



Test Review: LanguageCert IESOL B1 (Achiever) SELT

William S. Pearson

To cite this article: William S. Pearson (2022): Test Review: LanguageCert IESOL B1 (Achiever) SELT, Language Assessment Quarterly, DOI: [10.1080/15434303.2022.2103420](https://doi.org/10.1080/15434303.2022.2103420)

To link to this article: <https://doi.org/10.1080/15434303.2022.2103420>



© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 28 Jul 2022.



Submit your article to this journal [↗](#)



Article views: 132




View related articles [↗](#)



View Crossmark data [↗](#)

Test Review: LanguageCert IESOL B1 (Achiever) SELT

William S. Pearson 

University of Exeter, St Luke's Campus, Exeter, UK

ABSTRACT

The present article reviews LanguageCert's International English for Speakers of Other Languages (IESOL) *Achiever* Secure English Language Test (SELT). This high-stakes on-demand CEFR-linked exam has, since 2020, constituted Home Office-recognised evidence L2 English users can speak, write, and understand written and verbal English at B1 level. Passing the test facilitates enrolment onto a foundation or pre-sessional English course at a UK higher education institution, although some institutions set higher standards. As a neophyte SELT, there have been few descriptions and evaluations of the test beyond a range of sponsored studies. The current review indicated the Achiever test measures candidates' general abilities to understand, interact, and produce tasks that mirror real life. However, a lack of 'academicness' and validity concerns in listening raise questions over its suitability for predicting readiness for tertiary study. The test offers the benefits of efficiency in registration and communicating results, remote proctoring and invigilation, and numerous sample materials in the public domain. The provision of an innovative re-sit option may prove favourable to candidates, although could encourage repeat test taking and attempts to pass by a narrow margin, rather than investments in language learning.

In December 2019, it was announced that LanguageCert had secured a contract to become an approved provider of Secure English Language Tests (SELTs), joining Pearson and the well-established consortia of Trinity College London and IELTS (International English Language Testing System). The decision to widen the range of SELT providers to include LanguageCert has profound ramifications for three key stakeholders. For the parent company, PeopleCert Ltd., the contract is potentially extremely lucrative, signalling its status as a highly trusted and respected partner that delivers language tests both in the UK and globally which are officially recognised and condoned by the Home Office for visa and immigration purposes. The move offers candidates greater choice, increasing the amount of competition in a hitherto small market. The third group, test-users, are required to ensure staff interpret scores appropriately in light of the purposes for which the test is used. While there is an emerging body of sponsored research that addresses LanguageCert's IESOL (International English for Speakers of Other Languages) SELTs (e.g., Coniam et al., 2021; Green, 2019; Papargyris & Lampropoulou, 2020), these appear to be targeted at language assessment specialists. Perhaps owing to the tests' novelty as well as upheavals in language testing due to the Covid-19 pandemic, limited external attention has been paid to LanguageCert

CONTACT William S. Pearson  wsp202@exeter.ac.uk  University of Exeter, St Luke's Campus, Heavitree Road, Exeter EX1 2LU, UK

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

(see, Howell et al., 2020; Isbell & Kremmel, 2020). This review describes and evaluates the LanguageCert IESOL B1 (*Achiever*) SELT, taken for the purposes of demonstrating sufficient language readiness for English-medium academic study, and is targeted towards test-users (academic and admissions staff) and scholars unfamiliar with LanguageCert.

Test purpose and use

LanguageCert's IESOL SELTs are a suite of English language qualifications offered across 175 test centres globally (as of October 2021), mapped onto the Common European Framework of References for Languages (CEFR). Unlike language tests that assess the full spectrum of English language proficiency (e.g., IELTS, TOEFL), candidates undertake a test pegged to one of the six CEFR levels, A1-C2. The exams were acquired by LanguageCert from City & Guilds in 2015 and revised after an internal quality review (Green, 2019). Rolled out in April 2020, LanguageCert's IESOL SELTs aim to provide Home Office-recognised qualifications for non-native English-speaking visa applicants who need to prove their ability to speak, write, and understand spoken and written English to a required level (LanguageCert, 2020a). The *Achiever* test is of particular importance since demonstrating B1 in a SELT is a legal requirement in order for most L2 users to obtain UK permanent residency, naturalisation, skilled worker status, and for the purposes of the present study, enrolment onto a tertiary-level course via a sub-degree programme (e.g., pre-sessional or foundation course).

Test description

The LanguageCert IESOL B1 SELT consists of two discrete modules (termed the *spoken* and *written components*), which are administered separately at a secure test centre. The two components can be booked for the same day in some centres and must be undertaken within 14 days of each other. The spoken component comprises a 12-minute online interview remotely proctored by a live interlocutor. Discrete sections on listening, reading, and writing comprise the written component, for which candidates are permitted two hours and 40 minutes, of which 30 minutes is assigned to listening. There are no time limits for reading or writing, affording candidates greater flexibility to allocate time to the tasks as they wish. The written component is administered as a computer-based test. A maximum of 26 raw marks are achievable in both listening and reading, with 12 for writing and 12 for speaking. The marks from each of the four skills are scaled to a score out of 50. An independent evaluation and CEFR referencing exercise undertaken in 2018 by NARIC (National Recognition Information Centre, now the UK National Information Centre for the Recognition and Evaluation of International Qualifications, or UK ENIC) determined that, in order to claim a link to the CEFR at level B1, the cut score for a passing grade in each of the four skills should be set at 33. The method for scaling raw marks has not been publicly disclosed by LanguageCert. This may result in uncertainty for test-takers who wish to know the precise raw mark required to achieve a pass. The four sub-tests comprising LanguageCert IESOL B1 are henceforth described and outlined.

Listening

The main characteristics of the four-part listening section are summarised in [Table 1](#). Single-answer, three-option multiple choice questions (MCQs) feature in parts 1–2 and 4, although task demands vary. In part 1, candidates identify the appropriate response to complete a series of brief formal or informal dialogues, assessing learners' socio-pragmatic knowledge of speech acts and routines (Roever, 2006). In parts 2 and 4, the appropriate response that answers a question or completes a statement must be selected, testing candidates' skills of connecting and synthesising information to identify the appropriate topic, purpose, context, speakers, relationship between speakers, roles, attitudes, etc. in response to three 130-word conversations involving two participants (part 2) or a 500-word multi-speaker discussion (part 4). Part 3 stands out as the only non-multiple-choice task, assessing listeners' comprehension of main and supporting points by noting down specific information from a broadcast, narrative, or presentation onto a 'message pad' that features short prompts. All spoken texts are scripted. Analysis of four sample papers reveals a speech rate of 95 words per minute, factoring in pauses between parts and repeat listening.

Reading

As with listening, the reading section features 26 items, spread across a diverse range of passages of varying lengths ([Table 2](#)). Parts 1 and 3 encompass multiple, short extracts from an email, article, or advertisement, while 2 and 4 contain a continuous text of a narrative, descriptive, expository, biographical, or instructive nature. Part 2 and 4 passages are aimed at the general reader, often seeking to enlighten on a possibly unfamiliar person, place, or phenomenon. In part 1, test-takers select a word or short phrase from a multiple-choice selection (featuring two distractors) to fill a gap, testing their knowledge of lexis, cohesive devices, coherence, and textual layout (e.g., of addresses, signs, appointment cards, etc). Part 2 is undertaken by matching the appropriate sentence from a selection of six (with one distractor), employing candidates' knowledge of cohesion, coherence, and sentential functions. Test-takers scan four passages in part 3 to identify which fulfil certain communicative purposes and contain information that answers various questions. Finally, part 4 comprises a series of open-ended questions that are answered by scanning the passage for the specific information given (using a maximum of three words and paying attention to spelling in the written response).

Writing

The writing section comprises two communicative tasks. Part 1 requires a 70–100-word formal response to an input (e.g., a letter, poster, diary, timetable, or leaflet) that addresses three *content points* and is suitable for a public audience. Rhetorical requirements are not directly specified, varying noticeably across sample materials (e.g., making planning suggestions to local government, applying to participate in community programmes, entering competitions). Part 2 is a less formal, albeit longer task. Candidates are tasked with writing a 100–120-word letter on a topic of personal interest. The provision of sample candidate responses accompanied by brief examiner

Table 1. Task characteristics of the listening section.

	Part 1	Part 2	Part 3	Part 4
Times played	2	2	2	2
No. of items	7	6	7	6
Item type	MCOs	MCOs	Note completion	MCOs
Task requirements	Listen to 7 sentences and choose the correct reply to the speaker	Listen to 3 conversations and choose the correct response	Listen and note down specific information (in one to three words) onto a message pad	Listen to a discussion and choose the correct response
Combined transcript lengths in words*	160	400	300	500
No. of speakers	2	2	1	Multiple

*Calculated as an average across four papers.

Table 2. Task characteristics of the reading section.

	Part 1	Part 2	Part 3	Part 4
No. of items	5	6	7	8
Item type	MCOs	Matching	Matching	Short-answer questions
Task requirements	Complete gapped texts by selecting the appropriate word or short phrase	Complete a gapped text by selecting the appropriate sentence from a list	Match the passage which contains the answer to the given questions	Answer open-ended questions of one to three words
No. of passages	5	1	4	1
Combined passage lengths in words*	120	210	260	330

*Calculated as an average across four papers.

comments (LanguageCert, 2018) allow stakeholders to become more familiar with the rhetorical characteristics of the tasks, although performance outcomes are rarely justified with reference to textual features.

The tasks are assessed analytically by trained and certified marking examiners against four equally weighted criterion-referenced descriptions of performance (*task fulfilment, vocabulary, grammar, and organisation*). Task fulfilment requires candidates to ‘achieve communication’ and attend to the stated content points. The omission of content points may be penalised, with off-topic responses not receiving any credit. Organisation denotes coherence at the discourse level as well as surface-level cohesive devices, linking words, and punctuation. Vocabulary and grammar are assessed according to the range, frequency, and effect of errors with regard to structures typical at B1, outlined in the testing handbook (see, LanguageCert, 2020a). For each criterion, marks from 0–3 are awarded, with the four levels denoting a generalised conception of writing ability (e.g., “*meaning usually clear despite a more limited range of vocabulary and/or spelling errors*”, Vocabulary Accuracy and Range, Level 2).

Speaking

The speaking section contains four three-minute tasks. Part 1 is the only ‘independent’ section of the test, i.e., that assesses candidates’ ability to speak about familiar topics (Alderson, 2009). After asking test-takers to state and spell their name and give their country of origin, examiners select and ask one question from five frames (e.g., public transport) and only make brief responses or comments. Part 2 consists of two or three role play dialogues where the candidate employs their pragmalinguistic knowledge (Roever, 2006) to respond appropriately to the examiner’s questions or declarative statements and initiate interactions for a few turns, for instance: “*We’re at a science museum. I start. We’ve got about an hour. What shall we look at first?*”. A task sheet is provided in part 3, which features visual inputs aimed to facilitate a discussion evincing test-takers’ competence to make a plan, arrange, or decide something (e.g., transport and accommodation for a holiday). Part 4, the only test task where candidates are explicitly allowed to take notes, assesses their ability to sustain a 90-second monologue (after 30 seconds of preparation time) on a personal topic selected by the examiner (from a choice of three). Once the time has elapsed, the examiner asks some follow-up questions from a pre-written list.

The rating of the spoken component is carried out post hoc by an examiner using a recording of the test, with a criterion-referenced score between 0–3 awarded across four dimensions (as in writing). *Task fulfilment* and *coherence* aim to measure a candidate’s ability to manage the tasks adequately for the level and link utterances into coherent speech. The criterion *Pronunciation, intonation, and fluency* encompasses an amalgamation of three equally-weighted features that targets the individual’s competence to produce the sounds of English in order to be understood with appropriate stress and intonation and maintain the flow of speech. As with the writing section, test-takers are assessed on their range and accuracy of B1-level vocabulary and grammar.

Test quality

LanguageCert's B1 SELT was evaluated in the aspects of validity, reliability, test bias and fairness, practicality and access, and consequences, following the frameworks developed by Bachman and Palmer (1996) and Kunnan (2004), widely used among language assessment specialists.

Validity

Central to the fitness for purpose of a language test is its validity (Bachman & Palmer, 1996; Papargyris & Lampropoulou, 2020), incorporating the qualities of content representativeness and construct validity (Kunnan, 2004). There have been three independent validation studies of LanguageCert's IESOL SELTs (Green, 2019; National Recognition Information Centre, 2021; O'Sullivan, 2008). Green (2019) and O'Sullivan (2008) undertook an assessment of the B2 (*Communicator*) test, utilising the staged approach (comprising familiarisation, specification, standardisation, benchmarking/standard setting and validation) recommended in the Council of Europe's (2009) manual for linking tests to the CEFR. The assessment by the National Recognition Information Centre (2021), commissioned in 2018, was a pre-requisite for LanguageCert's eligibility for SELT status. NARIC determined CEFR alignment was evident at all stages of test development and delivery, although the methodology used is not stated in the executive summary, the only section of the report available to the public.

A validity argument is made that LanguageCert's IESOL SELTs assess a communicative conception of language ability by sampling users' reception, interaction, and production in everyday real life settings (Papargyris & Lampropoulou, 2020). The test directly links to the CEFR in determining what a test-taker is expected to be able to do in real world language use at B1 level, elaborated in LanguageCert's handbooks (LanguageCert, 2020a, 2020b). As such, in comparison to other SELTs, the test draws heavily on communicative/functional theories of language proficiency. There are also prominent structural underpinnings, since candidates are assessed on their use of B1 lexicogrammatical forms elucidated in the handbooks (LanguageCert, 2020a, 2020b). The breadth of functions and structures outlined may offer the benefit of dissuading candidates from focusing on a narrow range of rhetorical forms in an effort to develop test management strategies, while the explication of expected target forms provides clarity. Additionally, tailoring the test to the language abilities expected only at B1 simplifies its design and potentially reduces the prospect of measurement errors.

LanguageCert claims the Achiever test is, "suitable for candidates who are preparing for entry to higher education" (LanguageCert, 2020a, p. 3), yet there is a notable gap between the non-academic test topics, language, and task requirements and the academic target language use (TLU) domain, problematic for construct validity (Bachman & Palmer, 1996). Listening tasks typically comprise transactional talk in everyday settings (e.g., buying a ticket at the railway station), talks on general topics (a radio competition), and informal conversations (a couple discussing moving to a different house). Similarly, the topics of passages in the reading section are commonly pitched towards the general interest reader (e.g., a review of *March of the Penguins*). The written tasks are characterised by an absence of reading into writing, a mismatch in genre expectations and rhetorical requirements, and

the retrieval of only a limited sample of writing, while the topics of the spoken component are highly personalised (e.g., a time when you bought a special present, your favourite restaurant). As such, scores in the Achiever test should be considered as indicators of pre-study ability or language learning aptitude (Davies, 2008). Nevertheless, institutions' legal obligations to ensure accepted forms of English language proficiency (ELP) evidence clearly align with the CEFR – with potentially disastrous consequences on universities' lucrative right to recruit international students should they fall foul of Home Office regulations (Pearson, 2021) – may undercut the importance of the 'academicness' of an ELP admission test, potentially enhancing the attractiveness of Achiever.

There are a number of validity concerns associated with the Achiever listening test. Candidates are allowed to hear the recordings twice, which in spite of recent moves towards uploading recordings of lectures and seminars online, is rarely representative of ad hoc interactive encounters with other students or academic staff. Additionally, some candidates may move onto other sections of the written component after the first playback, potentially disturbing test-takers still listening. Three-option MCQs constitute 19 of the 26 items, although task demands vary (especially in part 2). While prevalent across high-stakes ELP assessments, their preponderance in the Achiever test could introduce invalidity through the increased reading load associated with gaining familiarity with the available options, test-takers focusing disproportionately on choices higher up the screen (i.e., option 'A'; Holzknicht et al., 2021), and the unknowable influence of informed or uninformed guesswork (Currie & Chiramanee, 2010; Yi'an, 1998). The preponderance of MCQs may have a harmful effect on learning by encouraging candidates to rehearse test management strategies to address such items (Alderson, 2009), rather than the expected competencies outlined in the CEFR. Finally, sample materials indicate a lack of visual clues that could help listeners picture the settings, number of speakers, and their roles (Alderson, 2009), potentially lowering anxiety and preventing confusion.

Reliability

Although a relatively new test, evidence attests to the reliability of listening and reading. Coniam and Lampropoulou (2020) employed classical test statistics (Kuder-Richardson Formula 20 [KR20], standard error of measurement [SEM], standard deviation) to investigate the consistency of 62 IESOL A1-C2 tests (it is not disclosed how many were B1). At B1, the authors uncovered overall KR-20 values of .86-.94 (.76-.89 for listening and .80-.90 for reading). Standard errors of measurement were 5–8%, indicative of test takers' 'true scores' occurring within a range of plus or minus two to four marks. Error variances of different examinees, calculated using a person-specific or conditional SEM, are yet to be reported and could provide insights into level-specific accuracy in light of the cut score. Higher than expected standard deviations were not considered a concern since, in theory, the Achiever test is open to any applicant. Coniam and Papargyris (2021) investigated examiner consistency in the writing component using Many-Facet Rasch Measurement (MFRM). Following a period of induction and training, 24 highly experienced examiners were asked to rate six Achiever scripts (along with six from the other IESOL tests), with opportunities for score confirmation and discussion as the cohort progressed through each level (starting at A1). Examiner fit to the model was considered

generally good, demonstrating raters marked consistently (albeit with possible severity or leniency), with only one examiner showing underfit. There has yet to be such an analysis of the speaking test.

Since LanguageCert IESOL B1 SELT is a computer-delivered test, the accuracy of listening and reading marking is enhanced through automatic scoring and elimination of the issue of illegibility (relevant for listening part 3 and reading part 4). Additionally, the post-hoc rating of candidates' speaking may improve the accuracy of scoring since examiners do not need to juggle assessing candidates' skills in conjunction with ensuring procedures are correctly followed. To address concerns over inaccuracies in the single marking of speaking and writing, "rigorous and frequent training and moderation" of examiners is emphasised (LanguageCert, 2020a, p. 99). Training encompasses marking sample tests, requiring would-be examiners to demonstrate they can consistently mark to standard before being certified (Coniam & Papargyris, 2021). During live writing marking, inaccurate or inconsistent rating may result in removal from the session (Coniam & Papargyris, 2021), suggesting examiner accuracy is continuously measured against pre-moderated seed tasks that resemble authentic scripts, inserted periodically into the marking session. Further measures include additional training or even dismissal (Coniam & Papargyris, 2021). Since such measures cannot prevent mistakes from being made, the test owners may wish to explore the introduction of double marking in speaking and writing.

Bias and fairness

As LanguageCert IESOL SELTs are designed to be undertaken by L2 English users from diverse backgrounds (e.g., gender, race, ethnicity, socioeconomic status, religion), the test's designers indicate that they have gone to trouble to ensure items, tasks, and texts do not exhibit bias against test-takers (LanguageCert, 2020a). Informational and literary passages in the reading section feature topics that are broad and neutral, potentially answerable by any demographic. Authentic sample tests (of which eight are currently available online) are notably absent of references to gender, race, ethnicity, and sexual orientation, limiting the risk of stereotyping sub-groups. The materials feature a diverse (although slightly Eurocentric) representation of cultures, reflected in the names and accents of speakers in the listening section and biographies of popular figures in reading. The material appears age appropriate for test-takers under the age of 16, who require the authority of a parent or guardian when booking (Isbell & Kremmel, 2020). Item bias in the listening and reading components across four key variables – gender, age (decade of birth), L1, and test centre – was investigated by Coniam and Lee (2021) using Differential Item Functioning (DIF) analysis. Moderate-to-large DIFs (a magnitude of >0.64 logits and a significance value of $p < .05$), cause for flagging up items for review, were reported for 30% of tests based on L1 (across a sample of 40 papers) and 23.8% for test centre (185). These were attributed to the scattering of different first languages across the large number of test centres and low sample sizes (partly due to omissions in the data). For gender, incidences of moderate-to-large DIFs were very low (6.7%), albeit a sample of only 15 tests was analysed.

Practicality and access

Home Office requirements compel LanguageCert to emphasise security over practicality. Notable measures include photographic ID checks to prevent impersonation, restrictions on bringing personal items into the test centre, a clear code of conduct, and continuous surveillance of test taking in the form of video and audio recording, physical walkthroughs, and observation windows (LanguageCert, 2020c). As a consequence of the costs involved in implementing such measures, LanguageCert SELT fees are relatively high (£200 for UK-based candidates, £169-£176 internationally). Clearly, as not all candidates can afford this, the test does not meet Kunnan's (2004) criterion of fair financial access. Nevertheless, as overseas tertiary education incurs far higher costs, test-takers' complaints are generally not afforded much credence.

As a SELT, LanguageCert requires travel to a secure test centre, which may involve time-consuming and costly journeys or accommodation for individuals not located in urban areas (where test centres are typically located). Furthermore, the processes of registration and test taking are carried out exclusively through a computer-based medium, assuming a certain amount of digital literacy. The possible differential effects on outcomes for candidates with limited digital expertise or familiarity with word processing written English are important future lines of inquiry. LanguageCert emphasises speed and flexibility in registration and administration. Candidates may take the Achiever test within four hours of registration (Isbell & Kremmel, 2020) and can re-arrange a pre-booked test for free up to 72 hours prior to the test. Results are relayed to candidates in just five business days, which compares favourably with other SELTs.

A novel feature of LanguageCert is the safety net of one re-sit in the event of unsuccessful outcomes, which the company provides, 'at a fraction of the exam price', known as Take². Take² allows test-takers to re-sit either the spoken or written component within 14 days. Sensibly, it is only available to borderline unsuccessful candidates, defined as those who achieved a minimum of 25/50 in the spoken component or an accumulated 75/150 in the written component, with the further requirement of 33/50 in either the writing or speaking sections. Test-takers must commit to Take² during initial registration, drawing on their level of certainty over whether they will meet the pass mark. Candidates are automatically notified of their eligibility for Take² upon receipt of their results. There appears to be no restriction on the number of times candidates can book a test with the pre-paid Take² option, which could encourage repeat test taking and attempts to pass by a narrow margin instead of investments in ELP. LanguageCert outlines that it is the responsibility of test centres to, "offer candidates the most appropriate LanguageCert International ESOL qualification", which depends on "each candidate's ability and needs" (LanguageCert, 2020a, p. 6). However, the selection of an appropriate test is determined by the type of visa required, along with the requirements publicised by the higher education institution (who may demand a pass in the Communicator test).

Consequences

Providing it can maintain its SELT status in the longer term, LanguageCert has the potential to exert a profound impact on the language testing landscape in the UK. Its B1 SELT offers prospective overseas students an alternative to IELTS and Trinity College London, who have long cornered the domestic and overseas SELT market. Achiever may appear more appealing

to candidates owing to the greater flexibility afforded in booking the test, the speed in which results are delivered, the prospect of undertaking a test targeted at their particular level instead of a more intimidating full spectrum test, and the option of pre-paying for the safety net of a re-sit. As a test of general communicative ability, Achiever may induce positive washback by encouraging students to focus on developing general communication skills, albeit perhaps at the expense of academic language. It may be the case that the test is perceived as a softer option by candidates unable to achieve test-users' requirements in a more academic ELP test, such as IELTS or the Pearson Test of English Academic. This could have longer-term repercussions for how admissions personnel come to view the test.

Conclusion

LanguageCert's IESOL B1 SELT offers increased competition in the provision of high-stakes ELP testing. While the test offers advantages to candidates in the form of a narrow B1 language focus, efficient results processing, and its innovative re-sit policy, it is a noticeably less academic option in comparison to other SELTs. As entrance standards for many sub-degree programmes are often only articulated in IELTS band scores, LanguageCert faces challenges in wresting market share away from its more established competitor. Since the test is pegged to the CEFR to meet Home Office visa requirements, there is a mismatch between language use in the test and the TLU domain in the academy. Therefore, while the test helpfully reveals what a candidate may be expected to do in real world language use at B1 level, it is crucial for staff involved in minimum language standard setting to interpret test outcomes cautiously, ideally as indicators of pre-study ability or language learning aptitude (Davies, 2008). If an institution is going to be making decisions about students' academic language capabilities (e.g., for in-session EAP support), supplementary in-house measures are necessary. Further studies are required to establish the reliability of the assessment of the spoken component (using for example, MFRM), as well as investigations into potential bias across key variables featuring larger sample sizes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

William S. Pearson  <http://orcid.org/0000-0003-0768-8461>

References

- Alderson, J. C. (2009). Test review: Test of English as a Foreign LanguageTM: Internet-based Test (TOEFL iBT[®]). *Language Testing*, 26(4), 621–631. <https://doi.org/10.1177/0265532209346371>
- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford University Press.
- Coniam, D., & Lampropoulou, L. (2020). *A review of LanguageCert IESOL listening and reading test reliabilities 2018-2020*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/coniam-d-lampropoulou-l-2020-review-of-languagecert-iesol-listening-and-reading-test-reliabilities.ashx>

- Coniam, D., & Papargyris, Y. (2021). *Examiner quality and consistency across LanguageCert Writing tests*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/coniam-d-papargyris-y-2021-examiner-quality-and-consistency-across-languagecert-writing-tests.ashx>
- Coniam, D., & Lee, T. (2021). *Potential bias in LanguageCert IESOL items: A differential item functioning analysis*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/coniam-d-lee-t-2021-potential-bias-in-languagecert-iesol-items.ashx>
- Coniam, D., Lee, T., Milanovic, M., & Pike, N. (2021). *Validating the LanguageCert test of English scale: The paper-based tests*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/coniam-d-lee-t-milanovic-m-pike-n-2021-validating-the-languagecert-test-of-english-scale-the-paperb.ashx>
- Council of Europe. (2009). *Relating language examinations to the common European framework of reference for languages: Learning, teaching, assessment (CEFR): A manual*. <http://www.coe.int/t/DG4/linguistic/Source/ManualRevision-proofread-FINAL.pdf%5Cnpapers3://publication/uuid/B6F2107B-A069-4C66-8DF7-572172E80D49>
- Currie, M., & Chiramanee, T. (2010). The effect of the multiple-choice item format on the measurement of knowledge of language structure. *Language Testing*, 27(4), 471–491. <https://doi.org/10.1177/0265532209356790>
- Davies, A. (2008). Assessing academic English language proficiency: 40+ years of U.K. language tests. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner, & C. Doe (Eds.), *Language testing reconsidered* (pp. 73–86). University of Ottawa Press.
- Green, A. (2019). *Relating LanguageCert Communicator to the CEFR*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/green-a-2019-relating-languagecert-communicator-to-the-cefr.ashx>
- Holzknicht, F., McCray, G., Eberharter, K., Kremmel, B., Zehentner, M., Spiby, R., & Dunlea, J. (2021). The effect of response order on candidate viewing behaviour and item difficulty in a multiple-choice listening test. *Language Testing*, 38(1), 41–61. <https://doi.org/10.1177/0265532220917316>
- Howell, B., Nation, P., Schmitt, D., Sinclair, C., Spencer, J., & Wigglesworth, J. (2020). *BALEAP guidelines on English language tests for university entry*. BALEAP. <https://www.baleap.org/wp-content/uploads/2020/08/AUG-BALEAP-Guidelines-on-English-Language-Tests-for-University-Entry-August.pdf>
- Isbell, D. R., & Kremmel, B. (2020). Test review: Current options in at-home language proficiency tests for making high-stakes decisions. *Language Testing*, 37(4), 600–619. <https://doi.org/10.1177/0265532220943483>
- Kunnan, A. J. (2004). Test fairness. In M. Milanovic & C. Weir (Eds.), *European language testing in a global context* (pp. 27–48). Cambridge University Press.
- LanguageCert. (2018). *International English for Speakers of Other Languages: Assessing writing performance - Sample scripts*. PeopleCert Qualifications Ltd. https://www.languagecert.gr/en/Documents/Assessing_Writing_Performance.pdf
- LanguageCert. (2020a). *International English for Speakers of Other Languages (Listening, Reading, Writing) qualification handbook*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/languagecert/product-information/preparation-material/english/languagecert-selt/selt-b2-lrws-2716/written-exam/qualification-handbook-iesol-listening-reading-writingver037.ashx>
- LanguageCert. (2020b). *International English for Speakers of Other Languages (Speaking) qualification handbook*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/languagecert/product-information/preparation-material/english/languagecert-selt/selt-b1-lrws-2714/spoken-exam/qualification-handbook-iesol-speakingver036.ashx>
- LanguageCert. (2020c). *LanguageCert International ESOL SELT - Exam instructions*. PeopleCert Qualifications Ltd.
- National Recognition Information Centre. (2021). *LanguageCert ESOL International qualifications: Independent CEFR referencing*. NARIC. <https://www.languagecert.org/-/media/feature/documents/research-team/publications-2021/naric-2019-languagecert-esol-international-qualifications-updated-independent-cefr-report-feb2019.ashx>

- O'Sullivan, B. (2008). City & Guilds Communicator IESOL Examination (B2) CEFR linking project: Case study report. City & Guilds. <http://docplayer.net/34526738-City-guilds-communicator-iesol-examination-b2-cefr-linking-project.html>
- Papargyris, Y., & Lampropoulou, L. (2020). *External validation of LanguageCert's English language examinations*. PeopleCert Qualifications Ltd. <https://www.languagecert.org/-/media/feature/documents/validation/external-validation-of-languagecert-esol-exams.pdf>
- Pearson, W. S. (2021). Policies on minimum English language requirements in UK higher education, 1989-2021. *Journal of Further and Higher Education*, 45(9), 1240–1252. <https://doi.org/10.1080/0309877X.2021.1945556>
- Roever, C. (2006). Validation of a web-based test of ESL pragmalinguistics. *Language Testing*, 23(2), 229–256. <https://doi.org/10.1191/0265532206lt329oa>
- Yi'an, W. (1998). What do tests of listening comprehension test? - A retrospection study of EFL test-takers performing a multiple-choice task. *Language Testing*, 15(1), 21–44. <https://doi.org/10.1177/026553229801500102>