

Designing an EAL Assessment and Testing Its Validity: the Case of AssessEP



Introduction and background

English language proficiency is undoubtedly one of the main predictors of academic success in schools with an English-based curriculum. As the need to assess learners with English as an Additional Language (EAL) in UK schools and internationally is rising, providing schools in the UK and internationally with the necessary resources to measure EAL learners' proficiency in English is vitally important. It helps not only to identify a learner's level of ability in English but also to inform subsequent lesson planning and target setting for learners.

To address the need for EAL learner assessment, the Across Cultures team has designed a robust assessment tool that assesses English language proficiency across all four language skills - listening, speaking, reading and writing. It approximately

aligns with well-known and widely used EAL frameworks, such as the BELL foundation, CEFR, ELLP and more. AssessEP was launched for its pilot implementation in June 2023 and was piloted in both UK national and international schools for a year until the end of July 2024. Meanwhile, it was gradually introduced to schools already using its sister programme, the Learning Village, an EAL blended learning programme designed for independent learning and small group teaching, as part of their membership in January 2024. This report outlines the process undertaken to design AssessEP and evaluates its validity as an assessment tool. It also includes details on moderating its levels and thresholds (i.e. pass marks) during its pilot implementation.

What is AssessEP?

AssessEP is a short and straightforward summative assessment tool that helps you accurately determine your EAL and multilingual learners' starting points in their English language proficiency journey and spot-check their progress in time. It assesses all four language skills (listening, reading, speaking and writing) preceded by an optional pre-diagnostic assessment that helps teachers determine whether low-level learners are capable of proceeding to the main assessment. The main assessment comprises two online and two offline modules as outlined below:

Short offline pre-diagnostic assessment
(speaking, listening, reading and writing)

Online reading assessment
⌚ (15-30 mins)

Online listening assessment
⌚ (15-30 mins)

Offline unaided writing assessment
⌚ (15-30 mins)

Offline speaking assessment
⌚ (10-15 mins)

AssessEP is an innovative assessment scheme that identifies learner levels and aligns them with its sister programme, Learning Village. It also approximately aligns with other EAL frameworks, such as the BELL foundation, CEFR and more on an international level. This makes it an assessment tool that can be used by teachers internationally, as it provides learner levels that are meaningful and align with widely recognised EAL frameworks that educators may already be familiar with.

How was AssessEP designed?

The assessment modules were designed through collaboration between our English language learning specialists based in the UK and New Zealand. These experts combined their knowledge to design reading and listening comprehension texts, along with multiple-choice questions, leveraging their extensive experience in English language

assessment. In addition, they designed writing and speaking questions with graded difficulty levels and detailed assessment criteria to help teachers benchmark their learners' performance. This enables educators to gain deeper insights into their students' language development and effectively support their learning journey.

The AssessEP reading/listening comprehension texts and multiple choice questions were designed based on the Common European Framework of Reference (CEFR) suggested topic areas and levels of language proficiency. Widely used across Europe and globally, the CEFR provides a common basis that defines levels of linguistic ability. It describes what language learners can do at each level, allowing progress to be measured and qualitatively defined across different stages of language development. The original reading and listening texts designed by the team were evaluated using two web-based tools: a) the Text Inspector, which provides over 200 metrics to evaluate the difficulty level of English texts using CEFR levels as a point of reference and b) the GSE teacher toolkit by Pearson, which allows you to level individual lexical items or whole texts against the CEFR levels. This process led to a number of revisions of the test content before the pilot implementation in schools.



The pilot implementation process

Prior to the official launch of AssessEP, the programme was used by schools on a pilot level that lasted twelve months.

During the pilot phase, the assessment was implemented in international schools, where there is a substantial need to assess the EAL learners' level of English. The aim of this pilot was to investigate educators' initial thoughts, views and perceptions of the assessment as well as potential challenges in the administration process. Semi-structured interviews were used as a research tool for this investigation.



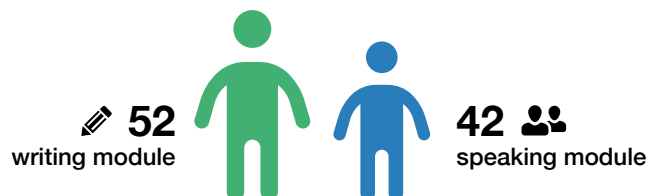
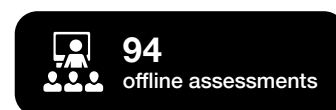
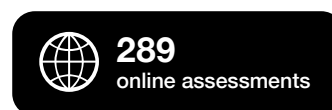
A group of 10 schools from different countries, all belonging to the International Schools Partnership (ISP), were introduced to AssessEP through short online meetings. Each school implemented AssessEP with 20 learner licenses. Following the initial implementation, 4 schools participated in semi-structured interviews with the AssessEP team. These interviews focussed on teachers' perspectives regarding the effectiveness of AssessEP as an assessment tool, the validity of results based on learners' subsequent classroom performance, and any challenges or issues encountered during the programme's implementation.

Following the pilot, the AssessEP team focussed on moderating the assessment thresholds and successfully determining the exact percentages that mark proficiency levels within this assessment. To achieve this, the programme was implemented at the Centre of Applied Linguistics and Language Studies at the Hellenic American Union in Athens, Greece. The participants were language learners aged 8-18 years old, who were already attending English language courses at this centre and had been accurately placed according to CEFR levels. To ensure the accuracy of learner levels, students at the Hellenic American Union undergo continuous assessment, including internal evaluations (e.g. diagnostic assessments upon entry), ongoing teacher judgements, and external examinations that

certify their English language proficiency - all aligned with CEFR levels.

To ensure that participants entered the pilot with sufficient contextual information and a legacy level to be compared with their AssessEP result, teachers provided our team with:

- **an indication of the CEFR level the learners are currently working at**
- **an indication of their CEFR level across all four language skills (reading, listening, speaking and writing)**
- **a teacher's judgement of whether each learner was high, medium or low performing within their level**
- **details on additional learning needs or extra information that may have affected the assessment result (e.g. a child may be too shy, bilingual etc.)**



191 learners were enrolled to sit the online modules of the assessment in controlled test conditions in the computer lab. Out of this wide sample, 52 learners also completed the writing assessment and 42 sat the speaking assessment module. This led to the administration of 289 online assessments and 94 offline assessments in total.

With research confirming that a proper sample size for exam threshold testing is 46 assessments (N= 46), the number of assessments administered as part of this research can be considered a safe sample which allows the generalisation of results for future administrations of the assessment (Mustapha & Robillos, 2020).

Outcomes of the pilot study

Teacher feedback

Upon completion of the study, an analysis of teacher feedback showed that teachers' views on AssessEP were clearly positive. They found it to be a valuable tool that provided accurate summative information on learner levels. Teachers also considered it to be straightforward and user-friendly. Bridget Bova, EAL coordinator at Mile School, in Italy, described the administration of AssessEP as

“A smooth experience with no issues. The online modules were straightforward and easy to navigate. The results were valid and matched learner performance across the board.”

At Mile School, teachers

“really liked the offline speaking assessment. It was image-based with a variety of images suitable for different cultures and able to activate prior schemata and knowledge effectively.”



Other teachers also acknowledged

“the importance of having access to online assessment tools which are and should be the future of English language proficiency assessment”

(Konstantina Karouni, EFL teacher, Hellenic American Union, Greece). Sutha Sunkaran, Secondary EAL Teacher and Head Of Multilingualism at Straits International School, reported that

“AssessEP was a user-friendly assessment with clear teacher instructions and all staff involved can access and share learner results easily. We used it in synergy with the BELL foundation. Thank you so much for creating such a wonderful platform. It has been very useful.”

Processing learner results

With regard to processing the volume of learner results at the end of the pilot implementation, common themes and patterns in learner performance were identified, which led to slight adjustments of some of the thresholds. While the majority of learners scored at the expected level based on prior teacher input, young learners seemed to consistently fare better on the pre-diagnostic and CEFR A1 Reading and Listening assessment. Raising the pass mark slightly on these assessments ensured that AssessEP results aligned more consistently with the actual level of learners. A similar adjustment was necessary in the B level thresholds. This process also led to adjustments to the content of the pre-diagnostic module of the assessment.

Our team of specialists also marked 52 writing scripts. Each sample was marked by two script raters, and in case of discrepancies, a third rater was brought in to ensure accuracy. This process led to the development of useful teacher training material and the refinement of our writing marking criteria, ensuring a clearer and more accurate representation of writing skill progression across English language proficiency levels. Finally, interesting ideas and suggestions were considered for incorporating more academically-orientated writing tasks into the assessment.

As a result of overseeing the assessment administration process and observing learner experience in action, the team decided to implement some technical changes to improve fairness and overall user experience. A back button was introduced, allowing learners to go back and edit their answers. Additionally, the option to pause the audio in the listening assessment was introduced, while the original audio slowdown feature was removed to ensure a fairer assessment - ensuring all learners experienced the same speech delivery speed when testing their listening skills. Finally, an option for learners to revisit the same assessment was introduced to accommodate unforeseen interruptions, such as a school fire drill.

Conclusion

In conclusion, the validity testing process for AssessEP has confirmed that it is a robust and valuable English language proficiency assessment tool.

Through rigorous analysis of the assessment design, threshold verification and ongoing teacher feedback from our schools, we have validated that AssessEP effectively measures language skills and supports educators in future planning, interventions and additional learner support. AssessEP will continue to be updated to reflect current best practices in language assessment. Our team is committed to conducting regular reviews to ensure it remains a reliable resource for teachers in the hugely important task of assessing their EAL learners.



References

Cummins, Jim. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire*. Buffalo, NY:Multilingual Matters Ltd.

Department for Education. (2024). *Schools, pupils, and their characteristics: Academic year 2023/24*. Explore Education Statistics. <https://explore-education-statistics.service.gov.uk/find-statistics/school-pupils-and-their-characteristics>

Mustapha, F. & Robillos, R.J, (2020). "Proper sample sizes for English language testing: A simple statistical analysis" *Humanities & Social Sciences Reviews* 8(4): 442-452. https://www.researchgate.net/publication/343650160_Proper_sample_sizes_for_English_language_testing_A_simple_statistical_analysis

Strand, S., & Demie, F. (2005). *English language acquisition and educational attainment at the end of primary school*. *Educational Studies*, 31(3), 275–291. <https://doi.org/10.1080/03055690500236613>

Click on the links below if you need more information on any of the following:

- **Learning Village** click [here](#)
- **CEFR** click [here](#)
- **Text Inspector** click [here](#)
- **GSE Teacher Toolkit** click [here](#)
- **ISP schools** click [here](#)
- **Centre of Applied Linguistics and Language Studies at the Hellenic American Union in Athens** click [here](#).

Acknowledgements

We appreciate and acknowledge the hard work and cooperation of all members of the AssessEP team, English language learning specialists and IT specialists, who contributed to the development of this assessment tool:

- i. **Caroline Scott** (BA, PGCE, DipM, NPQH, MA), EAL Author and Advisor, experienced Class Teacher, EAL Teacher, EFL Teacher & Head Teacher.
- ii. **Simon Lobo Morell**, Developer.
- iii. **Martha Giannakaki** (BA, MA), EAL specialist, class teacher, EAL teacher & linguist.
- iv. **Heidi Brunner** (BA), Content development lead, English language teacher & class teacher.
- v. **Breda Matthews** (MA B.Ed, Dip TESSOL), EAL specialist and experienced ELL, ESOL & EAP Teacher, Programme & Resource Creator, Consultant.
- vi. **Sally Hay** (BA, MA, PGrad Language Teaching, Mindlab, GradDip(Sec)) ESOL, ELL & English Teacher, Secondary Education Project Lead, Programme designer.
- vii. **Miranda Howell** (Ed. BA Hons, Trinity Dip TESOL, Mindlab), EAL specialist with years of experience as an ESOL, EAL, EAP Teacher & Curriculum Coordinator.
- viii. **Iva Miteva** (PGDip, MA, CertTESOL), EAL specialist with years of experience as an EAL teacher, linguist & Head of EAL.

We would like to express our sincere gratitude to Sarah Brown, Head of Learning First, and Ruth Radwan, Group Head of Multilingualism, Learning and Innovation Team for our collaboration with International Schools Partnership (ISP), a global community of 87 international schools across 24 countries, committed to using innovative and transformative approaches to learning. Special thanks go to the schools that participated in the AssessEP trial: Newton College, Mile School, Lynn-Rose School, Straits international school Penang, Asia Pacific International School, Straits International School, Colegio Pedro de Valdivia Agustinas, Star Mirdif, Tenby Penang, Boston School, whose insights and expertise were invaluable throughout the research process. More information on the ISP schools can be found [here](#).

We are also deeply grateful for our partnership with the Hellenic American Union in Athens, whose support was invaluable and made this study possible. Special thanks go to Bessy Geka, Deputy Director of the Centre for Applied Linguistics and Language Studies, Konstantina Karouni (EFL teacher & Academic Assistant of the English Language Programme) as well as Aggela Megariti and Kirki Psoni (members of the administration team) for consistently doing their absolute best in facilitating the pilot implementation of AssessEP with a large cohort of English language learners. Last but not least, we appreciate the time and cooperation of all the teachers, who participated in interviews and shared their experience with AssessEP and whose feedback greatly enhanced the quality of our work.