https://research.adra.ac.id /index.php/ijen/
P - ISSN: 2988-1579
E - ISSN: 2988-0092

Development and Evaluation of a Cognitive Instructional Design (CID) Model for English Proficiency in Higher Education

Yepi Sedya Purwananti[1], Punaji Setyosari[2], Bambang Yudi Cahyono[3], Sulthon[4]
[1]Universitas Negeri Malang, Universitas Bhinneka PGRI, Indonesia
[2] Universitas Negeri Malang, Indonesia
[3] Universitas Negeri Malang, Indonesia
[4] Universitas Negeri Malang, Indonesia

ABSTRACT

Background. English proficiency is an important competency in higher education, yet some students struggle to meet institutionally required TOEFL standards.

Purpose. This study developed and evaluated a Cognitive Instructional Design (CID) model and an integrated learning package for proficiency-oriented English instruction. The package comprised a model book, Student's Book, Assessment Book, Teacher's Book, and Online Review Materials delivered through Google Classroom.

Method. A development study using the Dick, Carey, and Carey instructional design model was conducted at STKIP PGRI Tulungagung, Indonesia. A purposive sample of 16 undergraduate students enrolled in an Introduction to TOEFL course completed a one-group pretest–posttest evaluation after six instructional sessions. Data were analyzed using descriptive statistics and paired-samples t-tests; normality was assessed using the Kolmogorov–Smirnov test.

Results. Mean total proficiency scores increased from 409.13 (SD = 48.73) at pretest to 470.25 (SD = 31.24) at posttest, a mean increase of 61.13 points, $t(15) = -9.034$, $p < .001$, 95% CI [−75.55, −46.70] for the pretest-minus-posttest difference. The within-group effect size was Cohen's $d_z = 2.26$. Scores also increased across listening comprehension, structure and written expression, and reading comprehension.

Conclusion. The findings show a substantial pre–post improvement in this group. Because the study involved only 16 students at one institution and had no control group, the results provide preliminary evidence of potential rather than definitive causal evidence of effectiveness. The model offers a structured framework for proficiency-oriented instruction that warrants further evaluation in larger controlled studies.

KEYWORDS

Cognitive Instructional Design, English for Proficiency, Instructional Design, TOEFL

Citation: Purwananti, S. Y., Setyosari, P., Cahyono, Y. B., & Sulton, Sulton. (2026). Development and Evaluation of a Cognitive Instructional Design (CID) Model for English Proficiency in Higher Education. *International Journal of Educational Narratives*, *4*(5), 2458–2466. https://doi.org/10.70177/ijen.v4i5.4194

Correspondence:
Yepi Sedya Purwananti,
yepisedya@gmail.com

Received: July 12, 2026
Accepted: October 24, 2026
Published: October 31, 2026

INTRODUCTION

English proficiency supports participation in higher education, research, and international professional networks. Cross-national studies associate English proficiency with educational and socioeconomic opportunities, although these relationships should not be read as direct causal effects (Si et al., 2025; Terasawa, 2025). For Indonesian universities, proficiency can be

important for graduation, scholarships, academic mobility, and employment.

In Indonesia, English is generally learned as a foreign language, and some university students continue to fall short of institutional proficiency requirements despite completing English courses (Raja et al., 2022; Azhari et al., 2022). The institutional needs analysis for this study indicated that fewer than 60% of students achieved the required English-proficiency score, and students who did not meet the threshold needed remedial support. This local problem motivates instruction designed specifically around the skills assessed in proficiency tests, rather than relying on general English instruction alone. TOEFL-oriented learning commonly addresses listening comprehension, structure and written expression, and reading comprehension (Kurniati et al., 2025; Dalimunte et al., 2023).

TOEFL-oriented instruction requires alignment among learning objectives, materials, instructional activities, assessments, and learner support. Instructional design offers a systematic means of aligning these elements with learner characteristics and intended outcomes (Abuhassna et al., 2024). For online and blended learning, course components and learning activities also need to be coordinated across delivery modes (Fayed & Cummings, 2025). The current definition from the Association for Educational Communications and Technology frames educational technology as the ethical study and application of theory, research, and practice to advance knowledge, improve learning and performance, and empower learners through the strategic design, management, implementation, and evaluation of learning experiences and environments using appropriate processes and resources (Association for Educational Communications and Technology [AECT], 2023). These principles support a proficiency-oriented learning package that connects cognitive theory to TOEFL-focused practice and feedback.

Cognitive learning theory provides a relevant foundation for language learning because it examines how learners acquire, organize, store, and retrieve information (Brown & Lee, 2025). Schema theory explains how new information is integrated into existing knowledge structures. Rumelhart and Norman (1978) describe three modes of schema change: accretion, tuning, and restructuring. Cognitive Instructional Design (CID) offers a framework for analyzing the cognitive demands and relationships among components of complex instructional designs (Loibl et al., 2025). For proficiency-oriented English learning, these ideas suggest sequencing activities so that learners activate and extend prior knowledge, refine their understanding of linguistic patterns, and reorganize knowledge through transfer-oriented practice. Recent work on cognitive strategies for foreign-language acquisition also emphasizes purposeful use of learning strategies and tools (Broda et al., 2025).

Prior studies have investigated the relationship between language-learning strategies and achievement (Bećirović et al., 2021), strategy-supported flipped EFL instruction (Öztürk & Çakıroğlu, 2021), and technology-supported language learning (Shadiev & Wang, 2022). Other work has examined diagnostic assessment with individualized feedback (Anisah & Amreta, 2024) and documented English-proficiency levels among Indonesian university students using EnglishScore-based assessment (Kurniati et al., 2025). Together, these studies establish the relevance of learning strategies, assessment, and technology, but they do not document how schema-based cognitive instructional design can be translated into and preliminarily evaluated as a coordinated TOEFL-oriented package that includes a CID model, student materials, assessment tools, teacher guidance, and online review. The gap addressed here is therefore specific: the integration and preliminary testing of these components in one institutionally targeted proficiency course.

The contribution of this study is the development and preliminary evaluation of a TOEFL-oriented CID package that links schema processes to three instructional stages understanding,

associating, and discovering—and coordinates five product components: a model book, Student's Book, Assessment Book, Teacher's Book, and Online Review Materials delivered through Google Classroom. The development follows the Dick, Carey, and Carey instructional design model. Specifically, the study aims to (1) develop a theory-informed CID model and integrated learning package, (2) refine the product using expert and user feedback, and (3) examine preliminary pre–post changes among 16 students. The contribution is design-based and context-specific; the one-group evaluation is intended as a preliminary effectiveness test and cannot, by itself, establish that the model caused the observed changes.

RESEARCH METHODOLOGY

This research used a research-and-development approach to develop and preliminarily evaluate a Cognitive Instructional Design (CID) model and learning package for proficiency-oriented English instruction. The Dick, Carey, and Carey model was selected because it provides a systematic sequence for aligning instructional goals, learner characteristics, assessment, strategy, materials, and evaluation (Dick et al., 2015). The study was conducted at STKIP PGRI Tulungagung, Indonesia, where English for Proficiency is a compulsory course intended to support students in meeting institutional TOEFL requirements. The implementation sample comprised 16 undergraduate students from the English Education Study Program who were enrolled in an Introduction to TOEFL course. Participants were selected through purposive sampling because they represented the intended users of the product. Most were 20–21 years old and had completed a general English course, although their initial proficiency levels varied.

The development followed the ten stages of the Dick, Carey, and Carey model: identifying instructional goals; conducting instructional analysis; analyzing learners and learning contexts; writing performance objectives; developing assessment instruments; developing instructional strategies; developing and selecting instructional materials; conducting formative evaluation; revising the product; and conducting summative evaluation. For reporting purposes, the stages were grouped into analysis, development, and evaluation–revision. The analysis examined institutional needs, learner characteristics, prior knowledge, learning contexts, and instructional objectives. The needs analysis indicated that fewer than 60% of students met the institution's required proficiency threshold. On the basis of schema theory, the CID sequence was organized into three stages—understanding, associating, and discovering—intended to support accretion, tuning, and restructuring through listening comprehension, structure and written expression, and reading comprehension activities. The package comprised a model book, Student's Book, Assessment Book, Teacher's Book, and Online Review Materials in Google Classroom.

The study used needs-analysis instruments to document institutional requirements, learner characteristics, and the learning context; expert-review questionnaires and user-evaluation forms to collect formative feedback; structured and unstructured interviews to elicit comments and suggestions from students and instructors; and TOEFL-oriented pretests and posttests covering listening comprehension, structure and written expression, and reading comprehension. The product-review criteria addressed theoretical alignment, relevance and accuracy of content, alignment with learning objectives, instructional-material quality, and usability.

Product validation involved three experts with complementary areas of expertise: English-language content, instructional design and learning models, and instructional media. The content expert reviewed the organization and coverage of the instructional materials and recommended reorganizing the Structure and Written Expressions component to ensure that the extensive range of skills could be accommodated within the available instructional time. The instructional design

expert recommended explicitly incorporating the instructional effects and nurturant effects of the proposed model. Meanwhile, the instructional media expert suggested revising the layout of the instructional books to comply with established book-design conventions. In addition to expert reviews, the course instructor provided feedback emphasizing the need to provide students with opportunities for independent practice. These recommendations informed revisions to the instructional package, including improvements to the model description, content organization, book layout, and opportunities for self-directed practice.

Following expert review, formative evaluation included a one-to-one tryout with three English Education students representing high, moderate, and low academic achievement, followed by a small-group tryout involving six Mathematics students, with two students at each achievement level. These tryouts provided additional user feedback to identify areas requiring refinement before the subsequent field implementation. Feedback from the experts, course instructor, and student tryouts was summarized descriptively and used to guide product revisions. The validation findings were reported qualitatively through expert comments and recommendations; the dissertation does not report numerical validator ratings, a content-validity index, or a test-reliability coefficient. Accordingly, the validation process is presented as expert-informed product review and formative refinement rather than as a quantified assessment of psychometric validity.

A one-group pretest–posttest design was used for a preliminary field evaluation of the developed product. The same 16 students completed a pretest before instruction and a posttest after six instructional sessions. That quantitative results were processed using SPSS. This design permits examination of within-participant change, but without a comparison group it cannot rule out alternative explanations such as test familiarity, maturation, or events occurring during the intervention. Descriptive statistics (means and standard deviations) summarized performance. Normality was assessed with the Kolmogorov–Smirnov test at $\alpha = .05$, and paired-samples t-tests compared pretest and posttest scores for the total proficiency score and each TOEFL component. The effect size was calculated as Cohen's dz, using the mean paired difference divided by the standard deviation of paired differences.

RESULT AND DISCUSSION

The development process produced a CID learning model and an integrated learning package for proficiency-oriented English instruction. The results are presented in three parts: the needs analysis and product development; the evaluation and revision process; and the change in total and component-level TOEFL-oriented scores from pretest to posttest.

Analysis Stage

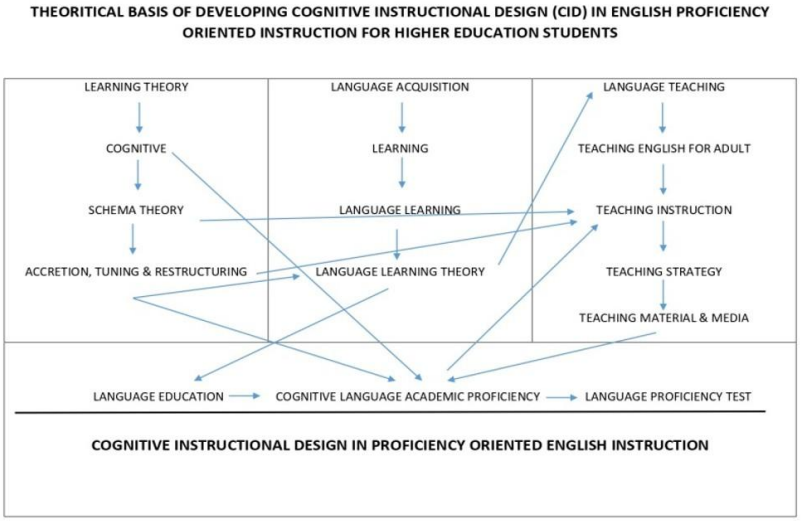
The needs analysis showed that fewer than 60% of students met the institution's required English-proficiency score. Students entered the course with heterogeneous English skills, and the proficiency requirement created a clear need for structured support. The resulting learning objectives emphasized identifying the major TOEFL components, applying problem-solving strategies, and improving performance through a sequenced process of understanding, associating, and discovering.

Development Stage

The development stage produced a CID model for proficiency-oriented English learning. Its instructional sequence consisted of three stages understanding, associating, and discovering—designed to help learners form, refine, and reorganize knowledge structures. The stages were

mapped to the listening-comprehension, structure and written-expression, and reading-comprehension components of the TOEFL-oriented course.

The integrated package comprised five components: a model book, Student's Book, Assessment Book, Teacher's Book, and Online Review Materials delivered through Google Classroom. The materials addressed listening comprehension, structure and written expression, and reading comprehension. Activities in each component were organized around the CID sequence, moving from understanding concepts to associating them with prior knowledge and discovering solutions through proficiency-test practice.



Picture 1. Syntax of CID Learning Model

Evaluation and Revision Stage

As illustrated in Figure 1, the CID framework connects cognitive learning and schema theory with language acquisition, instructional strategy, learning materials, and proficiency testing. During expert review, the learning-model expert recommended adding the model's instructional and nurturant effects. The English-content expert considered the materials sufficiently organized and comprehensive across proficiency-test components, but recommended reorganizing the structure and written-expression section because its many skills required substantial instructional time. The media expert recommended revising the book layout to conform to book-publication conventions. The course instructor also recommended giving students access to independent practice because the course requires sustained attention from both teachers and learners. Formative evaluation further included one-to-one review by three English Education students (high, moderate, and low achievement) and small-group review by six Mathematics students (two at each achievement level). The feedback informed revisions before the 16-student field implementation. Numerical validator ratings or a content-validity index are not reported in the dissertation.

Paired-Samples t Test

The total-score pretest and posttest results were compared using a paired-samples t-test in SPSS version 21.0, as specified in Chapter III. Because the same 16 students completed both administrations, the analysis tested the mean within-participant difference. Results are reported with the direction of the difference stated explicitly; p-values displayed by SPSS as .000 are reported as $p < .001$ rather than as zero.

Table 1. Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	Pretest	409.1250	16	48.72901	12.18225
	Posttest	470.2500	16	31.24420	7.81105

Table 2. Paired Sample T Test

Paired Differences									
					95% Confidence Interval of the Difference				
		Mean	Std . Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	Pretest-Posttest	61.12500	27.06505	6.76626	-75.54694	-46.70306	-9.034	15	.000

The total proficiency score increased from a pretest mean of 409.13 (SD = 48.73) to a posttest mean of 470.25 (SD = 31.24), corresponding to a mean increase of 61.13 points. For the contrast coded as pretest minus posttest, the paired-samples test was statistically significant, $t(15) = -9.034$, $p < .001$, 95% CI [−75.55, −46.70]. The within-group standardized change was Cohen's $dz = 2.26$, calculated as the absolute mean paired difference (61.125) divided by the standard deviation of paired differences (27.065). This is a large standardized within-group change; it should not be interpreted as a controlled causal effect because the study had no comparison group.

Table 3. T-Test Results for Comparison of Pretest and Post Test Scores

Variable	Pre-test		Post-test		Average difference	Correlation test	Sig. (t-test)
	Mean	Std. deviation	Mean	Std. deviation			
Listening scores	43.31	4.81	48.69	3.93	-5.38	r=0.940 p=0.001	0.000
Structure scores	38.56	6.74	45.19	3.43	-6.63	r=0.589 p=0.016	0.000
Reading scores	40.88	5.34	47.19	3.04	-6.31	r=0.815 p=0.000	0.000

For listening comprehension, the mean increased from 43.31 (SD = 4.81) at pretest to 48.69 (SD = 3.93) at posttest, an increase of 5.38 points. The pretest and posttest scores were strongly correlated ($r = .940$, $p < .001$), and the paired-samples t-test indicated a statistically significant difference ($p < .001$).

For structure and written expression, the mean increased from 38.56 (SD = 6.74) to 45.19 (SD = 3.43), an increase of 6.63 points. The pretest and posttest scores were moderately correlated ($r = .589$, $p = .016$), and the paired-samples t-test indicated a statistically significant difference ($p < .001$). For reading comprehension, the mean increased from 40.88 (SD = 5.34) to 47.19 (SD = 3.04), an increase of 6.31 points. The pretest and posttest scores were strongly correlated ($r = .815$, $p < .001$), and the paired-samples t-test indicated a statistically significant difference ($p < .001$).

The increase across the total score and all three TOEFL components suggests that the CID package was associated with improved performance during the six-session implementation. The pattern is consistent with literature indicating that language-learning strategies, self-regulated

learning in flipped EFL classrooms, and technology-supported language learning can contribute to language development (Bećirović et al., 2021; Öztürk & Çakıroğlu, 2021; Shadiev & Wang, 2022). However, those studies differ in design and outcome measures, so a numerical effect-size comparison would require directly comparable data that are not available in the cited literature as currently assembled.

Cohen's dz = 2.26 indicates a large standardized pre–post difference, but does not address internal-validity limits. Without a control group, testing/practice effects, maturation, and events during implementation remain alternative explanations; the purposive sample of 16 students at one institution limits generalization. Chapter V also notes three product-level constraints: the model's scope is limited to proficiency-oriented courses, implementation depends on instructor/tutor competence, and outcomes are influenced by learners' self-awareness, motivation, and self-management. The observed gain is promising preliminary evidence, not proof that CID alone caused the change. Expert identities and revision comments are documented, but the dissertation does not report numerical validator ratings, a content-validity index, or a test-reliability coefficient. Add these only if original records exist; report formal observation findings only if a separate protocol and analysis exist.

For practical adoption, institutions could pilot the package by mapping each TOEFL component to explicit objectives, training instructors to use the three CID stages, combining classroom practice with structured Google Classroom review, and monitoring component-level pre–post scores against institutional thresholds. The next evaluation should use a larger and more diverse sample, include a comparison group where feasible, report instrument validity and reliability evidence, and replicate the intervention across institutions. These steps would make it possible to assess both the transferability of the model and the stability of the observed gains.

CONCLUSION

This study developed a Cognitive Instructional Design (CID) model and an integrated learning package for TOEFL-oriented English instruction. The package includes a model book, Student's Book, Assessment Book, Teacher's Book, and Online Review Materials delivered through Google Classroom. Among 16 students at STKIP PGRI Tulungagung, total and component-level scores were higher after six instructional sessions. These results constitute preliminary evidence of positive within-group change, while the one-group design and small single-institution sample do not support a definitive causal conclusion.

The design contribution lies in translating schema processes—accretion, tuning, and restructuring—into the instructional sequence of understanding, associating, and discovering, connected to TOEFL-aligned materials and assessments. Institutions considering adoption should first conduct a local needs analysis, prepare instructors to implement the three stages, and monitor outcomes by skill component before scaling the package. Future studies should test the model with larger samples, a control or comparison group, replicated implementations, and reported evidence of instrument validity and reliability. Qualitative observation findings should be included only if an observation protocol and documented analysis were actually used.

DECLARATION OF AI AND AI ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work the author(s) used ChatGPT and Perplexity in order to Article Search. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

ACKNOWLEDGEMENT

The author expresses his deepest gratitude to STKIP PGRI Tulungagung for providing institutional support and research facilities during this research. The authors also thank the experts who participated in the validation process and the students who voluntarily participated in the implementation and evaluation of the developed learning model. Their valuable contributions and constructive feedback were essential to the successful completion of this research.

AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Project administration; Software; Validation; Visualization; Writing – original draft.
Author 2: Conceptualization; Methodology; Supervision; Validation; Writing – review and editing.
Author 3: Formal analysis; Investigation; Validation; Writing – review and editing.
Author 4: Resources; Supervision; Validation; Writing – review and editing.

DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

Abuhassna, H., Adnan, M. A. B. M., & Awae, F. (2024). Exploring the synergy between instructional design models and learning theories: A systematic literature review. *Contemporary Educational Technology, 16*(2), Article ep499. https://doi.org/10.30935/cedtech/14289

Anisah, G., & Amreta, M. Y. (2024). Cognitive diagnostic assessment instrument with individual feedback for language receptive learning on multimodal text. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya, 10*(1), 366–387. https://doi.org/10.22219/kembara.v10i1.32974

Association for Educational Communications and Technology. (2023). *AECT definition of educational technology*. https://aect.org/aect/about/aect-definition

Azhari, T., Sahputri, J., & Rasyimah. (2022). Analysis of student perceptions on the implementation of required TOEFL test before graduation. In *Proceedings of the 2nd International Conference on Social Science, Political Science, and Humanities (ICoSPOLHUM 2021)* (pp. 125–130). Atlantis Press. https://doi.org/10.2991/assehr.k.220302.019

Bećirović, S., Brdarević-Čeljo, A., & Polz, E. (2021). Exploring the relationship between language learning strategies, academic achievement, grade level, and gender. *Journal of Language and Education, 7*(2), 93–106. https://doi.org/10.17323/jle.2021.10771

Broda, M., Kovalova, K., Bezugla, I., Stoika, O., & Mulyk, K. (2025). Enhancing foreign language acquisition: Innovative cognitive strategies and tools. *Conhecimento & Diversidade, 17*(45), 166–182. https://doi.org/10.18316/rcd.v17i45.12371

Brown, H. D., & Lee, H. (2025). *Principles of language learning and teaching: A course in second language acquisition* (7th ed.). Routledge.

Dick, W., Carey, L., & Carey, J. O. (2015). *The systematic design of instruction* (8th ed.). Pearson.

Dalimunte, A. A., Tika, A., Dalimunte, M., Tanjung, A. P., & Suryani, I. (2023). Evaluating TOEFL ITP test: A critical review. *Humanitatis: Journal of Language and Literature, 9*(2), 245–254. https://doi.org/10.30812/humanitatis.v9i2.2329

Fayed, I., & Cummings, J. (2025). Cross-modalities instructional design model (CIDM) for online and blended teaching and learning. *International Journal on E-Learning, 24*(3), 301–333. https://doi.org/10.70725/157616jbyghw

Kurniati, D., Riyono, A., Solaeman, A., Millan, A. R., & Noordin, N. (2025). Exploring English proficiency levels of Indonesian university students through EnglishScore-based assessment. *Indonesian EFL Journal, 11*(3), 539–552. https://doi.org/10.25134/9kjdz259

Loibl, K., Leuders, T., Glogger-Frey, I., & Rummel, N. (2025). CID: A framework for the cognitive analysis of composite instructional designs. *Instructional Science, 53*(6), 1485–1509. https://doi.org/10.1007/s11251-024-09665-9

Öztürk, M., & Çakıroğlu, Ü. (2021). Flipped learning design in EFL classrooms: Implementing self-regulated learning strategies to develop language skills. *Smart Learning Environments, 8*, Article 2. https://doi.org/10.1186/s40561-021-00146-x

Raja, P., Flora, Putrawan, G. E., & Razali, A. B. (2022). English as an international language: Perceptions of EFL preservice teachers in higher education institutions in Indonesia. *Education Research International, 2022*, Article 3234983. https://doi.org/10.1155/2022/3234983

Rumelhart, D. E., & Norman, D. A. (1978). Accretion, tuning, and restructuring: Three modes of learning. In J. W. Cotton & R. L. Klatzky (Eds.), *Semantic factors in cognition* (pp. 37–54). Lawrence Erlbaum Associates.

Shadiev, R., & Wang, X. (2022). A review of research on technology-supported language learning and 21st-century skills. *Frontiers in Psychology, 13*, Article 897689. https://doi.org/10.3389/fpsyg.2022.897689

Si, K., Zhang, Y., & Zhang, D. (2025). English proficiency and national competitiveness in sustainable development research based on a bibliometric analysis. *Discover Sustainability, 6*, Article 547. https://doi.org/10.1007/s43621-025-01447-8

Terasawa, T. (2025). Relationship between English proficiency and socioeconomic status in Asia: Quantitative cross-national analysis. *World Englishes, 44*(3), 492–516. https://doi.org/10.1111/weng.12705.

Copyright Holder :
© Yepi Sedya Purwananti et al. (2026).

First Publication Right :
© International Journal of Educational Narratives

This article is under:

