ideally lead to the same results across different contexts and achieve high congruence for all scales.

Researchers

LPS can provide a theoretical framework, methodology, and source of descriptors for developing new tools. The CEFR is well-known for benchmarking the design of a contextualized language assessment framework or systems. CSE also proved to be a practical reference for constructing new scales and assessment models (Y. Liu et al., 2021; Ma & Chen, 2021; Yang et al.,2021). And the new tools help to operationalize the practice of LPS (Yang et al., 2021), simplify their implementation, and usher in novice use of LPS (Schmidt et al., 2019). Finally, the new tools could contribute to students' reflection on their errors, tracking progress, and facilitating learning (Yannakoudakis et al., 2018).

Conclusion

This research adopted the PRISMA systematic review method for an in-depth review of 48 articles regarding how different stakeholders use LPS in the education domain and the effect of applying LPS. The findings revealed that policymakers used LPS selectively to backbone their education decisions and reforms (Nishimura-Sahi, 2020). The adoption of CEFR should consider specific social contexts, such as language ecology, streamed schooling, the national curriculum, ongoing exam reform, and concrete economic and political situation (Nguyen & Hamid, 2021). Teachers used LPS to assess students' language proficiency to gain an overall profile of students' competence and diagnose their problems, achieving goals of assessment for learning (He et al., 2022; Shi & Zheng, 2021). New teaching models based on LPS were also adopted to improve learning confidence and outcomes (Shi & Zheng, 2021; Xiong & Liu, 2020; Zhong, 2019). Students used LPS as a benchmark to provide feedback from their self-assessments, track their progress, and set learning goals (He & Zhang, 2021; Li, 2022; Zhang & Wang, 2022). The intervention of LPS in learning could help cultivate self-regulated learners and enhance students' motivation and learning

outcomes. Curriculum designers tailored descriptors and scales from CEFR to develop new curricula, align the existing ones, and make revisions where necessary (Kalnberzina, 2018; Mohamed, 2021). Test developers aligned LPS with tests to make the results more reliable and credible, measure students' ability more accurately, and provide evidence for improving language tests (Hidri, 2021; Sufi et al., 2021). Raters used LPS to assess more accurately (HoIzknecht et al., 2018; Silveira & Martins, 2020). With the methodology and theoretical framework, LPS could offer references for developing new rubrics and frameworks and assessing models for researchers (Ma & Chen,2021; Yang et al.,2021).

Despite the potential in education, to ensure the successful implementation of LPS, trainers, teachers, testers, and raters alike, all the stakeholders should update their knowledge and information, improve their language skills, and take their responsibilities; the government should make more effort (Aziz et al., 2018). The application of LPS in education is a systematic project supported by wide-ranging stakeholders and updated concepts and behaviors; it calls for capturing political interests, developing contextualized relevant resources, and providing sufficient teacher training (Nishimura-Sahi, 2020).

This review offers a panoramic view of how LPS are used in education, providing evidence for the application validity LPS. However, there are some limitations in this study, for it mainly concerns CEFR and CSE. Besides, it explored the application and impact of LPS by adopting a model based on language testing for the paucity of theory in LPS application. Studies theorizing the validation of the aftereffects of other LPS should be encouraged.

Acknowledgments

I would first like to thank my supervisor Dr Samah Ali Mohsen Mofreh, whose expertise was invaluable in formulating the research questions and methodology. Her insightful suggestion pushed me to sharpen my thinking and brought my work to a higher level. I also would like to thank Dr Zhong for helping appraise the articles and the anonymous peer reviewers whose feedback offered insightful points in revising this article.

Author contributions

AZ writing the original draft and revising the article. SM conceived the initial idea and supervised the literature and methodology. Moreover, she revised the first complete draft. SS helped in conceiving the idea and provided guidance and supervision on the theory and empirical side. He offered some suggestions on the final revision.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Aihua Zhu https://orcid.org/0000-0002-0034-1222
Samah Ali Mohsen Mofreh https://orcid.org/0000-0002-7980-2940
Sultan Salem https://orcid.org/0000-0002-2593-6046

References

Al Habbash, M., Alsheikh, N., Liu, X., Al Mohammedi, N., & Al Othali, S. (2021). A UAE standardized test and IELTS Vis-À-Vis international English standards. *International Journal of Instruction*, *14*(4), 373–390. https://doi.org/10.29333/iji.2021.14422a

Aziz, A. H. A. A., Rashid, R. A., & Zainudin, W. Z. W. (2018). The enactment of the Malaysian common European framework of reference (CEFR): National master trainer's reflection. *Indonesian Journal of Applied Linguistics*, *8*(2), 409–417.

Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. *Quantitative Science Studies*, *1*(1), 377–386. https://doi.org/10.1162/qss_a_00019

Baharum, N. N., Ismail, L., Nordin, N., & Razali, A. B. (2021). Aligning a university English language proficiency measurement tool with the CEFR: A case in Malaysia. *Pertanika Journal of Social Sciences & Humanities*, *29*, 157–178. https://doi.org/10.47836/pjssh.29.s3.09

Bailey, K. M. (1996). Working for washback: A review of the washback concept in language testing. *Language Testing*, *13*(3), 257–279. https://doi.org/10.1177/026553229601300303

Birkle, C., Pendlebury, D. A., Schnell, J., & Adams, J. (2020). Web of Science as a data source for research on scientific and scholarly activity. *Quantitative Science Studies*, *1*(1), 363–376. https://doi.org/10.1162/qss_a_00018

Bolton, K., & Bacon-Shone, J. (2020). The statistics of English across Asia. In K. Bolton, W. Botha, & A. Kirkpatrick (Eds.), *The handbook of Asian Englishes* (pp. 49–80). Wiley. https://onlinelibrary.wiley.com/doi/10.1002/9781118791882.ch3

Brunfaut, T., & Harding, L. (2020). International language proficiency standards in the local context: Interpreting the CEFR in standard setting for exam reform in Luxembourg. *Assessment in Education-Principles Policy & Practice*, *27*(2), 215–231. https://doi.org/10.1080/0969594X.2019.1700213

Byram, M. (2022). Politics, origins and futures of the CEFR. *The Language Learning Journal*, *50*(5), 586–599. https://doi.org/10.1080/09571736.2020.1845392

Carloni, M., Tsenkulovsky, T., Carloni, M., & Mangan, R. (2018). *Web of science core collection descriptive document*. https://www.google.com.hk/search?client=aff-cs-360se&ie=UTF-8&q=Web + of + Science + Core + Collection + Descriptive + Document

Choong, E. E. C., Manoharan, P., & Rethinasamy, S. (2021). Speaking assessments by Japanese English teachers pre and post implementation of CEFR in the midst of a global

pandemic. *Pertanika Journal of Social Sciences & Humanities*, *29*, 335–349. https://doi.org/10.47836/pjssh.29.s3.17

Council of Europe. (Ed.). (2020). *Common European framework of reference for languages: Learning, teaching, assessment ; companion volume*. Council of Europe Publishing.

Deygers, B., Zeidler, B., Vilcu, D., & Carlsen, C. H. (2018). One framework to unite them all? Use of the CEFR in European university entrance policies. *Language Assessment Quarterly*, *15*(1), 3–15. https://doi.org/10.1080/15434303.2016.1261350

Eaquals. (2018, June 20). *The CEFR companion volume launch conference*. https://www.eaquals.org/2018/06/20/the-cefr-companion-volume-launch-conference-may-2018-tim-goodier/

Fleckenstein, J., Keller, S., Kruger, M., Tannenbaum, R., & Koller, O. (2020). Linking TOEFL iBT (R) writing rubrics to CEFR levels: Cut scores and validity evidence from a standard setting study. *Assessing Writing*, *43*, 33–47. https://doi.org/10.1016/j.asw.2019.100420

Franz, J., & Teo, A. (2018). 'A2 is Normal' – Thai secondary school English teachers' encounters with the CEFR. *RELC Journal*, *49*(3), 322–338. https://doi.org/10.1177/0033688217738816

Green, A. (2018). Linking tests of English for academic purposes to the CEFR: The score user's perspective. *Language Assessment Quarterly*, *15*(1), 59–74. https://doi.org/10.1080/15434303.2017.1350685

Harsch, C. (2014). General language proficiency revisited: Current and future issues. *Language Assessment Quarterly*, *11*(2), 152–169. https://doi.org/10.1080/15434303.2014.902059

Harsch, C., & Kanistra, V. P. (2020). Using an innovative standard-setting approach to align integrated and independent writing tasks to the CEFR. *Language Assessment Quarterly*, *17*(3), 262–281. https://doi.org/10.1080/15434303.2020.1754828

He, L., Jiang, Z., & Min, S. (2021). Diagnosing writing ability using China's Standards of English Language Ability: Application of cognitive diagnosis models. *Assessing Writing*, *50*, 1–14. https://doi.org/10.1016/j.asw.2021.100565

He, L., & Zhang, J. (2021). The application of the China's Standards of English Language Ability in remedial instruction and learning. *Foreign Language Testing and Teaching*, *3*, 1–11.

Hidri, S. (2021). Linking the International English Language Competency Assessment suite of examinations to the Common European Framework of Reference. *Language Testing in Asia*, *11*(1), 1–24. https://doi.org/10.1186/s40468-021-00123-8

HoIzknecht, F., Huhta, A., & Lamprianou, I. (2018). Comparing the outcomes of two different approaches to CEFR-based rating of students' writing performances across two European countries. *Assessing Writing*, *37*, 57–67. https://doi.org/10.1016/j.asw.2018.03.009

Infante Mora, E., Muñoz Andrade, J., Greenwood, D., Feldman, R., Ivanchikova, M., Cívico Gallardo, J., & García Saez, P. (2019). Part 2: The development of an active pedagogy. *Learning and Teaching*, *12*(3), 18–45. https://doi.org/10.3167/latiss.2019.120303

Jie, W. (2019). Relating English speaking tests to China's Standards of English: CET-SET 4 as a case study. *Foreign Language World*, *190*(1), 71–80.

Jin, Y., & Jie, W. (2020). The application and impact of language proficiency scales: A case of the CSE Speaking Scale. *Foreign Language World*, *198*(3), 52–60.

Jones, N. (2014). *Multilingual frameworks* (*Vol. 40*). Cambridge University Press.

Kalnberzina, V. (2018). Intercultural/pluricultural communication construct and its levels. *Baltic Journal of English Language Literature and Culture*, *8*, 40–55. https://doi.org/10.22364/BJELLC.08.2018.03

Li, J. (2022). Effects of CSE-based assessment for learning and task value on self-regulated learning in college English writing classroom. *Journal of China Examinations*, *12*, 19–26. https://doi.org/10.19360/j.cnki.11-3303/g4.2022.12.003

Little, D. (2018). Functional approaches to syllabus design: From the threshold level to the common european framework of reference for languages. In A. Faravani (Ed.), *Issues in syllabus design* (pp. 99–109). Brill.

Liu, H. (2016). *Language policy and practice in a Chinese junior high school from global Englishes perspective* [PhD Thesis]. University of Southampton.

Liu, J. (2017). China's standards of English and English learning. *Foreign Languages in China*, *14*(6), 4–11. https://doi.org/10.13564/j.cnki.issn.1672-9382.2017.06.002

Liu, J., & Yang, M. (2021). Exploring the applicability of the CSE in language testing and assessment. *Foreign Language Testing and Teaching*, *42*(2), 1–11.

Liu, M., & Huang, M. (2019). Primary English story teaching based on China's standards of English language ability. *Basic Foreign Language Education*, *21*(3), 91–98 + 111.

Liu, M., & Liu, Y. (2022). Building college English translation teaching model based on China's standards of English language ability. *Journal of Tonghua Normal University*, *43*(03), 139–144. https://doi.org/10.13877/j.cnki.cn22-1284.2022.03.022

Liu, Y., Li, X., & Liu, Z. (2021). Development of foreign language proficiency standards for adult learners and certification of their learning outcomes based on China's standards of English proficiency—A project on the development of English proficiency standards at the open university of China. *Technology Enhanced Foreign Language Education*, *199*(03), 63–69 + 10.

Ma, Y., & Chen, X. (2021). An empirical study on construction of the assessment model of college English pragmatic competence based on China's standards of English. *Journal of Northeast Normal University*, *310*(02), 151–164. https://doi.org/10.16164/j.cnki.22-1062/c.2021.02.020

Mazlaveckienė, G. (2018). Assessment of university students' English grammar proficiency in terms of CEFR criterial achievement levels: The case of Lithuanian university of educational sciences. *Theoria et Historia Scientiarum*, *15*, 35–50. https://doi.org/10.12775/ths.2018.003

Min, S., & Jiang, Z. (2020). Aligning an in-house listening test to China's Standards of English Language Ability. *Foreign Language Education*, *41*(4), 47–51. https://doi.org/10.16362/j.cnki.cn61-1023/h.2020.04.009

Mohamed, S. (2021). The development of an Arabic curriculum framework based on a compilation of salient features from CEFR level descriptors. *The Language Learning Journal*, *51*(1), 1–15. https://doi.org/10.1080/09571736.2021.1923781

Nguyen, V. H., & Hamid, M. O. (2021). The CEFR as a national language policy in Vietnam: Insights from a sociogenetic analysis. *Journal of Multilingual and Multicultural Development*, *42*(7), 650–662. https://doi.org/10.1080/01434632.2020.1715416

Nikolaeva, S. (2019). The common European framework of reference for languages: Past, present and future. *Advanced Education*, *6*(12), 12–20. https://doi.org/10.20535/2410-8286.154993

Nishimura-Sahi, O. (2020). Policy borrowing of the common European framework of reference for languages (CEFR) in Japan: An analysis of the interplay between global education trends and national policymaking. *Asia Pacific Journal of Education*, *42*(3), 574–587.

Nishimura-Sahi, O. (2022). Assembling educational standards: Following the actors of the CEFR-J project. *Globalisation, Societies and Education*, *21*(3), 392–404. https://doi.org/10.1080/14767724.2022.2037071

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, *372*, n71. https://doi.org/10.1136/bmj.n71

Peng, C. (2021). Aligning the CSE with the CEFR: Level alignment in writing ability. *Foreign Language World*, *5*, 84–93.

Peng, C., & Liu, J. (2021). The listening skill level alignment of the CSE with the CEFRA. *Foreign Language education*, *42*(5), 43–50. https://doi.org/10.16362/j.cnki.cn61-1023/h.2021.05.008

Peng, C., Liu, J., & Cai, H. (2022). Aligning China's standards of English language ability with the common European framework of reference for languages. *The Asia-Pacific Education Researcher*, *31*(6), 667–677. https://doi.org/10.1007/s40299-021-00617-2

Piccardo, E., North, B., & Maldina, E. (2019). Innovation and reform in course planning, teaching, and assessment: The CEFR in Canada and Switzerland, a comparative study. *Canadian Journal of Applied Linguistics*, *22*(1), 103–128. https://doi.org/10.7202/1060908ar

Poonpon, K., Satthamnuwong, B., & Sameephet, B. (2022). The effectiveness of task-based and genre-based integrated learning on English language proficiency of Thai rural secondary school students. *Theory and Practice in Language Studies*, *12*(9), 1736–1747. https://doi.org/10.17507/tpls.1209.05

Pranckutė, R. (2021). Web of science (WoS) and scopus: The titans of bibliographic information in today's academic world. *Publications*, *9*(1), 1–59. https://doi.org/10.3390/publications9010012

Rehner, K., Popovich, A., & Lasan, I. (2021). How the CEFR is impacting French-as-a-second-language in Ontario, Canada: Teachers' self-reported instructional practices and students' proficiency exam results. *Languages*, *6*(1), 15. https://doi.org/10.3390/languages6010015

Sahib, F. H., & Stapa, M. (2022). Global trends of the common European framework of reference: A bibliometric analysis. *Review of Education*, *10*(1), 1–26. https://doi.org/10.1002/rev3.3331

Savski, K. (2019). Putting the plurilingual/pluricultural back into CEFR: Reflecting on policy reform in Thailand and Malaysia. *Journal of Asia Tefl*, *16*(2), 644–652. https://doi.org/10.18823/asiatefl.2019.16.2.13.644

Schmidt, M. G., Nagai, N., Naganuma, N., & Birch, G. (2019). Teacher development: Resources and devices to promote reflective attitudes toward their profession. *Language Learning in Higher Education*, *9*(2), 445–457. https://doi.org/10.1515/cercles-2019-0024

Shak, P., & Read, J. (2021). Aligning the language criteria of a group oral test to the CEFR: The case of a formal meeting assessment in an English for occupational purposes classroom. *Pertanika Journal of Social Sciences and Humanities*, *29*(S3), 133–156. https://doi.org/10.47836/pjssh.29.s3.08

Shermis, M. D. (2018). Establishing a crosswalk between the common European framework for languages (CEFR) and writing domains scored by automated essay scoring. *Applied Measurement in Education*, *31*(3), 177–190.

Shi, J., & Zheng, B. (2021). Practice of the intelligent autonomous diagnostic applications based on CSE in English teaching for sports major students. *Journal of Guangzhou Sport University*, *41*(3), 114–119. https://doi.org/10.13830/j.cnki.cn44-1129/g8.2021.03.028

Sidhu, G. K., Kaur, S., & Chi, L. J. (2018). CEFR-aligned school-based assessment in the Malaysian primary ESL classroom. *Indonesian Journal of Applied Linguistics*, *8*(2), 452–463. https://doi.org/10.17509/ijal.v8i2.13311

Silveira, R., & Martins, T. (2020). Assessing second language oral proficiency development with holistic and analytic scales. *Ilha Do Desterro-A Journal of English Language Literatures In English And Cultural Studies*, *73*(3), 227–249. https://doi.org/10.5007/2175-8026.2020v73n3p227

Sufi, M., Abu, K., & Ibrahim, E. H. E. (2021). Mapping IIUM students' English language writing proficiency to CEFR. *Pertanika Journal of Social Sciences & Humanities*, *29*(S3), 85–101. https://doi.org/10.47836/pjssh.29.S3.05

Wang, H. (2020). Aligning school-based English proficiency tests with China's standards of English language ability: A case study. *Foreign Language World*, *200*(5), 72–79.

Xiong, Y., & Liu, Y. (2020). Development of foreign language proficiency standards for adult learners and certification of their learning outcomes based on china's standards of English proficiency—A project on the development of English proficiency standards at the open university of China. *Chinese Vocational and Technical Education*, *199*(8), 81–86.

Yang, Y., Zhang, R., & Mu, L. (2021). Classifying typical interpreting activities based on CSE-interpreting competence scales. *Foreign Language Research*, *220*(3), 96–102. https://doi.org/10.16263/j.cnki.23-1071/h.2021.03.016

Yannakoudakis, H., Andersen, O. E., Geranpayeh, A., Briscoe, T., & Nicholls, D. (2018). Developing an automated writing placement system for ESL learners. *Applied Measurement in*

Education, *31*(3), 251–267. https://doi.org/10.1080/08957347.2018.1464447

Yüce, E., & Mirici, I. H. (2022). Self-assessment in EFL classes of secondary education in Türkiye: The Common European Framework of Reference for Languages (CEFR)-based implementations. *Pegem Journal of Education and Instruction*, *13*(1), 349–359. https://doi.org/10.47750/pegegog.13.01.38

Zhang, W., & Wang, X. (2022). The application of CSE in students' self-assessment in a college English writing course: From the perspective of dynamic assessment theory. *Foreign Languages in China*, *19*(1), 71–78. https://doi.org/10.13564/j.cnki.issn.1672-9382.2022.01.003

Zhao, H., & Zhao, B. (2023). Co-constructing the assessment criteria for EFL writing by instructors and students: A participative approach to constructively aligning the CEFR, curricula, teaching and learning. *Language Teaching Research*, *27*(3), 765–793. https://doi.org/10.1177/1362168820948458

Zhong, L. (2019). The application of CSE in listening and speaking courses in vocational colleges in China. *Chinese Vocational and Technical Education*, *17*, 92–96.

Zhou, Y., & Liu, J. (2021). Developing a scale of language proficiency for specific purposes. *Journal of Foreign Languages*, *44*(4), 33–41.

Zhu, Z. (2016). A validation framework for the national English proficiency scale of China. *China Examinations*, *8*, 3–13. https://doi.org/10.19360/j.cnki.11-3303/g4.2016.08.001

ARTICLES FOR FACULTY MEMBERS

THE COMMON EUROPEAN FRAMEWORK OF REFERENCE DISSEMINATION IN ENGLISH LANGUAGE EDUCATION

Title/Author	**The CEFR-aligned curriculum execution in Malaysia and other countries: a conceptual paper / Bakar, R. A.; A. M., & Bakar, R. A.**
Source	***Malaysian Journal of ELT Research*** **Volume 19 Issue 1 (2022) Pages 1-15** **https://doi.org/10.52696/TGCT6849** **(Database: ResearchGate)**

Article

https://doi.org/10.52696/TGCT6849
Reprints and permission:
The Malaysian English Language Teaching Association
✉ Noor Azli Affendy Lee noor.azli@uitm.edu.my

The CEFR-Aligned Curriculum Execution in Malaysia and Other Countries: A Conceptual Paper

Noor Azli Affendy Lee
Academy of Language Studies,
Universiti Teknologi MARA, Cawangan Pulau Pinang
Malaysia

Aini Akmar Mohd Kassim
Academy of Language Studies,
Universiti Teknologi MARA, Shah Alam
Malaysia

Rofiza Aboo Bakar
Academy of Language Studies,
Universiti Teknologi MARA, Cawangan Pulau Pinang
Malaysia

ABSTRACT

The Common European Framework of Reference for Languages (CEFR) impacts language education, learning and evaluation in the European nations as well as in different nations around the world. The recently presented CEFR-aligned educational plan through the Malaysian English Language Roadmap (2013-2025) could set up a fundamental and reliable arrangement of learning guidance and evaluation in Malaysia. As the CEFR has been broadly embraced by numerous nations before its selection in Malaysia, there is a need to look at the issues faced by different nations to guarantee a superior arrangement of the CEFR in the Malaysian educational program. This paper aims to examine the executions of the CEFR in a few nations to satisfy their respective education policies in order to compare with the development and execution of the CEFR in Malaysian schools and universities. In this investigation, a review of 25 research papers published in journals from the year 2010 to 2019 related to the CEFR transformation and execution issues for English language from different nations all around the world, including Malaysia, was conducted. Utilising Google Scholar, these papers were selected with important keywords such as "CEFR" and the name of the chosen country. In view of the current writing, a few differences just as qualities and constraints of the CEFR-aligned executions were underscored, which propose required data to rethink the execution of the CEFR in the Malaysian education curriculum in order to accomplish the significant goal of refining English instructing, learning and assessment. The

paper ends with proposals on the need to normalise academic practice to improve the CEFR-aligned educational program change endeavours.

KEYWORDS: CEFR, Curriculum, Execution, Review

INTRODUCTION

The Common European Framework of Reference for Languages (CEFR) was set up in 2001 to offer a practicable system that assigns the learning entailments of language students to utilise a language adequately in practice (Council of Europe, 2018). Initially, its expectation was to offer evaluation and showing approaches for all languages in Europe. Notwithstanding, because of its straightforwardness of strategy in various regions, the CEFR has also been endorsed in nations outside Europe. It has been embraced and adjusted to suit different necessities and demands mostly in numerous nations' instructive strategies, homeroom education, language testing and assessment, language educational plan improvement and other significant territories in language education. Various nations are handling the CEFR in their own respective ways. Some decide to directly embrace its structure, particularly when a nation's background has recognizable affiliations and likenesses to local English-talking nations, while others choose to adjust the CEFR to be in tandem with their distinct cultures and local language acquisition approaches. In any case, the parallel differentiation among nations is that the CEFR is chiefly utilised by instructors to check how well their pupils are performing with respect to the CEFR scale levels against international standards.

The vital markers of the success of the educational plan and schedule made from an instructive strategy lie on how instructors utilise assessments. This involves measuring students' achievements using the required benchmark of the learning curriculum or the course outcome at the completion of a fixed timeline (Bharati & Lestari, 2018). Malaysian students have been acquainted with the English language starting from preschool (5 or 6 years old), and the language keeps on being instructed all through their schooling stages from elementary to tertiary level. The English language is a mandatory subject in the Malaysian schooling educational program and it is generally perceived as a significant second language in Malaysia. Despite having been educated under the English curriculum for a long time, a majority of Malaysian ESL pupils have not been able to accomplish a healthy degree of competency in the language (Azman, 2017) remarkably in communication and composition abilities (Hamzah et al., 2018). In the wake of the development of the CEFR in some nations' educational policies, the Ministry of Education of Malaysia has joined the trend to use the CEFR in its essential execution for English language teaching and learning in the country. The establishment of the Malaysian Education Blueprint 2013-2025 that filled in as the all-inclusive strategy that guides forward for the Malaysian schooling framework (Kaur & Shapii, 2018) has made an educational programme reform in the Malaysian English as a Second Language (ESL) syllabus, instruction and evaluation. This adjustment of the educational programme is expected to fill in as a way for a foundational change of English language instruction in Malaysia.

The execution of the CEFR in Malaysia is being done in three methodically essential stages (Ministry of Education Malaysia, 2013). The primary wave in 2013 to 2015 had zeroed in on strengthening the current instruction educational plans by raising the English language capability of teachers (Ministry of Education Malaysia, 2013). Then, the second wave in 2016 to 2020 introduced key movements which incorporate CEFR-aligned educational programmes, instruction and learning as well as evaluation development for teachers (Ministry of Education Malaysia,

2013). This would set and approve the fitting CEFR levels against every instruction level in Malaysian primary and secondary schools as well as tertiary institutions (Ministry of Education Malaysia, 2013). Beginning from 2017, the new CEFR-aligned educational programmes have begun to be carried out in standard one and form one English language syllabi respectively and the execution of the newly introduced syllabi proceeds to the following grade level every year. Simultaneously, ESL instructors actually go for professional development training and workshops to further adapt to the CEFR-align educational programmes (Zuraidah Mohd Don & Mardziah Hayati Abdullah, 2019). Ultimately, the third wave in 2021 onwards will focus on assessing, reviewing and modifying the CEFR-aligned English language curricula (Ministry of Education Malaysia, 2013) for the English Language Standards and Quality Council (ELSQC).

To help in understanding the CEFR Malaysia Roadmap 2025, it is also essential for those involved to be closely monitored and guided both in formulation and direction. It is likewise fundamental for them to be included all the more intently to permit their voices to be heard. It is inadequate to just depend on official proclamations of how assessment should be outlined and comprehended. Hence, literature should also be enhanced with a comprehension of how the CEFR is actually adopted and adapted throughout various nations.

This paper intends to look at the executions of the CEFR in a few nations in order to realise their particular education policies. The findings can then be compared with the progress and execution of the CEFR in Malaysia.

LITERATURE REVIEW

Due to the distressing concern in deciding standardised guidelines for English language teaching and assessment corresponding to worldwide benchmarks, for the last 15 years, the Common European Framework of Reference for Languages (CEFR) has been heartily utilised and embraced by European countries in the area of language assessment and evaluation and has gradually affected the design of educational programmes as a whole (Read, 2019) and on the assessment of language learning results specifically (Holzknecht et al., 2018). The CEFR has been extensively famous past Europe because of its evident thoroughness and experimentally created and approved six-level marks (A1, A2, B1, B2, C1, C2) which impact the items in language learning programmes in various settings for various utilisations and purposes (Idris & Raof, 2017). The clear-cut six-level scale suggests a progressive advancement in language learning from novice, intermediate to higher proficiency level (Read, 2019). There have been various CEFR-related studies conducted by researchers in order to explore the efficiency of the scales in differing national education curricula. A few investigations included a search for insights on the acknowledgment and responses on the utilisation of CEFR. Others were more attracted to investigating educators' understanding of the CEFR and its practical uses in classrooms. There were also studies relating to textbooks, educational plans, and instructional techniques.

The CEFR was first given in two draft reports in 1996. In 2001, it was further revised to be made accessible in French and English languages. Since its distribution, the CEFR has immediately accomplished a powerful capacity in language instruction all through Europe (Read, 2019). It was found that the most consistently used segment of the CEFR was the common reference levels of the 6-level rating scales. Subsequently, these rating scales have ended up being the ‘common currency’ in numerous nations in Europe and have started to get recognised in other countries across the globe. The establishment of the CEFR rating scales is illustrated in Figure 1.

Figure 1

The development process of the CEFR scales (Council of Europe, 2018)

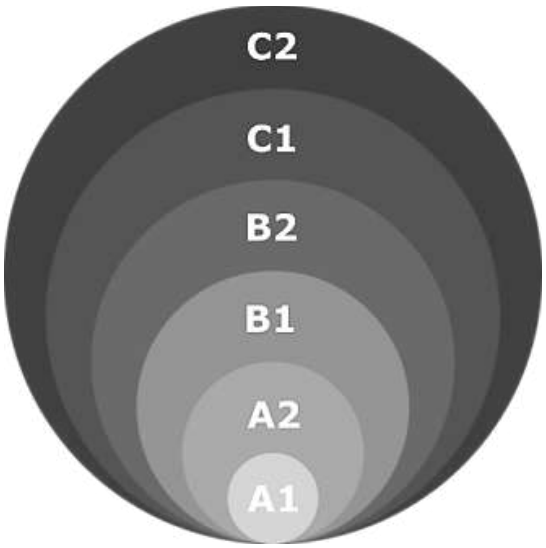
Phase 1 Step 1: Collection of 2000 descriptors from over 30 scales in use around the world. Step 2: Classification of each descriptor according to categories of communicative language ability and writing additional descriptors to fill perceived gaps.
Phase 2 Step 3: Pairs of teachers are given sets of descriptors typed onto confetti like strips of paper and asked to sort them into categories. Step 4: The same pairs are asked to comment on the "usefulness" and "relevance" of each descriptor for their students. Step 5: Teachers are given the same sets of descriptors and asked to separate them into three levels: 'low', 'middle' and 'high', and then divide each of these into two categories to create the familiar six level scale. Step 6: The descriptors most consistently placed in the same level of the scale are used to create overlapping 'questionnaires' of descriptors, with the overlap items operating as anchors.
Phase 3 Step 7: A rating scale is attached to each descriptor on the questionnaire. Step 8: A group of teachers is asked to rate a small number of their learners from their classes on the rating scale for each of the descriptors on the questionnaire. Step 9: This data is used to construct scales of unidimensional items using Rasch analysis, rejecting any items that misfit the Rasch model. Step 10: Items that behave statistically differently across languages or sectors are identified and removed. Step 11: Cut scores are established using difficulty estimates in order to achieve equidistant bands.
Phase 4 Step 12: Conduct the study again using a different group of teachers.

The CEFR addresses an exertion by language teachers and testing experts in Europe to build up a typical system to help relate language courses and evaluations to one another (Cox et al., 2017). The fundamental standards of the CEFR incorporate viewing language learning as a long-lasting experience and recognising the capability of the students' language proficiency in a given coursework. The widespread utilisation of the CEFR has made its rating scales a truly agreeable instrument as it proposes a more exact and steady method of procuring the stage at which the students' proficiency in the language is, rather than using general characterizations of language

learners (refer to Figure 2, Council of Europe, 2018). It is a way to build global arrangement, advance deep rooted learning and boost the quality and practicality of language learning and development in educational institutions. Despite all the materials that have been produced, the Council of Europe has been working in effort to keep on creating parts of the CEFR, especially the illustrative descriptors for L2 and FL proficiency. Subsequently, the importance of studies that examined the CEFR scales for scoring L2 or FL learners' proficiencies need to be further explored to compare the CEFR scales in various locales. However, the CEFR must also be utilised rationally to augment the quality of the assessments of learners on a scale that is more visible than other methods employed to identify their proficiency (Holzknecht et al., 2018). By doing this, it reflects the expanding awareness of the need for an incorporated method to language education across the curriculum.

Figure 2

The six common reference levels (Council of Europe, 2018: 34)



The widespread use of the CEFR has made its proficiency scale a very likeable tool as it proposes a more accurate and consistent way of acquiring the phase at which the students of interest are than using broad classifications of learners such as 'beginners', 'intermediate' or 'advanced' (Holzknecht et al., 2018). Thus, the emphasis of analyses that have exploited CEFR scales for scoring second (L2) or foreign (FL) language learners' proficiencies has not been to explore the comparability of the CEFR scale in diverse settings, but to use it as a rational tool to enhance the quality of the assignments of learners or their performances on a scale that is more apparent than other systems used to define learners' proficiency (Holzknecht et al., 2018).

Much work done by the other institutions and professional bodies since the publication of CEFR has confirmed the validity of the initial research conducted. To build on the widespread adoption and use of the CEFR, the Council of Europe has published an extended version of the illustrative descriptors that complements the original ones in 2018 (Council of Europe, 2018). This extension takes the CEFR descriptors beyond the area of modern language learning to encompass aspects

relevant to language education across the curriculum in the extensive consultation process undertaken in 2016 to 2017. The summary of the major changes is captured in Figure 3.

Figure 3

Summary of the major changes in CEFR Companion Volume with New Descriptors 2018 (Foley, 2019: 32)

Summary of the major modifications/additions in CEFR 2001-2018
1. Developing the illustrative descriptors of second/foreign language proficiency
2. To produce versions of CEFR for young learners (7-10\11-15) and for sign language
3. To develop more detailed coverage in the descriptors for A1 and the C levels
4. Complement the original illustrative scales with descriptors for mediating a text, mediating concepts, and mediating communication
5. The provision of descriptors for plurilingual/pluricultural competence
6. The removal of any reference to 'native speaker' and it is being replaced with speakers of the target language
7. The proficiency level of speakers of the target language is not specified and uses the term 'partial' competence, arguing that language users are fundamentally uneven in different contexts

As elsewhere around the world, English language teaching, learning and assessment are undergoing substantive change towards the establishment of a common framework of English language ability scales. Challenges faced with regards to the implementation of CEFR have been studied extensively to reimagine language pedagogy and to improve the utilisation of CEFR in various domains in different countries. Due to the alarming concern in determining standards for English language instruction in relations to global benchmarks, since its establishment, the CEFR has been vigorously utilised and received to the region of language evaluation and testing in numerous nations and has impacted the plan of language educational programs when all is said in done (Read, 2019) and on the appraisal of language learning results specifically (Holzknecht et al., 2018). There have been numerous CEFR related studies done by scholars in many areas. Some studies involved looking for insights on the recognition and reactions on the use of the CEFR. Others were more attracted to investigating educators' arrangement and there were additionally contemplates identifying with CEFR-adjusted course books, educational programs and instructing strategies.

Holzknecht et al., (2018) in their study in comparing CEFR-based ratings of writing performances between raters of different national and educational contexts stated that writing-related CEFR descriptors for rating purposes may indeed produce equivalent results among raters in different European countries if those raters are trained and greatly knowledgeable in using systematically established CEFR-based rating scales. Hence, it should be said that the straightforward use of CEFR descriptors for rating purposes needs extensive training and experience and cannot be anticipated from classroom teachers. At the same time, the CEFR has come to serve as an administration means for government officers to exert control over language education by stipulating learning outcomes in general terms and a way of outlining minimum levels of language

aptitude in contexts such as higher learning, occupation and migration (Read, 2019).

METHODOLOGY

This study reviews 25 journal articles published between 2010 and 2019 which identify with the execution of the CEFR in selected countries and the issues each nation faced with regard to its implementation. Utilising Google Scholar, these papers were selected with important keywords such as "CEFR" and the name of the chosen country. The countries were chosen as they provide extensive literature with regard to the development and implementation of the CEFR. As the finding for the investigation, the paper will focus on the executions and alignment of the CEFR in European countries Canada, the Netherlands, and Sweden; Asian countries Taiwan, Japan, and China; and Southeast Asian countries Vietnam and Malaysia.

FINDINGS

Although initiated from a project of the Council of Europe, the CEFR soon demonstrated reasonable context-independence and was initiated in countries around the world. In 2006, the Canadian Council of Ministries of Education decided to found an operational team to deliberate the strengths and weaknesses of the CEFR in detail (Mison & Jang, 2011). In 2010, the council publicly proposed that provinces and territories of Canada utilise the CEFR for teaching, learning and assessment purposes. Going along with the government, educators from several other areas in education have adopted the CEFR in Canada. Prominently, the insertion of the CEFR in the Canadian context placed a positive progression in motion at the level of reconceptualisation of tools and frameworks connected to assessment, curriculum and pedagogy (Arnott et al., 2017). Findings by Mison and Jang (2011) suggested that assessment transparency, consistency and plurilingualism in the classroom are noticeable and current concerns of FL and L2 teachers should be reflected in order to boost teacher's support and partaking in the implementation of CEFR in Canadian classrooms.

In the Netherlands, since its introduction, CEFR is gradually recognised and utilised in Dutch secondary education. Findings from Moonen et al., (2013) stated that the use of CEFR is most prevalent in the use of CEFR-related textbooks. However, the fraction of teachers who use CEFR more comprehensively is rather small even though commonly Dutch FL teachers have the essential grasp of CEFR and welcome its function as an instrument to evaluate target language proficiency on a universal level. This is because the Dutch government does not officially impose the usage of CEFR and the schools can choose to implement CEFR however they prefer.

The forthcoming interpretation of language learning following the CEFR has prompted a shift in teacher education for EFL teachers in Sweden and how foreign languages are taught in Swedish schools (Baldwin, 2018). However, the choice to establish learning outcomes linked to the CEFR as a preliminary point for managing teaching and learning was considered outside of the teaching cluster and subsequently, there were diverse viewpoints within the group about employing CEFR and its descriptors. It was stated in Baldwin's (2018) study that the teachers deemed it problematic to utilise the CEFR descriptors when measuring examples of pupil product and pupils would have difficulties comprehending the CEFR descriptors as they were unclear.

As previously stated, the effect of the CEFR has circulated well outside Europe and it is remarkable to compare how various countries in Asia have reacted to it. Many non-speaking English countries such as China, Thailand, Vietnam, Malaysia and many more have embraced English as a language of communication to partake and contend in the globalised economy (Uri & Aziz, 2018b) to the extent of implementing and aligning CEFR into their own national education policies.

In 2005, the Ministry of Education in Taiwan was determined to adopt CEFR as a common benchmark of English language proficiency in the country and all national tests were progressively standardised against the CEFR so that its marks could be interpreted in terms of the levels on the framework (Read, 2019). The Language Training and Testing Centre at National Taiwan University commenced a project to plot the test levels on to the 6 levels of CEFR and were capable of displaying a very satisfactory level of alignment. However, a number of problems arose as the Ministry of Education in Taiwan did not have the technical capability to assess the validity of the assertions attained by test producers that their tests have been aligned with the CEFR and there is a lack of transparency on the grading standards employed by universities to measure their students' attainment in English through their course work.

Since the late 1900s, the Ministry of Education, Culture, Sports, Science and Technology in Japan has tried to foster the English communication proficiencies of Japanese students (Kimura et al., 2017) and to offer consistency and transparency in language learning (Fergus, 2015). An 8-year project called CEFR-J was established by a team of language scholars at the Tokyo University of Foreign Studies and its goal was to adapt the CEFR to the Japanese context using in-depth evaluation of a set of 647 descriptors so that they would better mirror the level of complexity that Japanese learners faced in doing diverse communicative tasks in English and the chances to apply English in the Japanese context (Read, 2019). This project also examined bodies of texts and of Japanese learner language corpus to recognize grammatical and lexical traits and followed the effect of the CEFR-J through inventive usage of "big data analysis" (Read, 2019). However, Japanese teachers of English met with difficulties in the lack of clarity on what English language use should be like in the classroom (Kimura et al., 2017). This concurs with Fennelly's (2016) statement that Japanese practitioners are not prepared, not taught and not eager to accept and use the CEFR curriculum as teachers and students are equally inclined to look into excelling in exams compared to concentrating on communication skills. This leads to Kimura et al., (2017) examining the ideal assessor situations to recommend appropriate standard measurements to utilise in continuing teacher professional development.

Meanwhile, China proposed the development of a Common Chinese Framework of Reference for Languages (CCFR) with a precise emphasis on the teaching of English to offer the chance to scrutinise some essential questions about language education in the country such as what motivates Chinese learners to study foreign languages, at what age level, what target language should be taught and what proficiency levels to aim (Read, 2019). However, the exceedingly segmented Chinese education structure gave rise to lack of transparency and management among stakeholders and many of them are refusing the new CCFR system. Liu and Jia (2017) reported there is a discrepancy between the learner "can do" and what they "do" in testing conditions as they may focus on their own performance rather than focusing on reacting to their peers' actions and feedback dynamically, overlooking the collaborative quality of the assessment. Despite that, the experience acquired from developing and integrating CEFR into Chinese education has given rise to the development of Test for English Majors (TEM) that was designed to assess English language

proficiency of their undergraduate English majors. TEM is likely to aid the implementation of the teaching syllabus and to enhance the quality of language teaching and learning for English majors across China (Zou & Zhang, 2017). In another study by Zheng et al., (2016), they stated there is a possibility of positive effects on teachers' teaching and evaluation habits if teachers are more acquainted with the CEFR scale via appropriate instruction.

Vietnam is the first country in South East Asia to adopt CEFR in 2008 (Uri & Aziz, 2018b). In Vietnam, English was first introduced and taught during the French times but it was not as significant as studying French. Nevertheless, English has developed to be an exceptionally significant foreign language for economic reforms in later years especially during the 90s Asian financial crisis (Uri & Aziz, 2018b). Understanding the significance of English to Vietnam's economic development, English has become a staple and obligatory course to both undergraduate and postgraduate students in Vietnam. Numerous efforts have been made to restructure the foreign (especially English) language teaching system, among which is the adoption of CEFR into the Vietnamese local context of language teaching and learning as a quick-fix solution to reorganise the national foreign language education system (Le, 2018). Six years after the adoption of CEFR, in 2014, a Vietnamese version of CEFR was approved to all levels of education in Vietnam from kindergarten to higher education (Khang, 2018). This reformed framework for foreign language proficiency was established to fit the native contexts and lessen Eurocentric elements of CEFR. This framework is employed to measure the standard and quality of English learning as well as to encourage educational institutions to dynamically improve and execute bilingual programmes. However, the framework caused the opposite intended reaction among the Vietnamese teachers who were not accustomed to its rubrics and were uninterested in using the CEFR descriptors in their classroom activities. Even after nearly 10 years of its first introduction in Vietnam, the adoption of CEFR still meets challenges and difficulties from restricted human resources to complications in teacher professionalism (Le, 2018) such as teachers have not been assessed or been trained thoroughly and systematically on CEFR-based materials and assessment (Khang, 2018). Khang (2018) also found that the Vietnamese government has some official standards for EFL teachers but they usually differ from the descriptors in the CEFR.

English has been utilised and taught in proper educational instructions for years in Malaysia. Despite all the efforts put forth by the government to enhance English proficiency levels of Malaysians, the standard is still inadequate compared to other developed countries and Malaysia has yet to deliver highly-skilled graduates who have solid control of the language (Uri & Aziz, 2018b). Therefore, the CEFR has been systematically adopted by Malaysia, aligning the framework with the syllabus, curricula and assessments in the Malaysian education system. The implementation of CEFR in Malaysia is planned to take place in 3 waves starting from 2013 to 2025 with the first wave in 2013 to 2015 to focus on consolidating the existing education system and curricula which includes tackling minimal English competence among English teachers. The second wave (2016 to 2020) would present a fundamental shift which includes CEFR-aligned curricula, teaching and learning as well as assessment development. After 4 years of implementation, the outcomes of CEFR-aligned English language curricula will be evaluated and revised in the third wave from 2021 to 2025.

The Malaysian cascade training on CEFR was documented by Aziz et al., (2018). The reported CEFR cascade training model can be referred to in Figure 4. The first cohort of teachers who underwent training consisted of those who would be teaching English for primary 1, primary 2,

form 1, and form 2 in 2017. The familiarisation stage lasted from October to November 2016 (exposing participants to language learning pedagogy perspectives in the CEFR and interpreting action-oriented perspectives on curriculum, teaching methodology and assessment, reflection on how CEFR could impact areas of education), learning material evaluation, adaptation and design (understanding principles of materials evaluation, differentiation strategies, adaptation and design, integrated learning skills) was combined with the curriculum induction stage (understanding content and learning standards, scheme of work, lesson outlines and procedures resources including new textbooks and non-textbook materials, differentiation strategies and teachers;' feedback) which was held from July to September 2017. The item writing and formative assessment stage was conducted from January to March 2018. However, at the time of the report, the fourth stage was still ongoing, hence, it was not reported. Aziz et al., (2018) reported the third-tier course suffered greatly in terms of content delivery as there was insufficient training due to time constraint and lack of organisation and funding.

Figure 4

The Malaysian CEFR Cascade Training Model (Aziz et.al. 2018)

Tiers	Trainers	Familiarisation (Stage 1)	Learning Material Evaluation, Adaptation and Design (Stage 2)	Curriculum Induction (Stage 3)
First Tier	Cambridge English Super Trainers (CEST)	5 days 5 to 7 CEST 25 NMT each	5 days 5 to 7 CEST 25 NMT each	5 days 5 to 7 CEST 25 NMT each
Second Tier	National Master Trainers (NMT)	5 days 200 NMT 100 DT each	5 days 200 NMT 25 DT each	5 days 200 NMT 25 DT each
Third Tier	District Trainers (DT)	Not stated 6000 DT Not stated	1 day Concurrent with Stage 3 100 teachers each	2 days Concurrent with Stage 2 100 teachers each
Fourth Tier	ESL teachers			

In its preliminary stage in which policymakers and stakeholders are yet to get used to the framework, Malaysia has chosen to adopt CEFR into its language curriculum development and gradually examine its expansion and adapt to its outcomes. Noteworthy alterations have been put together in lesson plans and the ESL syllabus and the stakeholders have begun adopting particular benchmarks and content to match the needs of Malaysian teachers and learners. While the CEFR framework and notions establish an affinity to Malaysian education instruction, voices from language teachers in classrooms are critical and need to be given consideration (Mison & Jang, 2011).

Even though the CEFR-aligned curriculum and syllabus have just been formally started to be implemented in all Malaysian primary and secondary schools in 2017, there were several studies done to investigate how the CEFR would affect teaching and assessment practices in the Malaysian classroom contexts. These studies ranged from views and effects of curriculum reforms, curriculum alignment, CEFR impact or influence (Uri & Aziz, 2018a), and many more. However, most studies done in relation to CEFR and the Malaysian education curriculum are quantitative studies on tertiary educators and teachers' views or beliefs and other viewpoints from stakeholders such as government officials (Uri & Aziz, 2018a), parents (Iber, 2014) and CEFR trainers (Aziz et al., 2018a). Most of these studies reported that teachers' beliefs on the implementation of CEFR do not correspond with their classroom practices. However, almost all of these studies have reported the CEFR implementation in the early stages by teachers due to the implementation (second wave) only started in 2017 or gathered perceptions from educators that have yet to fully utilise the CEFR in their respective institutions. Connecting to Aziz et al.'s (2018) reflective report as national master trainer for the CEFR cascade training, the insufficient practice by teachers may be due to the ineffective training done during the third-tier stages in which the District Trainers had to train the ESL teachers. It will be interesting to see whether these perceptions and practices have changed or have gained prominence after a few years of implementation by the teachers, coupled with recently reported ongoing professional development training by the trainers.

Sidhu et al., (2018) in their study on CEFR-aligned school-based assessment in Malaysian primary ESL classrooms stated that the implementation of CEFR in schools is still insufficient. Even though teachers are positive and generally receptive of the CEFR framework and its advantages (Lo, 2018; Uri & Aziz, 2018a), most teachers indicated lack of understanding and awareness of the incorporation of CEFR into classroom assessment as teachers needed more guidance and training for them to fully understand and utilise more innovative CEFR-aligned assessments in their classrooms (Aziz et al., 2018). Aziz et al. (2018) indicated that despite several efforts made in training Malaysian teachers to apply and practice CEFR-aligned teaching and assessment in the past two years, there are still various aspects that need improvement.

In Sidhu et al. (2018) and Moonen et al.'s (2013) studies, it was found that many teachers are still falling back to the conventional textbook exercises as their standard practice and guide in developing students' ESL proficiency. The conventional methods used by teachers from the findings indicate that the teaching and learning in standard classrooms are still teacher-based where teachers act as instructors. Many are still primarily focused on task outcomes rather than developing their students' proficiency (Lo, 2018) due to time constraint in finishing the syllabus

DISCUSSION

From the review of relevant literature, it can be summarised that there are several similarities and differences between the implementation of the CEFR in various countries around the world. Both European and Asian countries use the CEFR to gauge their own English language learners' proficiency and many studies focused on rater and inter-rater professional development. It can also be seen that there is ongoing development in all countries to improve the CEFR implementation such as the inclusion of alternative assessments, peer and self-evaluation in classrooms and more CEFR-aligned tools that educators can use in their English language classrooms. In European countries, the CEFR is mostly not a compulsory element for schools and teachers are not being forced to adopt the CEFR into their classrooms. It serves as an alternative benchmark; hence, there is a lack of willingness to adopt the CEFR. However, in most Asian countries, the CEFR is being

forcibly implemented by each country's government and educational bodies. Most of these rapid implementations in which the government expected its citizens to improve their English language proficiency and communication skills have led to several drawbacks which were discussed in the literature review. Even though the CEFR has been introduced as early as 2001, Malaysia and several other Asian countries have started to adopt the framework only recently as early as 2011. Malaysia adopted the CEFR in 2013 and it is still in the first cycle of its implementation and evaluation phases.

There is a common theme or problem which appears in almost each of the countries mentioned in the review of literature of this paper. Even though CEFR has been positively accepted by stakeholders across each country, especially teachers and policymakers, there exists a lack of transparency of the purpose of implementing CEFR between the ministry or policymakers (top level) and teachers who will be the ones implementing CEFR at the bottom level. The echoing common problem faced by the countries is insufficient teacher training and professional development.

The CEFR has been so prominent worldwide and its benefits to policymakers and educational administrators are difficult to overlook (Read, 2019). This shows that it is simple to believe that the framework can be utilised comprehensively to second language learning situations. However, many language educationists have come to recognize that the CEFR must be adapted if it is to play a prominent part in outlining language objectives and curricula in their own education systems. The epitome of universal benchmarking of learner attainments in developing second languages has been balanced against the variety of social and educational contexts in specific countries (Read, 2019). As noticed in various countries, the means to align and implement CEFR requires a long time and governments are currently improving it based on several drawbacks. Nevertheless, countries could have a better implementation of CEFR by looking into each other's development and evading the mistakes made.

Top-down change, rather than unclear changes in educational aspirations, is the only sensible means to produce progressive change throughout the system (Fennelly, 2016; Zheng et al., 2016). When the CEFR is presented into individual educational contexts, teachers frequently face difficulties in comprehending and employing the theoretical principles without tangible examples (Mison & Jang, 2011). In response to this, teachers claimed that CEFR was still at its infancy, too general and too theoretical to be adapted to classrooms in both hypothetical and pragmatic senses. This is due to the lack of an organised administration or official group to control such an initiative at the state level; thus, it deterred its wider acceptance by teachers (Arnott et al., 2017). There also exists a group of teachers in the grey area, those still contemplating the CEFR's viability for their classroom. The alterations made by the implementation of CEFR may be viewed as a threat to current practices established from their experience (Baldwin, 2018). Professional development workshops for teachers can help to augment the implementation of CEFR and should take into account the level to which CEFR is carried out in order to present language teachers the support and emphasis that are most achievable for their professional context (Moonen et al., 2013). Currently, though the ministry is executing more training for teachers, the teachers are not appropriately prepared to authentically achieve the specified education targets or to adapt to the CEFR influence. Teacher training connecting to how to foster students' ability to achieve the CEFR grading levels is crucial (Fennelly, 2016). All in all, more devotion should be given to boost teachers' crucial capability to satisfy the requisites and benchmarks in the CEFR (Khang, 2018).

LIMITATION OF STUDY

As much as the study achieved the researchers intended for this study, there were several limitations to it, from the ability to generalise conclusion based on a mere 25 journal papers on the execution of the CEFR in various countries. Nevertheless, the papers selected generally show the similar issues faced by Malaysia and various countries that need to be taken into account and relevant to the implementation of the CEFR into a national curriculum and educational syllabus.

CONCLUSION

Taking into account the research and data reviewed by Arnott et al. (2017), it is proposed that certain significant areas should be focused on by future research such as ongoing macro- and micro- policy developments, CEFR-informed initial teacher education programmes, supervision of L2 teachers and language testing. Possible reasons for teachers' challenges with the CEFR include: the abstract nature of the CEFR document, lack of research into school-based uses of the CEFR, and the teachers' beliefs and cultures. In other words, it is not a document of policy, curriculum or assessment that obstruct the implementation of CEFR but the lack of support by its direct stakeholders that fail to comprehend the complexity of the theoretical framework and its application of principles which can lead to the failure of its successful implementation (Mison & Jang, 2011).

In conclusion, this paper has revealed that, generally, teachers have a rudimentary grasp of CEFR and value its usefulness as a means to measure target language competence on a universal level. However, transparency and consistency are required to further reinforce the usage of CEFR in schools and classrooms. This can be obtained in educational approaches and professional developments of teachers who serve as the groundwork to ensure the CEFR-aligned implementation in any country is a success. Hence, their voices need to be heard and their perspectives on the CEFR framework need to be shared in order to reach the high expectations and requirements of what is presently lacking in Malaysian ESL education. The CEFR-aligned syllabus and assessment can then be further standardised for clearer comprehension and implementation of CEFR in Malaysian classroom teaching and assessment.

ACKNOWLEDGEMENT

The researchers would like to acknowledge Universiti Teknologi MARA, Cawangan Pulau Pinang for supporting this paper.

REFERENCES

Arnott, S., Brogden, L.M., Faez, F., Peguret, M., Piccardo, E., Rehner, K., Taylor, S.K. & Wernicke, M. (2017). The Common European Framework of Reference (CEFR) in Canada: A research agenda. *The Canadian Journal of Applied Linguistics, 20*(1), 31-54.

Aziz, A. H. A. A., Ab Rashid, R., & Zainudin, W. Z. W. (2018). The enactment of the Malaysian Common European Framework of Reference (CEFR): National master trainer's reflection. *Indonesian Journal of Applied Linguistics, 8*(2), 409-417. doi:10.17509/ijal.v8i2.I3307.

Azman, H. (2017). Implementation and challenges of English language education reform in Malaysian primary schools. *3L: The Southeast Asian Journal of English Language Studies, 22*(3), 65-78.

Baldwin, R. & Apelgren, B. M (2018). Can Do and Cannot Do – CEFR inspired examination and assessment in a Swedish higher education context. *Apples – Journal of Applied Language Studies, 12*(2), 19-35. http://doi: dx.doi.org/10.17011/apples/urn.201809144127.

Bharati, D.A.L. & Lestari, T. (2018). A need analysis in developing project-based writing assessment module to stimulate students' critical thinking and creativity. *KnE Social Sciences,* 130-138. Doi: 10.18502/kss.v3i18.4706

Cox, T.L., Malone, M.E. & Winke, P. (2017). Future directions in assessment: Influences of standards and implications for language learning. *Foreign Language Annals, 51*(1), 104-115. doi: 10.111/flan.12326

Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment: Companion Volume with New Descriptors*. Retrieved on 7 December 2020 from https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989

Fennelly, M.G. (2016). The influence of CEFR on English language education in Japan. *Bull Shikoku University, 46*(A), 109-122.

Fergus, O. (2015). Toward critical, constructive assessments of CEFR-based language teaching in Japan and beyond. *Osaka University Knowledge Archive: OUKA, 41,* 191-204. doi: 10.18910/51427.

Foley, J.A. (2019). Issues on assessment using CEFR in the region. *LEARN Journal: Language Education and Acquisition Research Network Journal, 12*(2), 28-48.

Hamzah, M.S.G., Abdullah, M.Y., Hussin, A.R., Abdullah, S.K. Wong, K.T. & Ismail, S. (2018). Analisis pelaksanaan program memartabatkan Bahasa Malaysia (BMB) Sekolah Rendah dan Menengah di Malaysia. *Journal of Kesidang, 3*(1), 40-61.

Holzknecht, F., Huhta, A. & Lamprianou, I. (2018). Comparing the outcomes of two different approaches to CEFR-based rating of students' writing performances across two European countries. *Assessing Writing, 37,* 57-67. doi: https://doi.org/10.1016/j.asw.2018.03.009.

Iber, G. (2014). English language teaching in Malaysia: The case for a dual-track English curriculum [Short communication]. *Advances in Language and Literacy Studies, 5*(4), 68-71. doi: 10.7575/aiac.alls.v.5n.4p.68

Idris, M. & Raof, A.H.A. (2017). The CEFR rating scale functioning: An empirical study on self- and peer assessments. *Sains Humanika, 9*(4-2), 11-17.

Kaur, P. & Shapii, A. (2018). Language and nationalism in Malaysia: A language policy perspective. *International Journal of Law, Government and Communication, 3*(7), 1-10.

Khang, N.D. (2018). Today's teachers' CEFR competence in the classroom – A view of critical pedagogy in Vietnam. *Theria et Historia Scientiarum, 15,* 121-138. doi: http://dx.doi.org/10.12775/ths.2018.008.

Kimura, Y., Nakat, Y., Ikeno, O., Naganuma, N. & Andrews, S. (2017). Developing classroom language assessment benchmarks for Japanese teachers of English as a foreign language. *Language Testing in Asia, 7*(3), 1-14. doi: 10.1186/s40468-017-0035-2.

Le, H. T. T. (2018). Impacts of the CEFR-aligned learning outcomes implementation on assessment practice. *Hue University Journal of Science: Social Sciences and Humanities, 127*(6B), 87-99. doi:10.26459/hueuni-jssh.v127i6B.4899.

Liu, L. & Jia, G. (2017). Looking beyond scores: validating a CEFR-based university speaking assessment in Mainland China. *Language Testing in Asia, 7*(2), 1-16. doi: 10.1186/s40468-017-0034-3.

Lo, Y. Y. (2018). English teachers' concern on Common European Framework of Reference for Languages (CEFR): An application of CBAM. *JuKu: Jurnal Kurikulum & Pengajaran Asia Pasifik, 6*(1), 46-58.

Ministry of Education Malaysia (2013). Malaysian Education Blueprint 2013-2025 (Preschool to Post-secondary Education).Retrieved on 7 December 2020 from https://www.moe.gov.my/menumedia/media-cetak/penerbitan/dasar/1207-malaysia-education-blueprint-2013-2025/file

Mison, S., & Jang, I. C. (2011). Canadian FSL teachers' assessment practices and needs: Implications for the adoption of the CEFR in a Canadian context. *Synergies Europe, 6*, 99-108.

Moonen, M., Stoutjesdijk, E., de Graaff, R., & Corda, A. (2013). Implementing CEFR in secondary education: impact on FL teachers' educational and assessment practice. *International Journal of Applied Linguistics, 23*(2), 226-246.

Read, J. (2019). The influence of the Common European Framework of Reference (CEFR) in the Asia-Pacific Region. *LEARN Journal: Language Education and Acquisition Research Network, 12*(1), 12-18.

Sidhu, G.K., Kaur, S. & Lee, L.C. (2018). CEFR-aligned school-based assessment in the Malaysian primary ESL classroom. *Indonesian Journal of Applied Linguistics, 8*(2), 452-463. DOI: 10.17509/ijal.v8i2.13311

Uri, N. F. M., & Aziz, M. S. A. (2018a). Implementation of CEFR in Malaysia: Teachers' awareness and the challenges. 3L: *Language, Linguistics, Literature, 24*(3), 168-183. doi:10.17576-2018-2403-13.

Uri. N.F.M. & Aziz, M.S.A. (2018b). Neoliberalism and its impact on English as a medium of instruction in Malaysia, Thailand and Vietnam. *Journal of Creative Practices in Language Learning and Teaching (CPLT), 6*(1), 17-29.

Zheng, Y., Zhang, Y. & Yan, Y. (2016). Investigating the practice of the Common European Framework of Reference for Languages (CEFR) outside Europe: A case study on the assessment of writing in English in China. *ELT Research Papers, 16*(1), 3-20.

Zou, S. & Zhang, W. (2017). Exploring the adaptability of the CEFR in the construction of a writing ability scale for test for English majors. *Language Testing in Asia, 7*(18). doi: 10.1186/s40468-017-0050-3.

Zuraidah Mohd Don & Mardziah Hayati Abdullah (2019, May 23). *The reform of English language education in Malaysia*. Free Malaysia Today. Retrieved on 7 December 2020 from https://www.moe.gov.my/en/menumedia/printed-media/newspaper-clippings/the-reform-of-english-language-education-in-malaysia-free-malaysia-today-22-mei-2019

ABOUT UMT FACULTY

SDI

Selective Dissemination of Information (SDI) service is a current-awareness service offered by the PSNZ for UMT Faculty Members. The contents selection criteria include current publications (last 5 years), highly cited and most viewed/downloaded documents. The contents with pdf full text from subscribed databases are organized and compiled according to a monthly theme which is determined based on the topics of specified interest.

For more information or further assistance, kindly contact us at 09-6684185/4298 or email to psnz@umt.edu.my/sh_akmal@umt.edu.my

Thank you.

Perpustakaan Sultanah Nur Zahirah
Universiti Malaysia Terengganu
21030 Kuala Nerus, Terengganu.

Tel. : 09-6684185 (Main Counter)
Fax : 09-6684179
Email : psnz@umt.edu.my