"Every teacher was an island": Teacher perceptions of a CEFR alignment project to implement a standardized approach to assessment

Carolyn Westbrook, British Council, Great Britain

Aidan Holland, British Council, Great Britain

Implementing fair and valid assessment is a fundamental part of a teacher's job as assessments enable learners to demonstrate progress as well as highlighting strengths and weaknesses (Rahman 2018). Results should be reported in an accessible way so learners and other stakeholders understand the outcomes and learners' progress. This article reports on a collaboration between assessment researchers and teachers to standardize the approach to assessment in a global secondary language programme. Teaching materials were mapped to the CEFR, standardized set-up notes were created for assessment tasks, assessment tools and training were developed, and teachers implemented the approach in a practical way. Feedback showed the system improved objectivity and clarity in assessment, though challenges around feasibility and alignment with the CEFR Companion Volume (CEFR/CV) (Council of Europe [CoE] 2020) remained.

Keywords: CEFR alignment, classroom-based language assessment, teacher assessment, teacher-researcher collaboration, standardized approach to assessment

Acknowledgements

1 Introduction

Assessment is a key part of teaching and learning. As Race et al. (2005: xi) note, "[n]othing we do to, or for our students is more important than our assessment of their work and the feedback we give them on it". Effective assessment highlights learners' strengths and weaknesses, enables them to demonstrate progress (Rahman 2018), and shapes classroom practice (Bachman and Palmer 2010). To be meaningful, assessments must be valid, reliable, and practical for teachers to use, while results should be reported in accessible ways that promote transparency and support learner progress.

In 2016, the British Council began standardising its secondary teaching materials across 50 countries. Courses were based around CEFR levels, yet the teaching materials were not formally aligned to the

CEFR. The teaching materials are based around topic modules, each of which is presented in the form of a magazine, culminating in a task-based project. Each CEFR (sub-)level comprises 10 magazines. However, assessment remained inconsistent: there were no shared criteria, limited guidance on evaluating projects, and no explicit alignment of tasks to the CEFR, leaving teachers uncertain about standards, and stakeholders unclear about learner progress.

This article reports on a collaborative project between researchers and teachers to create a CEFR-aligned assessment system for the British Council's *Secondary Plus* courses, delivered to learners aged 11-17 worldwide. The project involved mapping course materials to the CEFR, developing a standardized approach to assessment, developing assessment and reporting tools, training teachers, and collecting feedback on implementation. The article presents teachers' perceptions of the new system and the challenges encountered, particularly regarding use of descriptors from the CEFR/CV (CoE 2020).

2 Literature review

Traditionally, curriculum, delivery (i.e., the operationalization of the curriculum in a specific context) and assessment have been seen as separate entities: teachers focus on teaching and learning but curricula are designed by publishers or education boards while assessment experts design assessments (O'Sullivan 2021). However, Bunch (2012: 1) argues that "[a] key component of educational achievement test validation is alignment of the test to both curriculum and instruction". Similarly, O'Sullivan (2021) posits that curriculum, delivery and assessment must be inextricably linked for a learning system to function:

> Within the system, the three core elements (curriculum, delivery, assessment) must be based on a single philosophy of learning, supported by clearly defined models of language ability and progression, and underpinned by a measurement model. Failure to ensure that all three are fully in harmony is likely to lead to the failure of the system. (O'Sullivan 2021: 2)

The philosophy of learning in the CEFR is the action-oriented approach. Learners are seen as 'social agents' who have tasks to accomplish in a particular context (CoE 2001: 9). The CEFR provides "a common basis for the explicit description of objectives, content and methods" (CoE 2001: 1), thereby enabling curricula, teaching materials and assessments to be aligned to external standards such as the CEFR.

The document *Aligning Language Education with the CEFR: a handbook* (British Council et al. 2022) outlines five stages for aligning materials to the CEFR:

1. Familiarization—participants are familiarized with the CEFR descriptors so they can apply them accurately during specification.

2. Specification—descriptors that match the skills and competences targeted by materials are identified.

3. Standardization—experts analyse materials and establish benchmark samples that exemplify performance at each CEFR level.

4. Standard setting—minimum standards for each level are determined.

5. Validation—gathers evidence to support the alignment claims.

Although traditional assessment has focussed on assessing the four skills plus grammar and vocabulary, the (2001) CEFR and its *Companion Volume* highlight four modes of communication: reception, production, interaction and mediation. These reflect authentic language use (Plakans 2020). Technological advances have also enabled multimodal assessments, which enhance motivation, promote critical thinking (Varaporn and Sitthitikul 2019) and support inclusivity for learners with Special Educational Needs and Disabilities (SEND) (Ellis 2024).

3 Project phases

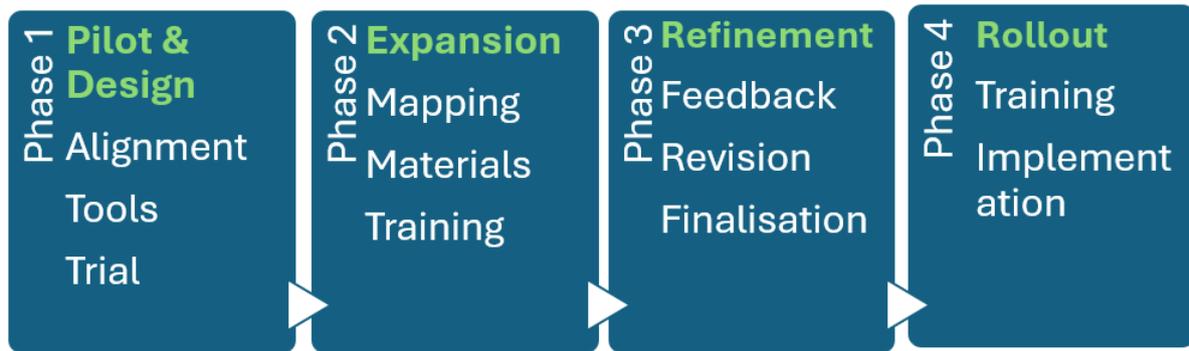The project progressed through four main phases. These are summarized in Figure 1.

Figure 1. The four project phases

3.1 Phase 1

Phase 1 was a pilot phase which involved mapping one set of ten magazines at the B1.1 level to the CEFR using the British Council's internal CEFR database, the British Council Curriculum Framework (BCCF), and the CEFR/CVfollowing the steps outlined in the Handbook (British Council et al. 2022). The BCCF was created before the publication of the CEFR/CV so descriptors for some scales were missing, hence the need to use the CEFR/CV too. Since each magazine culminated in a multimodal project, these projects were used as the basis for the assessment and amended as necessary to reflect the content of the magazine, the specific learning outcome and the level of the magazine. Phase 1 also involved creating standardized project set-up notes for teachers, assessment and reporting tools using CEFR Can Do descriptors and specially developed Performance Indicators (PIs) which needed to be feasible for teachers with limited time for marking. The Performance Indicators are included in the BCCF and provide additional information about performance at the different CEFR levels, broken down into six areas: focus, fluency, range, accuracy, discourse, and appropriateness. This phase culminated in a small-scale trial using the projects and the assessment and reporting tools produced.

3.2 Phase 2

In the second phase, the CEFR mapping was extended to all Secondary Plus levels from A1 to C1. In total, 12 levels, each with 10 magazines per level, were mapped, resulting in 120 magazines being mapped to the CEFR. Standardized set-up notes for teachers, assessment and reporting tools for all 120 projects were created as well as a training pack for teachers including a handbook and a training presentation. These materials were then trialled, and samples were collected for standardization and benchmarking. An expert panel of six teachers and researchers led the standardization and standard setting phases, producing benchmark samples to support future training and ensure consistency across levels.

3.3 Phase 3

The third phase involved finalising the assessment sheets and teacher training materials based on the feedback received.

3.4 Phase 4

In this final phase, rollout commenced. This involved training the teachers to implement the new assessment approach using the teacher training materials developed previously. Training was delivered

to 600 teachers and academic leads, and the approach has been implemented in 25 countries around the world. The rollout is currently on hold due to ongoing transformation within the organization.

4 Results

The teachers who were trained as part of the initial rollout were asked to complete a brief pre-training questionnaire to gauge their level of teaching experience and the amount of prior assessment training. After undertaking the training, they were asked to complete a post-training questionnaire. A selection of the results from both questionnaires is reported below.

4.1 Pre-training questionnaires

The pre-training questionnaire was completed by 195 teachers. Questions 1-5 and 16-17 collected consent and background data from the respondents. Due to space limitations, only a selection of key results is presented here.

Question 6 asked about teaching experience. Sixty-six percent of respondents had between 11 and 30 years' experience teaching, 8% had over 31 years' experience, and the remainder (26%) had been teaching for 10 years or less.

Despite all respondents confirming that they assessed their learners regularly, 23% had not had any training in assessment (Q.7). Of the remaining 77% who had had some training in assessment, this was often only short sessions of up to two days (Figure 2).
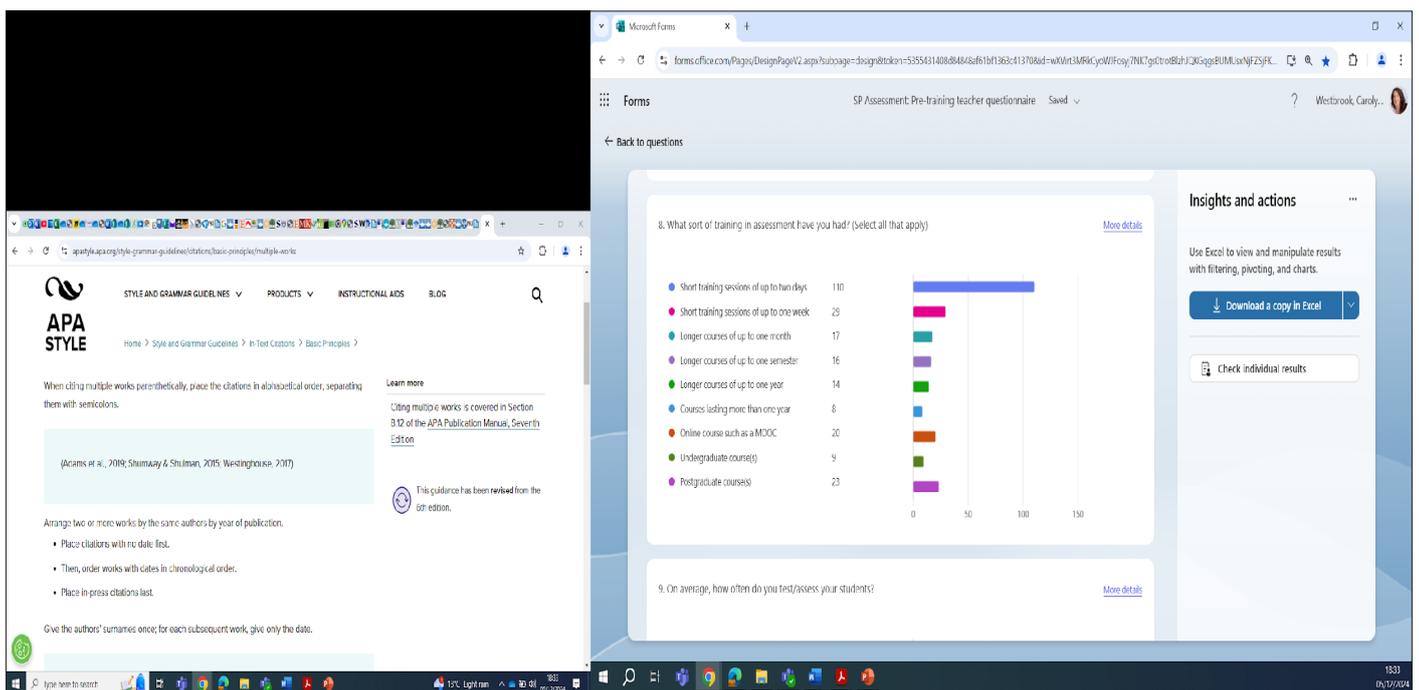

Figure 2. Responses to "What sort of training in assessment have you had?"

Question 13 asked if the assessments helped teachers teach; 79% responded yes. Only 5% of participants replied that this was not the case, while 16% were unsure. Question 14 expanded on the previous question by asking how the assessments helped teachers. Many respondents stated that the assessments provided information about learners' weaknesses, which helped teachers to plan for future lessons and to see progress:

I can see how well the ss [sic] have understood the material covered and what they need for the future. So it helps me mold the course to the needs of my ss [sic]. (T18)

[Assessments] help to see if my learners are making progress ... (T119)

Finally, question 15 looked at teachers' satisfaction with assessment up to that point (Figure 3). 50% were either very satisfied or satisfied, while 41% were neutral. The remaining 9% were either dissatisfied or very dissatisfied.
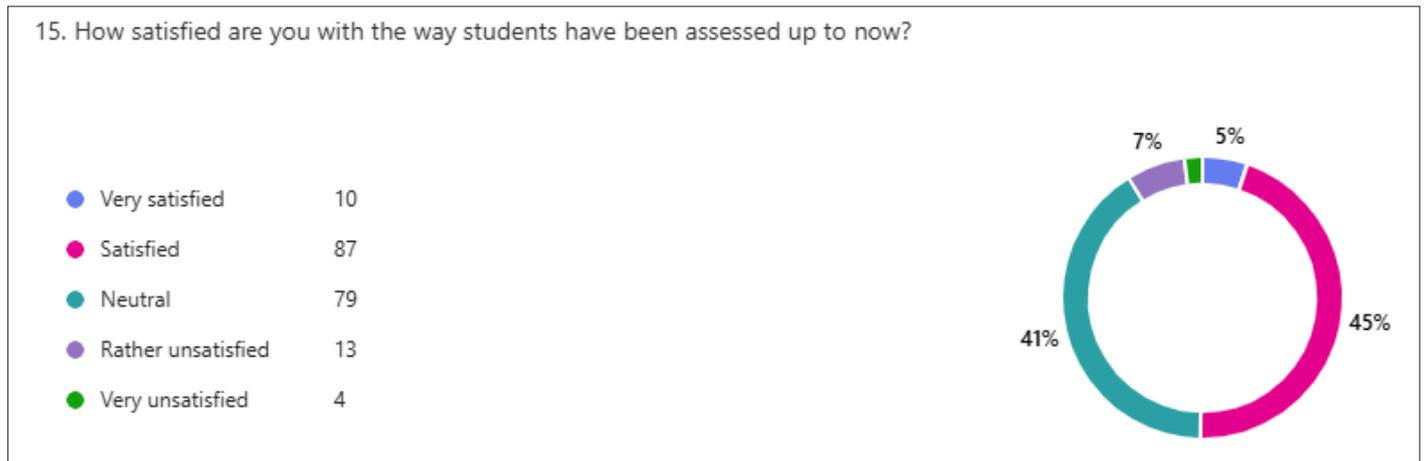

15. How satisfied are you with the way students have been assessed up to now?

Very satisfied	10
Satisfied	87
Neutral	79
Rather unsatisfied	13
Very unsatisfied	4

Figure 3. Responses to "How satisfied are you with the way students have been assessed up to now?"

4.2 Post-training questionnaires

Ninety-four teachers completed the post-training questionnaire (59 of them had completed the pre-training questionnaire), providing feedback on their experiences with the new assessment approach.

Questions 1-3 and 9-10 collected consent and background information. Questions 4-8 investigated teachers' perceptions of the new approach. Figure 4 shows that 72% of respondents agreed or strongly agreed that the new assessment system would enable them to assess more objectively (question 4).
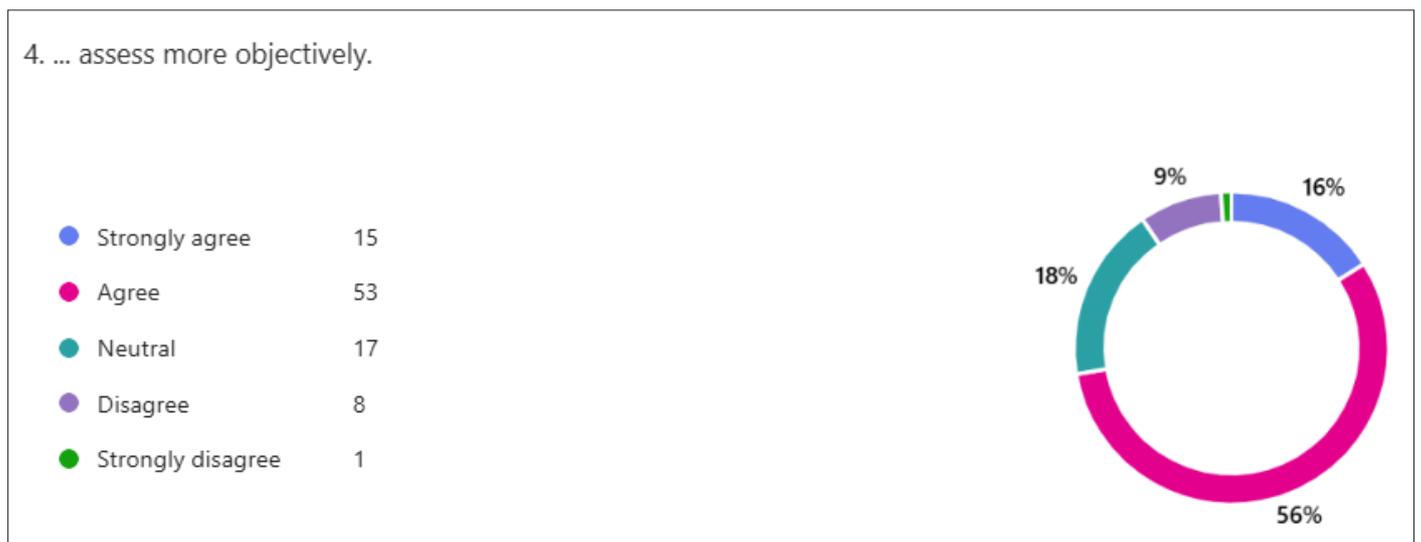

4. ... assess more objectively.

Strongly agree	15
Agree	53
Neutral	17
Disagree	8
Strongly disagree	1

Figure 4. Responses to "The new assessment system will enable me to assess more objectively"

For question 5, 62% either agreed or strongly agreed that the new approach would improve student outcomes/results while 30% were neutral (see Figure 5).
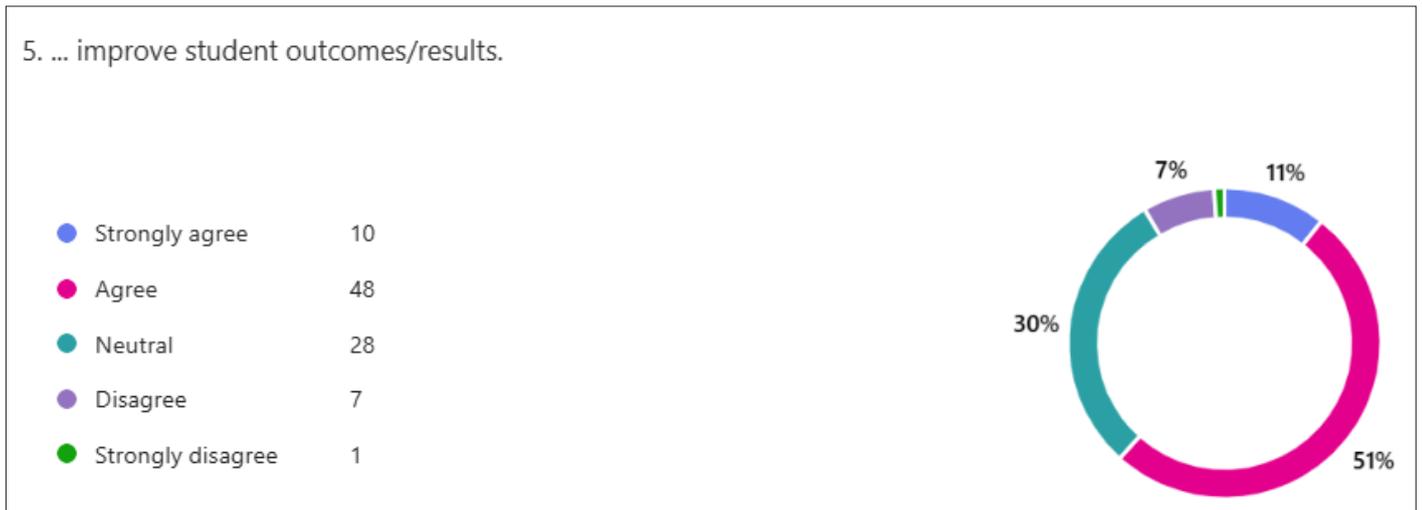

Figure 5. Responses to "The new assessment system will enable me to improve student outcomes / results"

Question 6 (Figure 6) asked whether the new approach would enable teachers to provide better feedback to students and parents. 78% either agreed or strongly agreed that this was the case.
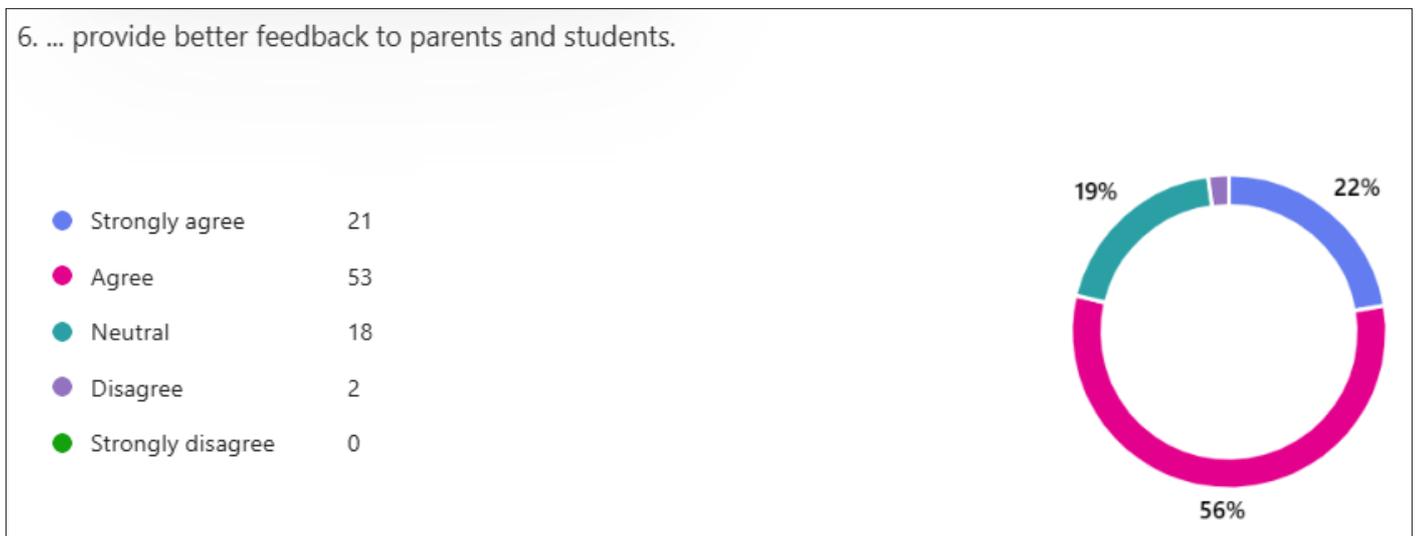

Figure 6. Responses to "The new assessment system will enable me to provide better feedback to parents and students"

Similarly, 78% either agreed or strongly agreed that the new approach would enable them to better prepare their students for the project/task (question 7, Figure 7).
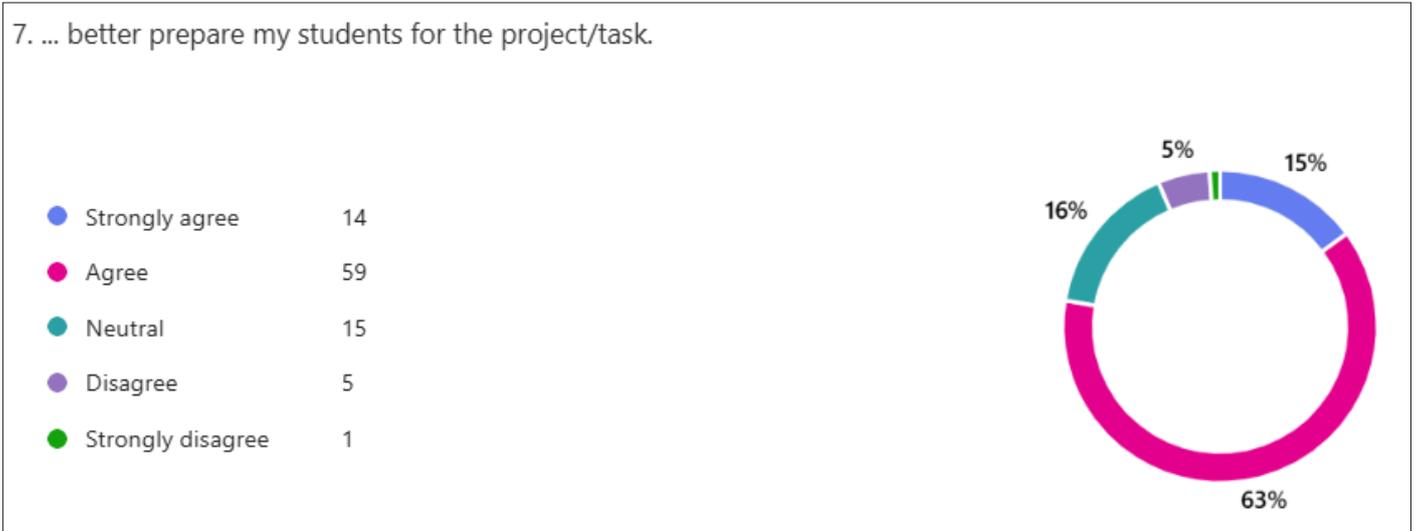
Figure 7. Responses to "The new assessment system will enable me to better prepare my students for the project/task"

For question 8, 54% either agreed or strongly agreed that the new approach would enhance how they teach; 32% were neutral (Figure 8).
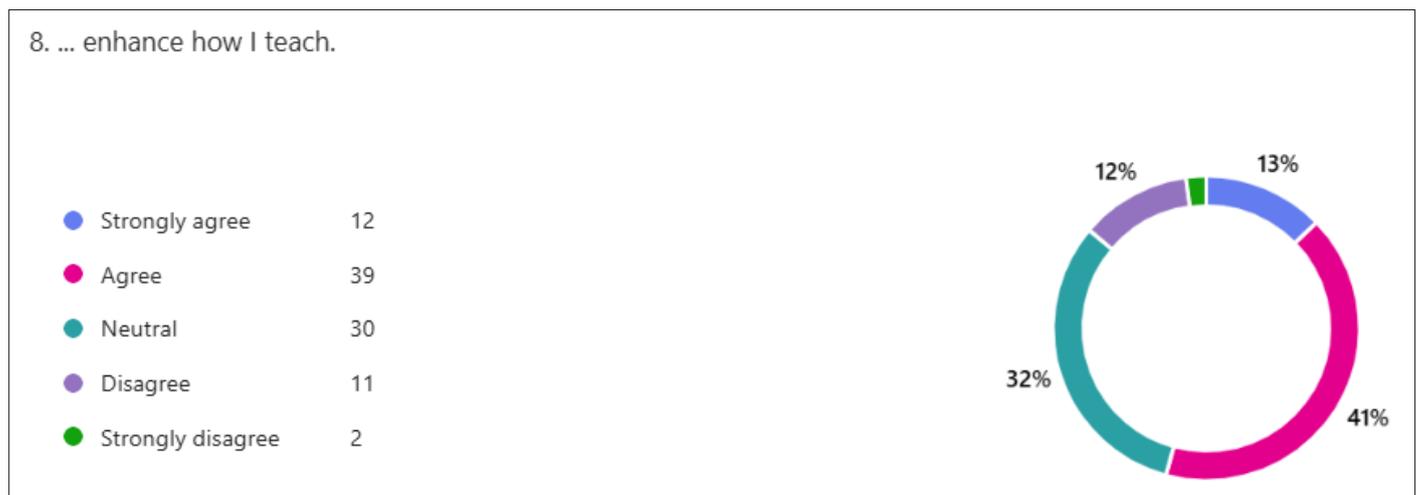

Figure 8. Responses to "The new assessment system will enable me to enhance how I teach"

For questions 5 and 8, 30% and 32% of respondents respectively were neutral. This may have been because they had not yet implemented the new approach so they were unsure of its impact.

4.3 Qualitative analysis of feedback from focus groups

Two focus groups and two interviews were carried out with teachers who had implemented the approach with two or more magazines to collect their feedback about standardization in the new approach, the number of assessments, and the ease of use of the Can Dos and the PIs.

4.3.1 Standardization in the new assessment approach

Some participants reported issues with the previous approach which led to bias and inconsistency, and welcomed the standardization in the new approach:

> *[The previous assessment approach] had its pluses and minuses. Obviously, standardization was the biggest minus because we all had our different ideas and … we know our students well. So … we already have an idea of what grade to give our students. I think there's a bit of bias based on their past performance …* (I1M)

> *… [previously] everybody was their own island … that is the beauty of this new assessment system, because with my class, it was just me and what I saw fit for my students, … and there was inconsistency, right, between … me and other teachers.* (I2G)

> *… the parents will want it to be standardized because [that's] what they're used to [from school] …* (FG3H)

Others were less happy about the uniform approach, feeling like they were expected to "become robots" (P1-1) or that it took away some of the fun of teaching the Secondary Plus courses:

> *… when students come to do the First Certificate exam, … they might not enjoy it, but they know they're preparing [sic] the specific certificate exam [whereas] often with the [Secondary Plus] magazines, it's [sic] viewed … by teachers and also maybe by students that [they] should be … enjoyable …* (FG4J)

4.3.2 Number of assessments

The number of assessments that would be feasible in the new approach was discussed repeatedly during the project. There were differing opinions within the project team and among the teachers in the pilot phases. In the end, the decision was made to have only one standardized assessment per module. However, opinions among focus group and interview participants would suggest that more than one would be appropriate:

> *… I would say … test them on everything …* (FG3H)

> *I really don't think that parents would just wear one assessment per module.* (FG3K)

> *I don't think it's too tough to do two assessments.* (I1M)

4.3.3 Ease of use of Can Dos and PIs

For a standardized assessment approach to function, teachers need to feel confident that they understand and can apply the Can Dos and PIs when marking learners' work. Participants were generally positive about the grade bands, the Can Dos and PIs. They felt they were useful and easy to apply:

> *Just three categories makes it really, really simple. And I think it's quite clear as well.* (I1M)

> *I think they're fair … they're clear … they're achievable.* (I2G)

But many participants agreed that hitting all six PIs was challenging for learners:

> *… hitting all of them is really hard.* (I1M)

5 Discussion

The pre-training questionnaires demonstrated that most teachers had had some assessment literacy training before doing the training for the new assessment approach. Teachers assess their learners regularly and they see the value of assessments in identifying strengths and weaknesses (Rahman 2018) and helping to shape classroom practice (Bachman and Palmer 2010) by providing information to teachers regarding areas that learners struggled with so teachers can go over these again.

The post-training questionnaire responses and the comments from focus group participants suggest that teachers recognized the strengths of the new approach while at the same time noting challenges with the previous approach. One of the weaknesses of the latter lay in the lack of standardization and reliance on subjective interpretations whereas the new approach allows teachers to evaluate more objectively and provide clearer feedback to learners and carers.

Despite the perceived benefits of the new approach, some teachers felt constrained by the new system. The data suggests that this perception might have several causes. Firstly, the new assessment system was seen as potentially stunting opportunities for teacher creativity, akin to the effect which preparing for large-scale exams might have. Furthermore, the limited assessment literacy (Stiggins 2014) among teachers may have been another obstacle to accepting the new approach. Consequently, any resistance could be due to a lack of understanding of the importance of standardized, reliable marking. Overall, though, the results from the pre- and post-training questionnaires and the focus groups suggest that the system contributes to greater clarity and purpose in teachers' assessment processes and demonstrates the system's potential to positively impact teaching while underscoring the need for continued support and training.

The feedback also highlights the tension between achieving construct validity and practicality. While the assessment researchers emphasized the importance of multiple assessments to mitigate the risk of construct under-representation (Messick 1989), practical constraints led to a compromise of one standardized project assessment per module, carried out by teachers, with the option of an additional assessment for each module based on context. This underscores the challenge of balancing best practices in assessment with the realities of teaching contexts, particularly where workloads and contractual limitations are significant factors.

Teachers reported feeling comfortable and confident in using and applying the Can Dos and PIs. However, the general consensus that hitting all six would be a challenge for learners is something that could be reviewed for future academic cycles. Alternatively, additional information could be provided to better prepare learners for the assessments, which may occur organically as teachers become more accustomed to the new approach.

6 Challenges

As with any project, there were challenges along the way. The challenges were two-fold, relating to internal issues on the one hand and the use of the BCCF and CEFR/CV on the other.

6.1 Internal challenges

One challenge was that teachers frequently adapt Secondary Plus materials to suit their classes and their contexts due to cultural differences and interests. However, if the materials are amended too much, the learners may not have covered everything necessary to successfully complete the project, thereby leading to a mismatch between delivery and assessment (O'Sullivan 2021). To avoid this mismatch, the project team provided detailed notes so teachers would know what aspects to focus on during the magazine even if input was amended.

The next challenge was the debate around how many assessments would be feasible for teachers. Due to contractual constraints, many teachers are unable to spend much time marking outside of class,

so a key focus of the approach had to be on practicality. Further to many discussions within the project team, the decision was made to have only one standardized assessment per module. Some teachers considered this sufficient while others wanted more than one assessment, often because that is what parents in those contexts expect. As a result, a second assessment task was created for any centres requiring an additional assessment.

Finally, many of the projects originally culminated in presentations, so some projects were amended to have a different output. This may have been because the materials were not empirically aligned to the CEFR when they were designed, thus resulting in a lack of construct coverage across the possible CEFR scales. However, since the alignment exercise highlighted this construct under-representation, this can be addressed when revising and updating the materials.

6.2 BCCF & CEFR/CV challenges

When aligning the materials to the CEFR using the BCCF, there were issues with missing descriptors since the BCCF had been created before the CEFR/CV was published. As a result, the new scales and descriptors, most notably the mediation scales, were missing so we had to refer to the CEFR/CV directly. However, the CEFR/CV was not designed with young learners in mind so, in line with the recommendation in the CEFR/CV to "select ... [and] adapt the formulation of [descriptors] ... to better suit the specific context" (CoE 2020: 42), some descriptors were amended slightly.

Another challenge was that the CEFR/CV was also missing descriptors in some scales. For example, the *Explaining data* scale includes descriptors for the plus levels at A2 and B1 but not at B2. Therefore, differentiating between a B2.1 performance and one at B2.2 is challenging. Given that the Secondary Plus courses are broken down into B2.1 and B2.2, we adopted the B2 descriptors and tried to differentiate between levels using other descriptors which were also appropriate for the particular task.

Similarly, some scales had descriptors for one side of an interaction but not the other. For example, in the *Interviewing and being interviewed* scale, there are no descriptors for the interviewer role below B1.1 nor at B2.1. Conversely, there is no interviewee role at B2.2. Where the descriptors for certain roles were missing, we amended the assessment task so that the assessed role was the one for which there was a corresponding descriptor.

Finally, in both the BCCF and the CEFR/CV, there are often several descriptors for one CEFR level or descriptors with multiple examples, which are not all relevant. Again, the project team had to decide which descriptors to adopt and which to adapt.

7 Conclusion

This project contributed to fairer and more consistent assessment by mapping teaching materials to the CEFR to ensure the materials targeted the correct CEFR level, and developing a standardized approach to assessment in which teachers assess learners using the project at the end of each magazine, using standardized set-up notes, criteria, and reporting tools. Teachers were trained in the assessment approach and standardized using CEFR-aligned benchmarks. The project created stronger collaboration between researchers and teachers, integrating classroom insights into the design process. Despite challenges such as balancing workload with assessment demands and addressing gaps in the BCCF and CEFR CV, the results of the mapping exercise and the alignment of the assessments to the CEFR confirmed that the Secondary Plus materials broadly target the intended CEFR levels. Furthermore, the new tools should make learning more tangible for learners and promote objectivity in teachers' assessments.

8 Applications

Beyond the immediate context, the project illustrates how collaborative alignment initiatives can strengthen assessment literacy, improve transparency for stakeholders and build consistency in applying marking standards. It also shows that engaging teachers throughout the project increases the feasibility and acceptance of new approaches.

9 Future directions

Moving forward, the project highlights the need to update and extend frameworks such as the BCCF and CEFR/CV, including descriptors tailored for young learners. It also points to the value of continued teacher training and support, both to address perceptions of constraint and to sustain consistent application across contexts. More broadly, similar alignment projects could inform assessment practices in other large educational programmes, supporting fairer, more transparent learning systems internationally.

10 References

Bachman, Lyle F. & Adrian S. Palmer 2010. *Language assessment in practice.* Oxford: Oxford University Press.

British Council, UKALTA, EALTA & ALTE. 2022. *Aligning language education with the CEFR: A handbook.* https://www.britishcouncil.org/sites/default/files/cefr_alignment_handbook_layout.pdf (accessed 28 November 2024).

Bunch, Michael B. 2012. Aligning curriculum, instruction, and assessment. https://www.measurementinc.com/sites/default/files/AligningCurriculumAssessmentandInstruction.pdf (accessed 10 Sept 2025).

Council of Europe. 2020. *The Common European Framework of Reference for Languages: Learning, teaching, assessment: Companion Volume.* https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (accessed 29 Nov 2024).

Council of Europe. 2001. *Common European Framework of Reference for Languages: Learning, teaching, assessment.* Cambridge: Cambridge University Press. https://rm.coe.int/1680459f97 (accessed 31 Oct 2025).

Ellis, Paul. 2024. Cultivating inclusion: Strategies for embracing diverse learners. Cambridge University Press and Assessment, Partnership for Education. https://www.cambridge.org/partnership/cultivating-inclusion-strategies-for-embracing-diverse-learners (accessed 25 Nov 2024).

Messick, Samuel. 1989. Validity. In Robert L. Linn (ed.), *Educational measurement*, 3rd edn, 13103. American Council on Education; Macmillan.

O'Sullivan, Barry. 2021. *The Comprehensive Learning System.* London: British Council. https://www.britishcouncil.org/sites/default/files/the_comprehensive_learning_system_new_layout.pdf (accessed 8 Sept 2025).

Plakans, Lia. 2020. Assessment of integrated skills. In Carol A. Chapelle (ed.), *The concise encyclopedia of applied linguistics*. Chichester: Wiley Blackwell. DOI: 10.1002/9781405198431.wbeal0046.pub2.

Race, Phil, Sally Brown & Brenda Smith. 2005. *500 Tips on assessment*, 2nd edn. London: Routledge.

Rahman, Md. Mehadi. 2018. Exploring science teachers' perception of classroom assessment in secondary schools of Bangladesh. *European Journal of Education Studies* 4(9). 139-160. DOI: 10.5281/zenodo.1296835.

Stiggins, Rick J. 2014. Improve assessment literacy outside of schools, too. *Phi Delta Kappan* 96(2). 67-72. DOI: 10.1177/0031721714553413.

Varaporn, Savika & Pragasit Sitthitikul. 2019. Effects of multimodal tasks on students' critical reading ability and perceptions. *Reading in a Foreign Language* 31(1). 81-108. https://files.eric.ed.gov/fulltext/EJ1212804.pdf (accessed 25 Nov 2024).

11 Biographies

Carolyn Westbrook is a test development researcher at the British Council. Formerly an associate professor in EFL, she is a senior fellow of the Higher Education Academy (now Advance HE) in the UK. She delivers assessment literacy training for teachers and lecturers around the world and has worked on a number of assessment development projects. Her research interests are the assessment of integrated skills, EAP, ESP and EMI. She has worked as a teacher and teacher trainer for over 30 years.

Aidan Holland works as Global Assessment Solutions Manager, Wider Europe. He has been working for the British Council in a variety of roles since 2009. He has over 20 years' international experience in teaching and teacher training and almost a decade of experience in academic management positions. He holds an MA in TEFL/SL from the University of Birmingham and an MA in Language Testing from the University of Lancaster.