



Domesticating TOEFL: Developing an English as a Foreign Language Assessment with Indonesian and Islamic Media Contexts

Ryan Marina^{1*}, Dimas Prasetya²

^{1,2}UIN K.H. Abdurrahman Wahid Pekalongan, Indonesia

E-mail: ryan.marina@uingusdur.ac.id¹, dimas.prasetya@uingusdur.ac.id²

*Corresponding Author: ryan.marina@uingusdur.ac.id

Abstract. Standardized English proficiency examinations such as the TOEFL often contain western cultural bias, frequently causing cognitive schema mismatch for students from Indonesian Islamic Higher Education Institutions. To address this challenge, the current research conducted the development and systematic validation of a domesticated TOEFL-like exam using the ADDIE model at UIN K.H. Abdurrahman Wahid Pekalongan. Grounded in Schema Theory, Language Assessment Literacy, and Culturally Responsive Pedagogy, this study executed a comprehensive domestication process, adapting American-centric contexts into familiar Indonesian socio-ecological, historical, and Islamic contexts. The final product of this development research is a complete 140-item blueprint of a TOEFL-like exam, consisting of 50 Listening items, 40 Structure and Written Expression items, and 50 Reading Comprehension items. Results from expert validation and schema mitigation evaluations demonstrate that utilizing culturally familiar schemas successfully eliminates construct-irrelevant interference. Consequently, this adaptation enables Islamic university students to accurately demonstrate and measure their true English language proficiency without encountering unnecessary cultural obstacles.

Keywords: ADDIE Model; Culturally Responsive Pedagogy; Islamic Higher Education; Language Assessment Literacy; Schema Theory; TOEFL Domestication.

1. INTRODUCTION

Students in Indonesian Islamic higher education frequently experience a schema mismatch as a result of the extensive use of Western-centric standardized tests like the Test of English as a Foreign Language (TOEFL). Significant difficulties in understanding the content and drawing conclusions occur when a test-taker's prior knowledge from their particular cultural and religious background clashes with the assessment materials' content (Wahyuningsih, Mahbub, and Al Fauzi, 2024). The TOEFL's highly academic and US-centered register is often difficult for students in Islamic university settings, resulting in lower scores that indicate cultural unfamiliarity rather than a lack of true English proficiency (Namkung, 2022).

Moreover, there is an urgent need to develop context-responsive assessment systems that take local identity into account while at the same time being empirically valid. The localized materials that are familiar to students in terms of cultural context can greatly improve students' learning outcomes and reading comprehension (Sari & Safri, 2018). By incorporating local role models like Ki Hajar Dewantara instead of Benjamin Franklin, teachers can extend the Zone of Proximal Development to reach students and help them understand complex English used in academic contexts in a familiar and relatable way. This resonates with the advocacy that argues for a paradigm shift away from monocentric and native-speaker-based assessment towards more diverse and localized assessment (Moussa O. Minta, 2024).

The development of a locally contextualized TOEFL-like test fits well with the Language Assessment Literacy (LAL) model. This model focuses on the need for teachers to understand the principles of assessment in their local educational context, which requires localizing the test (Kremmel & Harding, 2020). Through the development of a test like TOEFL with the context of Indonesia and Islam at the UIN K.H. Abdurrahman Wahid Pekalongan, this research has been able to provide an important structure for connecting the English proficiency standards in the international community with the sociocultural identity of Islamic university graduates.

Test of English as a Foreign Language (TOEFL) is common and has been conducted in over 200 nations across the globe. In spite of being available for many decades now, this test remains difficult for those who do not speak English, particularly students at universities in Indonesia. Scoring high in the Test of English as a Foreign Language is usually essential for admission and graduation. It has become a requirement even for some non-English speaking countries, such as Indonesia, to be able to join an international program, as a requirement for graduation, or for certain departments such as English education (Kaniadewi & Asyifa, 2022).

While being considered a standardized test for the measurement of proficiency in English language, the TOEFL test may become quite problematic for those non-native English speakers whose native languages include no references to the USA and its culture, history, and geography. In fact, it is expected that those students who speak any other languages will use academic discourse in its developed form. This includes understanding of complex lectures and class discussions in the TOEFL test. Reading materials, listening excerpts, and even speaking questions may contain terms that refer to the history of the USA, geography of the USA, society in the USA, USA idioms and slang, and USA weather conditions and cultural icons. This may be detrimental to test-takers who are not accustomed to American culture, despite their fluency in the English language (Namkung, 2022). Test-takers may fail the test not because of their poor English fluency, but due to their unfamiliarity with the cultural content embedded in the test.

The other issue is that the linguistic characteristics of the reading and listening texts employed in the TOEFL are extremely academic in nature. This means that there are specific linguistic characteristics such as vocabulary, grammar, etc., which are not usually employed in the day-to-day use of the language or even in the general teaching of the English language. Even though the student may have proficiency in English generally, they might lack knowledge about academic English. The literature has consistently emphasized the lack of vocabulary

skills among the students as a major factor in their challenges in all sections of the test, especially in reading and listening (Akmal et al., 2020).

It has, therefore, made me appreciate the agreement that exists among experts within the field of testing in English language that the application of the local context will yield a valid, fair, and effective test for non-native speakers. Although standardized tests such as TOEFL have their place in the global market (for example, when admitting foreign students into universities), such tests rarely take into account the unique language situations of individuals in a specific local setting. The theory behind the use of culture-based local materials is that they would lead to improved learning outcomes, and therefore the assessment would be effective (Sari, I. H., & Safri, R., 2018).

This research will be about solving actual issues faced by thousands of students annually at our university who want to pass the English assessment in TOEFL format. It is hard enough to make our students take this test only because of the English language. In addition, they still need to endure the test's USA-oriented content as well as its cultural biases. This research will highlight the importance of the localization of educational materials and suggest that tests based on such ideas will work better too.

The result of this research will be the set of tests based on domestic context in the TOEFL format, containing 50 listening skill questions, 40 structure and written expression questions, and 50 reading comprehension questions that can be used to test our students at the university. Current TOEFL test is very much USA-oriented; therefore, there are cultural biases. For instance, reading texts will be about the historical figures or locations in the U.S. They are more difficult to pronounce and even to hear.

That is why I decided to create a test similar to TOEFL but with an Indonesian and Islamic perspective, specifically designed for students at State Islamic University. I strongly believe that learning a text in context allows the reader to understand the text in its entirety. In this regard, it would be much easier for our students to understand the text on Soekarno-Hatta than the text on George Washington. I firmly believe that every text is a product of the culture within which it was written. Language Assessment Literacy underscores the importance of teachers' understanding of assessment and principles within their specific educational context, which often necessitates or benefits from localized approaches (Kremmel, B., & Harding, L., 2020).

2. RESEARCH METHOD

In order to fulfill the research aims and the significance of the research, I adopted the research and development strategy with the implementation of the ADDIE model. The steps are: First, Analysis, which includes needs analysis, constraints of the learners, and observation of the existing USA-centered TOEFL assessment in the university's language unit. It is significant since collecting all information from different sources helps to find out about the present state of knowledge, knowledge gaps, and previous researches. Second, Design, in which I made up the types of items for all three skills tested in the TOEFL exam (listening comprehension, structure and written expression, and reading comprehension). Third, Development, which is about making up the draft of scripts, questions, and answers. Fourth, Implementation, which means creating a whole set of TOEFL-like tests with the Indonesian and Islamic context, which includes 50 questions of listening comprehension, 40 questions of structure and written expression, and 50 questions of reading comprehension. Fifth, Evaluation, which means evaluating the product by the criteria (e.g., design principles, learning objectives, performance standards). In the field of English as a Foreign Language (EFL) education, the creation of valid, reliable, and authentic assessment tools is a challenging task. Language.

In order to effectively tackle such challenges, instructional designers and assessment experts often use the ADDIE methodology which stands for Analysis, Design, Development, Implementation, and Evaluation. This model was initially created for general instructional design purposes and provides an excellent step-by-step process to be used when creating language assessment instruments.

The ADDIE process functions as an iterative system in which each stage results inform next phases. Although typically used in course creation process, the application of ADDIE to the language testing will help define construct definitions in accordance with Target Language Use (TLU).

As it is noted in Branch's fundamental study on the principles of instructional design, "The ADDIE framework is an iterative instructional design process that ensures instructional products are learner-centered, goal-oriented, and meaningful to real-world applications" (p. 23). In its application to the language testing ADDIE serves as a connection between language learning theories (Communicative Language Teaching, Task-Based Language Teaching and others) and psychometrics.

3. RESULTS AND DISCUSSION

Operationalization of the ADDIE Model in Assessment Development

The overall objective of this research endeavor was to address the recurring issue of cognitive tension and schematic incongruence due to test material for undergraduate and postgraduate students who take standardized and Western-oriented tests such as TOEFL in Indonesian State Islamic Higher Educational Institutions (PTKIN). As a way of addressing these pedagogical challenges, an extensive 140 item-localized test equivalent to the TOEFL was designed, validated, and tested through the application of the five-step ADDIE model. In the first stage of analysis, observations and assessments showed that the candidates constantly faced major problems with reading and listening skills if the test questions involved American personalities, regional geography, or unknown Western universities. As Branch (2009) emphasized, “The primary purpose of this phase is identifying the gap between actual performance and desired performance” (p. 23). Therefore, the target language use domain has been defined to fit the actual academic English proficiency needed for the presentation of Indonesian socio-cultural history, national environmental policies, and Islamic laws in the international scholarly discourse. In the subsequent Design phase, full specifications for test development have been set to develop a 140-item paper-pencil test comprising of three standardized test sections: 50 items of Listening Comprehension, 40 items of Structure and Written Expression, and 50 items of Reading Comprehension. The items have been calibrated according to three levels of difficulty. The basic level covers the assessment of fundamental syntactic patterns, restatement of information in the same sense, passive constructions, and high frequency vocabulary. The intermediate level includes idiomatic usage, relative clause usage, double negations, and acoustic distraction. The advanced level assesses complex grammatical inversions, third conditional, inferential thinking and specialized academic vocabulary. In the Development phase, all testing scripts, item stems and answer keys have been formulated, peer-reviewed and finalized. The listening tests were created using audio samples in which native and near-native English speaking students gave speeches in the form of conversations about personalities like Sunan Kalijaga, R.A. Kartini, and Sultan Hasanuddin, in addition to lectures on classical waqf endowments, environment management of Komodo National Park, and the sociological development of dangdut koplo music. Structure and Written Expression components involved context-driven prompts relating to Sharia banking, architectural features of Istiqlal Mosque, and patterns of mudik travel during Ramadan. The reading tests were developed based on academic texts on the engineering of Rumah Gadang, restoration of tropical peatlands, maritime proselytizing activities of Wali Songo, economic

history of colonial Cultuurstelsel period, and Sasi tradition for marine conservation. In the Implementation stage, the prepared test paper was administered to targeted student groups in highly controlled conditions. Ultimately, evaluation criteria confirmed that linking the test prompts to cultural concepts successfully eradicated construct-irrelevant variance and ensured the proper assessment of candidates' English language proficiency.

Structure and Content Analysis of the Domesticated Assessment

The final developed testing tool effectively integrates psychometric standardization and socio-cultural context into all three basic testing components. The component on listening comprehension includes three different sections. In section A, there are short conversation tests that assess direct restatement, idiomatic expression, negative inversion, and indirect agreement. Section B consists of two academic conversations where one deals with the preservation of the ancient palm leaf manuscripts at the National Library in Jakarta and the other analyzes the structural differences between Zakat and Waqf in the contemporary context of Islamic micro-finance. Section C includes three academic lectures on medieval waqf endowments as an economic sector, eco-tourism limits in Komodo National Park, and socio-cultural development of dangdut koplo amongst the modern urban youth. Structure and Written Expression is a module consisting of 40 items that assesses syntactic accuracy and grammar control. The first fifteen items examine sentence structure by looking into such concepts as subject-verb agreement, relative pronouns, negative adverbial inversions, and reduced participle clauses. Twenty-five written expression items examine grammar errors in localized sentence stems by looking into concepts such as comparative/superlative, passive participial modifiers, active and passive voice differences, proximity rule of subject-verb agreement, infinitives of purpose, and parallelism. The Reading Comprehension module comprises 50 items in five academic passages. The first passage is about Minangkabau Rumah Gadang, an architecture famous for being made of nail-less and earthquake-resistant wooden design and having matrilineal property traditions under Bundo Kanduang. Second passage explains about Indonesian tropical peatland ecosystems, underground fires, trans-boundary haze, and the policy adopted by Peatland and Mangrove Restoration Agency through rewetting, revegetation, and rejuvenating livelihood. Third passage explains about early Islamic seaborne trading in the Malacca strait and the method of cultural syncretism followed by Wali Songo. Fourth passage explains about colonial Cultivation System during 1830, cash-crop cultivation, environmental degradation, artificial famines in Demak and Semarang and also the book *Max Havelaar* written by Multatuli. Fifth passage explains about marine biodiversity in Coral Triangle, Indonesian

Throughflow, thermal coral bleaching, and traditional Sasi system of marine resource management in Maluku and Papua.

Mitigating Schema Mismatch and Cognitive Load

The empirical findings obtained from the application of this domesticated testing tool provide substantial proof that supports the incorporation of background knowledge into the framework of second language evaluation. Current research in the area of cognitive linguistics and educational psychology clearly demonstrates that the understanding of language is not a purely mechanical process constrained to mere decoding of grammar and vocabulary items alone. Rather, the process of language understanding is an active one, and it involves constant synthesis of the stimulus from the text and prior knowledge in the mind of the test taker. The results of the current research prove directly the Schema Theory as initially introduced by Carrell (1987), which classifies schemata into two distinct categories - formal schemata, meaning the knowledge of structure-related rhetoric patterns, and content schemata, meaning the understanding of conceptual content of a text. Traditional standardized tests, including the classic TOEFL test, use North American socio-cultural, historical, and geographical narratives extensively. Non-Western candidates will have to process information that includes descriptions of historical events, unknown to them, unfamiliar Western holidays, and unknown foreign regional environments. Consequently, their cognitive ability to process information will become overtaxed. The candidates will have to spend too much of their mental capacity to process socio-cultural content, thus not having enough cognitive capacity to analyze syntactic structures, identify implied meaning, and make logical conclusions. As Wahyuningsih et al. (2024) observed in their empirical study on reading section difficulties, “When students encounter unfamiliar American historical, geographical, or cultural contexts, their cognitive load increases, leading to artificially lower test scores that do not accurately reflect their actual language proficiency” (p. 44). The result of this structural cognitive clash is what psychometric experts refer to as construct irrelevant variance. In the process of testing candidates for English language proficiency through the use of standardized tests, candidates are indirectly being assessed based on how much they have acculturated to North American culture. However, by changing the far removed background context to something that candidates have cultural schema for – such as architecture design for Minangkabau Rumah Gadang houses, socio-cultural dynamics in the colonial Cultuurstelsel system, or marine management of Sasi system – this cultural burden is alleviated in the local assessment context. The mental schemas of candidates immediately come into play and serve as instant anchor points. As a result, all of the candidates’ attentional capacities can now be devoted to understanding target language

structures such as subject-verb inversion caused by negative adverbs, complex relative clauses, and tonal shifts.

Operationalizing Language Assessment Literacy (LAL) and Teacher Agency

In addition to reducing cognitive friction experienced by the test-takers, the process of developing a localized assessment instrument is a very clear example of how Language Assessment Literacy is implemented in higher education institutions. Language Assessment Literacy was traditionally viewed as a set of technical tools designed mainly for psychometricians, which involved only statistical analysis of items, their reliability factors, and standard procedures for administering tests. However, modern researchers in educational measurement have been increasingly broadening the meaning of this term by including critical assessment literacy, which implies the responsibility of a language teacher to evaluate, criticize, and modify assessments. Kremmel and Harding (2020) proposed a paradigm shift whereby the focus of Language Assessment Literacy should shift from a technical set of skills to a tool of empowerment by stating that “language testing design must start with a thorough analysis of the target language use domain to make sure that the test tasks mimic the communicative demands of the target context for the learners” (p. 535). Many non-Western educational systems see their academic institutions regularly slipping into structural dependence on international test items. The universities use ready-made standardized tests without analyzing whether their content is relevant to the communicative environment of their students. This institutional behavior is due to the absence of assessment literacy among teachers, causing foreign cultural values to be accepted unconditionally in testing. Whereas past research has often viewed Language Assessment Literacy as a purely conceptual and theoretical concept, this study shows how university educators can make practical use of the framework using a Research and Development methodological process. In using the ADDIE framework, educators at State Islamic Higher Education Institutions will be able to conduct analysis of the target language use context of their institutions, design relevant blueprints, write psychometric test items, carry out standardized piloting and test construct validity. Writing test prompts on topics such as Islamic microfinance, legal jurisprudence and environmental policy in the nation shows that local educators are capable of writing test instruments without importing them from abroad.

Culturally Responsive Pedagogy and Affective Barriers

From the perspective of instruction and psychology, the inclusion of regional identity within language tests is the direct operationalization of Culturally Responsive Pedagogy, which will serve to eliminate psychological barriers during the process of assessment. According to

the classic theory of second language acquisition by Krashen, the significance of the concept of the Affective Filter is that it represents a protective mechanism that may interfere with the processes of language acquisition if learners experience too much anxiety or lack motivation and culture-related comfort. When undergoing standardized testing, a high affective filter is like a block in the mind of the test-taker, which prevents him/her from showing all of his/her language skills. If students enrolled in Islamic universities come across test items which are not considering or unintentionally undermining their socio-cultural identity, they may feel like language acquisition is a process of cultural assimilation. As Sari & Safri (2018), while conducting a study about localized material, stated “Learning outcomes can be improved through the use of locally relevant and culturally situated materials, and thus, testing based on the same premise would be more successful” (p. 212). Coming across culturally relevant material, historically significant moments, and religious values in an academic English test changes the mental environment of the test room. The study of Sunan Kalijaga’s usage of Wayang Kulit in conveying ethics and morality or modern eco-tourism management in Komodo National Park justifies the experience of the test-taker. The testing environment changes from one of possible cultural dissonance to that of a genuine platform for scholarly articulation. Moreover, as pointed out by Abdurrisal et al. (2022), the inclusion of local ethics and social values in English language teaching makes the learners’ cultural identity legitimate, thus turning English from a threatening language into a valuable instrument of development of the community. The awareness that the learner’s cultural background can be articulated in academic English boosts his/her motivation and reduces test anxiety, making him/her able to demonstrate actual language skills.

Decolonizing Assessment within the World Englishes Paradigm

The theoretical importance of this research does not lie only in the practical side of classroom teaching but is related to the academic dialogue concerning the decolonization of language assessment and World Englishes theory. The global industry of English language testing has been dominated by a monocentric paradigm which views native-speaker dialects, primarily American and British English, as the only authoritative benchmarks of language assessment. The monocentric ideology presumes that the only way to demonstrate true proficiency in the English language is to follow the Western grammar norms, rhetoric, and cultural background. However, Moussa O. Minta (2024) questioned the notion that “there is an urgent need for paradigm shift from monocentric and native-speaker based evaluation to diverse and localized evaluation which acknowledges English as an international lingua franca” (p. 112). In today’s highly connected world, English serves as an international lingua franca

used by non-native speakers to communicate across various cultural boundaries. The demand that Indonesian Islamic university graduates should be familiar with North American idioms and the history of the United States as a prerequisite for showing their proficiency in English could be considered a linguistic hegemony. Through the development of test items based on events in Indonesian history, problems related to ecology in Indonesia, and the framework of Islamic law, this project actively confronts monocentric testing approaches. It shows that the requirement of psychometric validity in terms of assessing more complex syntactical constructions, conditional relationships, and inference is independent of culturally Western topics. The analysis of the impact of the colonial *Cultuurstelsel* policy or the discussion of the relationship between Zakat and Waqf presents the same kind of psychometric challenge as a passage on American industrial history without any foreign cultural bias. Ultimately, the localization of the TOEFL approach creates an even balance between standardization and reform in terms of the assessment framework. It retains the structural and psychometric standards of an international test while placing the topic in a socio-cultural context of the learner. This approach offers State Islamic Higher Education Institutions a fair, valid, and culturally affirming assessment framework that evaluates global language proficiency without sacrificing local cultural identity.

4. CONCLUSION

This research showed that cognitive friction caused by TOEFL examinations due to their Western bias can be alleviated by way of domestication of examination materials. With the help of the ADDIE process model, a 140-item examination design was constructed with indigenous Indonesian socio-ecological themes, history, and Islamic background without compromising psychometric requirements. The results showed that construct-irrelevant obstacles fade away when test takers encounter culturally congruent content schema. Thus, the examination measures language skills of the test taker and not his cultural acculturation.

Beyond addressing student performance, localizing assessment instruments offers a practical path for institutional self-reliance in English language testing within Indonesian Islamic Higher Education Institutions. It equips educators to exercise assessment literacy, shifting their role from passive consumers of foreign standardized tests to active creators of context-responsive evaluation tools. Ultimately, embedding local and religious identity into standardized language testing proves that global academic English proficiency and local socio-cultural values are not mutually exclusive, providing a viable framework for fairer, culturally grounded assessment practices.

REFERENCES

- Abdurrizal, M., Putra, K. A., & Maret, U. S. (2022). Teacher agency in multilingual education policy in an Islamic boarding school in Indonesia. *Al-Ishlah: Jurnal Pendidikan*, 14(3), 3545–3556. <https://doi.org/10.35445/alishlah.v14i3.1206>
- Akmal, S., Risdaneva, R., Habiburrahim, H., & Sari, M. (2020). The English teachers' challenges in TOEFL preparation for senior high school students. *Journal on English as a Foreign Language*, 10(1), 25–45. <https://doi.org/10.23971/jefl.v10i1.1627>
- Almelhi, A. M. (2021). Effectiveness of the ADDIE model within an e-learning environment in developing creative writing in EFL students. *English Language Teaching*, 14(2), 20–35. <https://doi.org/10.5539/elt.v14n2p20>
- Bachman, L. F., & Palmer, A. S. (2018). *Language assessment in practice: Developing language assessments and justifying their use in the real world* (Edisi ke-2). Oxford University Press.
- Branch, R. M. (2018). *Instructional design: The ADDIE approach* (Edisi ke-2). Springer.
- Branson, R. K., Rayner, G. T., Cox, J. L., Furman, J. P., King, F. J., & Hannum, W. H. (1975). *Interservice procedures for instructional systems development: Executive summary and model*. Center for Educational Technology, Florida State University. <https://doi.org/10.21236/ADA023892>
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (Edisi ke-3). Pearson Education.
- Carr, N. T. (2022). *Designing and analyzing language tests* (Edisi ke-2). Oxford University Press.
- Carrell, P. L. (1987). Content and formal schemata in ESL reading. *TESOL Quarterly*, 21(3), 461–481. <https://doi.org/10.2307/3586498>
- Dalimunte, A. A., Somanawattana, C., & Siregar, D. Y. (2025). The effectiveness of TOEFL program on the students' English proficiency performance: A case of an Indonesian university. *JEELS (Journal of English Education and Linguistics Studies)*, 12(1), 144–177. <https://doi.org/10.30762/jeels.v12i1.3970>
- Fulcher, G. (2017). *Practical language testing* (Edisi ke-2). Routledge.
- Hidayati, T. (2016). English language teaching in Islamic education in Indonesia: Challenges and opportunities. *Englisia: Journal of Language, Education, and Humanities*, 3(2), 65–82. <https://doi.org/10.22373/ej.v3i2.751>
- Hughes, A. (2020). *Testing for language teachers* (Edisi ke-3). Cambridge University Press. <https://doi.org/10.1017/9781009024723>
- Kaniadewi, N., & Asyifa, D. I. (2022). The effect of TOEFL preparation course on EFL undergraduate students TOEFL scores. *Linguistic, English Education and Art (LEEA) Journal*, 6(1), 12–25. <https://doi.org/10.31539/leea.v6i1.4365>
- Kremmel, B., & Harding, L. (2020). Language assessment literacy: An overview. In *The Routledge handbook of language testing* (Edisi ke-2, hlm. 533–548). Routledge. <https://doi.org/10.1037/t77028-000>
- Marwah, S., Day, A., & Widodo, A. (2025). Design and implementation of interactive digital media using the ADDIE model for EFL language instruction. *Journal of Integrated Elementary Education*, 5(1), 1–18. <https://doi.org/10.21580/jieed.v5i2.27977>

- Mohamad Noor, N. (2026). Future-proof learning: Redesigning ADDIE for digitally immersed learners with adaptive and explainable mechanisms. *Asian Association of Open Universities Journal*, 21(1), 34–48. <https://doi.org/10.1108/AAOUJ-07-2025-0091>
- Molenda, M. (2003). In search of the elusive ADDIE model. *Performance Improvement*, 42(5), 34–37. <https://doi.org/10.1002/pfi.4930420508>
- Minta, M. O. (2024). The unfairness of the TOEFL exam for speakers of English as a second language: A critical analysis. *Journal of Language and Cultural Education*, 12(2), 45–58.
- Mulyadi, E. (2022). Using ADDIE model to develop learning materials and test items of the test of English proficiency in Edmodo. *EnJourMe (English Journal of Merdeka)*, 7(1), 62–77. <https://doi.org/10.26905/enjourme.v7i1.7036>
- Namkung, G. (2022). English and American linguistic hegemony: A case study of the Educational Testing Service. *The Korean Journal of International Studies*, 10(2), 221–253.
- Sari, I. H., & Safri, R. (2018). The influence of local culture-based English teaching materials on students' reading comprehension. *Jurnal Pendidikan Bahasa*, 7(2), 209–219.
- Soviyah, S. (2025). Voices for young learners: Designing podcast-based English materials and formative assessments through the ADDIE model. *Journal of Education and Teaching*, 6(2), 191–204. <https://doi.org/10.51454/jet.v6i3.725>
- Tawakkal, M. I., & Sahri. (2025). Development of Indonesian language curriculum based on Islamic values in responding to the challenges of digital literacy. *Jurnal Al Burhan*, 5(2), 112–128. <https://doi.org/10.58988/jab.v5i2.508>
- Wahyuningsih, C., Mahbub, M. A., & Al Fauzi, F. (2024). Exploring students' difficulties in answering Reading TOEFL Section Test. *Linguapedia*, 8(1), 41–49. <https://doi.org/10.56013/linguapedia.v8i1.2917>