Developing SDG-Based Multimodal Texts for Graded Readers with GenAI to Support English Reading

Fela Dafitri[1]

Fauzi Miftakh[2]

Agung Susilo Yuda Irawan[3]

Putri Kamalia Hakim[4]

Universitas Singaperbangsa Karawang

fela.dafitri@fkip.unsika.ac.id[1]

Abstract

This study developed Sustainable Development Goal (SDG)-based multimodal texts for graded readers in an Indonesian vocational high school. It addressed three research questions concerning the development process using Generative AI, the characteristics of the resulting texts, and students' perceptions. The study employed the ADDIE model (Analysis, Design, Development, Implementation, and Evaluation). Data were collected from four English teachers through a needs-analysis questionnaire, one English education expert through a validation form, and 61 vocational students through a post-use questionnaire. Based on the identified needs, 15 texts were developed across five operational proficiency levels (A1, A2, A2–B1, B1, and B2), covering SDG 8, SDG 9, and SDG 12. Generative AI supported text generation, adaptation, and illustration production, while Google AI Studio was used to create audio narration. Pedagogical specifications and final revisions remained under researcher supervision. Expert validation yielded a mean score of 4.40/5.00, indicating that the product was appropriate for implementation. Students reported generally positive perceptions (M = 3.91/5.00), particularly regarding audio support (M = 4.13) and the contribution of texts to understanding SDG issues (M = 4.02). Audio speed received the lowest rating (M = 3.59), indicating a need for improvement. Overall, the findings demonstrate the potential of integrating graded linguistic progression, SDG content, multimodal features, and Generative AI in vocational English material development. However, the results reflect perceived usefulness and acceptability rather than measured learning gains or causal effects.

Keywords: *Generative AI; graded readers; multimodal texts; Sustainable Development Goals; vocational English*

Introduction

English learning in vocational education needs to address both language development and the communicative demands of students' prospective professional contexts. Vocational learners use English not only as a school subject but also as a resource for accessing occupational information, understanding workplace communication, and participating in professional activities. Consequently, English materials for vocational learners need to balance linguistic accessibility with contextual relevance.

This issue is particularly important in reading. Vocational students may encounter difficulty when texts contain unfamiliar vocabulary, complex structures, or topics that are distant from their study programs and future work. Syaifuddin et al. (2022) demonstrate the importance of needs analysis for identifying the reading requirements of vocational learners. Farida et al. (2025) similarly report that vocabulary, comprehension, technical terminology, motivation, and contextual relevance are important considerations in vocational material development. Nurhidaya et al. (2026) and Wang (2019) also highlights the importance of employment-oriented English in secondary vocational education. These findings suggest that vocational reading materials should begin with evidence about users rather than with a generic text.

The present study therefore treats needs analysis as the starting point for material development. The intention is not simply to make English texts easier, but to create texts whose linguistic demands are systematically controlled while their content remains meaningful to vocational learners. This orientation is consistent with materials-development principles that emphasize the alignment of learner needs, content, activities, and learning purposes (Tomlinson, 2023).

Graded readers are specially designed reading materials that present texts at different levels of linguistic difficulty to accommodate learners' language proficiency. They involve controlling vocabulary, grammatical structures, and text complexity to provide accessible reading input while

enabling learners to progressively develop their reading skills and lexical knowledge (Astutik et al., 2025). In addition to linguistic appropriateness, graded readers should consider learners' interests, preferences, and contextual backgrounds to create meaningful reading experiences. Astutik et al. (2025) emphasize the importance of aligning graded readers with learners' proficiency levels and incorporating culturally relevant content and multimedia features to enhance engagement. Similarly, Véliz-Campos et al. (2026) found that familiar vocabulary, accessible syntax, and comprehensible stories contributed to adolescent EFL learners' reading confidence and enjoyment. These findings highlight the importance of developing graded readers that balance linguistic accessibility, engaging content, and learner relevance to support meaningful and sustained English reading.

Grading, however, should not be equated with merely shortening or simplifying a difficult text. Its pedagogical value depends on systematic control of linguistic demands and on the relationship between text difficulty and learner ability. Nation (2015) and Nation and Waring (2019) emphasize the importance of accessible reading input in extensive reading, while Schmitt et al. (2011) demonstrate the importance of the proportion of known words for comprehension. Leitch (2023), in a comparative classroom study of decodable readers and leveled texts, found that both approaches supported reading fluency development, although students using decodable readers demonstrated greater improvements in words read per minute. This finding highlights the importance of controlling the linguistic features of reading materials and aligning them with learners' existing reading abilities rather than simply reducing text length. Recent work on e-graded readers further suggests that alignment between CEFR levels and learner proficiency is an important consideration in digital graded reading (Afdaliah et al., 2025).

The Common European Framework of Reference for Languages (CEFR) provides a systematic descriptive framework for understanding and organizing language proficiency. It classifies language users into six reference levels: A1 and A2 (Basic User), B1 and B2 (Independent User), and C1 and C2 (Proficient User). These levels are characterized by illustrative

descriptors that explain what learners can understand and accomplish through language use (Little, 2006). Beyond classifying proficiency, the CEFR serves as a reference for analyzing learners' needs, establishing learning objectives, designing instructional materials, and assessing language learning outcomes. Its descriptors cover different communicative language activities, including reading, and provide indicators of learners' ability to comprehend texts of varying complexity (Little, 2006; Li & Xu, 2024). In the development of graded readers, this framework offers a systematic basis for aligning text difficulty with learners' proficiency by considering vocabulary range, linguistic complexity, and reading comprehension demands. Accordingly, the present study adopts CEFR as a reference for organizing the progression of reading materials across five levels (A1, A2, A2–B1, B1, and B2), with specific linguistic and textual characteristics adapted to the vocational learning context.

For vocational learners, linguistic grading must be combined with contextual specificity. A text should not become so simplified that it loses the occupational or social meaning that makes it relevant. The present study therefore conceptualizes grading as a progression in text length, vocabulary range, sentence complexity, content depth, and reading demands.

The Sustainable Development Goals (SDGs) provide a meaningful content framework for connecting English learning with contemporary social, economic, and environmental issues. The United Nations (2015) established the 17 SDGs as part of the 2030 Agenda, creating a broad framework through which educational activities can engage learners with global challenges.

Within English language teaching, SDG integration can connect language learning with critical thinking, engagement, empathy, global awareness, and social responsibility. Azizah et al. (2025) identify these possibilities while also noting challenges involving resources, teacher preparation, and alignment between sustainability content and language objectives. Mambu (2023, 2025) similarly demonstrates that SDG issues can be integrated with language awareness and critical language work rather than treated as unrelated content. Research on SDG integration also indicates that

successful implementation depends on teachers' ability to connect sustainability goals with curricular and classroom objectives (Sulaiman et al., 2025). Thus, SDG content should be integrated with language and reading purposes.

Contemporary digital reading environments increasingly combine written language with visual and auditory representations. For EFL learners, illustrations can provide contextual cues while audio can provide access to pronunciation, rhythm, and the spoken realization of written language. Mayer and Fiorella (2021) explain that multimedia learning involves processing words and pictures and emphasize that multimedia elements should support meaningful cognitive processing rather than simply increase the amount of media. Supporting this perspective, Nabung (2023) highlights the integration of text, images, audio, video, and other interactive elements in EFL learning. The study found that interactive multimedia, considered alongside students' learning preferences, was significantly associated with their English learning outcomes. These findings suggest that integrating visual and auditory elements into digital reading materials can offer learners additional ways to engage with and process language input.

Multimodality is therefore understood in this study as the coordinated use of text, illustration, and audio to provide complementary access to meaning. This position is consistent with recent perspectives on multimodal communication, which emphasize the integration of visual, auditory, and textual elements in meaning-making (Al-Muttairi & Al-Alusi, 2025; Guo, 2023). Research also indicates that multimedia and interactive tools can support engagement and reading when they are deliberately connected to instructional purposes (Daulay & Dewi, 2025; Nur et al., 2025). In the context of EFL reading, Fathin and Mutiarani (2025) examined the integration of text, images, audio, and video in reading instruction, highlighting the use of multiple modes to support students' comprehension and engagement. The design challenge is therefore to make the modes complementary rather than redundant or distracting.. The design challenge is therefore to make the modes complementary rather than redundant or distracting.

In vocational settings, multimodality can also respond to learners' preference for contextualized digital materials. Farida et al. (2025) found that vocational students valued visual and contextual support in English learning materials. Similarly, Nurhidaya et al. (2026) emphasized the importance of aligning English instruction with students' future workplace communication needs, including the ability to understand workplace documents and information. These findings support the development of digital materials that combine visual support, relevant vocational contexts, and interactive features to make English learning more meaningful and applicable to students' future professional needs.

Generative AI has created new possibilities for language-material development because large language models can generate, adapt, and revise texts according to specified characteristics. This capability is particularly relevant when developers need multiple versions of related content at different proficiency levels. Arora et al. (2025) show that AI-supported content adaptation can contribute to language-learning material development, while Brierley and Ross (2025) illustrate possibilities for AI-enhanced extensive reading. Morita et al. (2025) further demonstrate how large language models and image-generation models can be combined in interactive digital reading environments.

The literature indicates that vocational English, graded reading, SDG-oriented ELT, multimodal learning, and Generative AI have each received growing research attention. However, these strands have often been examined separately. There is still limited research on how needs-based vocational English, systematically graded reading levels, SDG-oriented content, multimodal representation, and Generative AI-assisted production can be coordinated within a single pedagogically controlled development process.

The present study seeks to address this gap by developing SDG-based multimodal texts for graded readers in an Indonesian vocational school. To guide the development process, the study is guided by the following research questions:

1. How are SDG-based multimodal texts for graded readers developed with

Generative AI in an Indonesian vocational school?
2. What characteristics emerge from the development of SDG-based multimodal texts for graded readers with Generative AI support?
3. How do students perceive the use of the SDG-based multimodal texts?

Methods

Research Design

This study employed the ADDIE model as the primary research and development framework for developing SDG-based multimodal texts for graded readers. ADDIE consists of five interconnected stages: Analysis, Design, Development, Implementation, and Evaluation (Branch, 2009). The framework was selected because it provides a systematic pathway for translating user needs into product specifications, developing materials, implementing them in an authentic context, and using evaluation evidence to guide refinement.

In this study, ADDIE was treated as more than a sequence of production steps. Each stage generated information that constrained the next stage. The Analysis stage identified the needs of teachers and the vocational context; the Design stage converted those needs into a graded-reader blueprint; the Development stage used Generative AI as a supporting production tool within that blueprint; the Implementation stage exposed the product to intended users; and the Evaluation stage combined expert validation and student perceptions to identify strengths and areas for revision.

Participants and Context

The study was conducted at one of state vocational schools in Karawang, an Indonesian vocational high school. Four English teachers participated in the needs-analysis stage, 61 vocational students participated in the implementation and post-use evaluation, and one English education expert participated in product validation. The teachers were selected because of their direct experience teaching English in the vocational-school context. The students were the intended users of the developed graded readers.

Instruments

Three instruments were used: (1) a needs-analysis questionnaire for English teachers, (2) an expert-validation form, and (3) a student evaluation questionnaire. The teacher questionnaire contained background questions, 17 closed-ended statements using a five-point scale, and four open-ended questions. The statements addressed existing reading materials, SDG integration, graded readers, and multimodal/technology-supported materials.

The expert-validation form contained 10 five-point items covering material quality, language, SDG alignment and accuracy, reading pedagogy, comprehension questions, illustrations, audio, multimodal integration, and technology. It also included an overall judgment and open-ended comments. The student questionnaire contained 12 closed-ended statements and three open-ended questions concerning text difficulty, SDG understanding, language accessibility, illustrations, audio, multimodal integration, perceived SDG awareness, and perceived encouragement of SDG-supportive behavior.

ADDIE Procedure

1. Analysis

The Analysis stage identified the existing reading context, teachers' perceptions of student needs, relevant SDG themes, requirements for graded reading, and expectations concerning multimodal and AI-supported materials.

2. Design

The Design stage translated the needs-analysis findings into a blueprint specifying proficiency levels, text length, vocabulary and sentence characteristics, SDG content, reading focus, question types, and multimodal components.

3. Development

The Development stage produced the texts, illustrations, audio narration, and reading questions. ChatGPT supported text generation/adaptation and illustration production; Google AI Studio

supported audio narration. AI output was reviewed and revised against the predetermined specifications before integration.

4. Implementation

The developed materials were implemented with the intended student users in the vocational school. Students interacted with the graded texts and accompanying multimodal components and subsequently completed the evaluation questionnaire.

5. Evaluation

Evaluation involved expert validation and analysis of student responses. The evidence was used to identify the product's strengths and limitations and to establish priorities for refinement, particularly concerning audio delivery and technical usability.

Data Analysis

Quantitative data from the three questionnaires were analyzed descriptively using mean scores and percentages. Mean scores summarized responses to five-point items, while percentages described positive response distributions where appropriate. Open-ended responses were examined thematically to identify recurring needs, perceptions, strengths, and suggestions for improvement. The quantitative and qualitative findings were interpreted together so that product characteristics could be considered alongside user experience.

Findings and Discussions

The Development of SDG-Based Multimodal Texts Developed with Generative AI

Step 1 - Analysis: A Needs-Driven Starting Point

The Analysis stage established the need for more level-appropriate and innovative reading materials. Teachers rated the appropriateness of existing reading materials at M = 3.00/5.00, while the statement that students still experienced difficulty understanding English texts received M = 3.50. The

need for innovation in English reading instruction received M = 4.00. Although the teachers did not judge existing materials to be completely inappropriate, the combination of moderate appropriateness and continuing reading difficulty indicated a clear space for improvement.

These findings highlight the importance of developing vocational English materials based on teachers' understanding of learners' linguistic and contextual needs. Farida et al. (2025) identify vocabulary, reading comprehension, technical terminology, motivation, and contextual relevance as important considerations in vocational English material development. Similarly, Nurhidaya et al. (2026) emphasize the importance of aligning English instruction with vocational students' future workplace communication needs, including the ability to understand workplace documents and information. The present findings reinforce these perspectives by showing that teachers perceived a need for materials whose linguistic difficulty could be adjusted to students' proficiency. Consequently, differentiated reading levels became a central design requirement for the developed product.

The Analysis stage also generated strong support for SDG integration. The importance of SDG themes and their relevance to students' lives both received M = 4.00. Teachers rated the potential of SDGs to increase global awareness at M = 4.25, and both critical-thinking potential and workplace relevance received M = 4.50. Open-ended responses identified decent work, industry and innovation, responsible consumption and production, hospitality, workplace training, and skills as relevant contexts.

These findings explain why SDG 8, SDG 9, and SDG 12 were selected. The selection was not imposed as a generic sustainability package; it emerged from the vocational context. SDG 8 could be localized through employment, entrepreneurship, and workplace practices; SDG 9 through technology, infrastructure, and innovation; and SDG 12 through consumption, production, and waste. This needs-based localization is important because SDG integration becomes pedagogically stronger when sustainability content is

connected with learners' contexts and language objectives (Azizah et al., 2025; Mambu, 2023; Sulaiman et al., 2025).

Teachers also showed strong support for graded reading. The need to adjust text difficulty to students' proficiency received M = 4.25, while the potential of graded readers to increase reading interest and the need for more graded materials each received M = 4.00. Three of the four teachers reported that they had not previously used graded readers, indicating a gap between perceived need and classroom experience. This is a significant finding for development: teachers recognized the value of level-appropriate reading even though graded readers were not yet a routine resource in their practice.

This result is consistent with recent research emphasizing the importance of matching graded reading materials to learners' proficiency. Astutik et al. (2025) highlight the need to align graded readers with learners' language levels while incorporating relevant cultural content and multimedia features. Similarly, Véliz-Campos et al. (2026) report that familiar vocabulary, accessible syntax, and comprehensible stories contribute to adolescent EFL learners' reading confidence and enjoyment. These studies support the teachers' expectation that reading materials should be linguistically accessible without losing their meaningful content.

The strongest support concerned multimodality and technology. Images and audio, AI support for English reading, and the need for an SDG-based graded-reader application each received M = 4.50. Teachers' open-ended responses further requested translation, pronunciation support, audio, text-based exercises, and a level-based system. These findings indicate that teachers expected technology to function as instructional support rather than merely as a visual addition to reading materials. This expectation is consistent with Farida et al. (2025), who identify the importance of contextualized materials in vocational English learning, and with recent research emphasizing the pedagogical contribution of multimedia and interactive resources (Daulay & Dewi, 2025; Nur et al., 2025). Overall, the needs analysis established five important development requirements: linguistic accessibility, vocational relevance, SDG integration, multimodal support, and technology-assisted production.

Step 2 - Design: Translating Needs into a Five-Level Blueprint

The Analysis findings were translated into a five-level blueprint: A1, A2, A2–B1, B1, and B2. The progression was defined through text length, vocabulary range, sentence complexity, content depth, and reading focus. Lower levels emphasized accessible language and explicit information, while higher levels gradually introduced more varied vocabulary, complex syntax, inference, analysis, interpretation, and evaluation.

Level	**Text length**	**Vocabulary & sentence characteristics**	**Content focus**	**Reading focus**
A1	150–200 words	Very common/concrete vocabulary; short, simple sentences	SDG-related people, objects, places, or activities	Finding information
A2	200–250 words	Common vocabulary + topic terms; simple and compound sentences	Simple functions and benefits	Understanding information
A2–B1	250–300 words	More varied topic vocabulary; simple, compound, and basic complex sentences	Problems and solutions	Connecting information
B1	300–400 words	More varied/topic-specific vocabulary; complex sentences	Problems, causes, and impacts	Analyzing information
B2	400–450 words	Wider/partly abstract vocabulary; varied complex sentences	Issues, relationships, and broader impacts	Interpreting and evaluating information

Table 1. Design specifications of the graded readers.

The blueprint resulted in 15 texts, three at each level, distributed across the three selected SDGs.

Level	**SDG 8: Decent Work**	**SDG 9: Industry & Innovation**	**SDG 12: Responsible Consumption**
A1	A Farmer in My Village	A Solar Panel at School	My Reusable Bottle
A2	A Traditional Weaver	A New Bridge to Our Village	A Traditional Market
A2–B1	A Small Local Business	A Smart Village	Reducing Food Waste at School
B1	Young Entrepreneurs	A Local Innovation	From Waste to Useful Products
B2	Green Jobs for the Future	Innovation for Sustainable Communities	Responsible Consumption in Modern Society

Table 2. Distribution of the 15 reading texts by level and SDG.

Each text was accompanied by three questions addressing content, language, and comprehension, producing 45 questions in total. The question design also followed the level progression. Lower levels emphasized explicit information, vocabulary, reference, and supporting detail, while higher levels increasingly required inference, analysis, interpretation, and evaluation.

The design demonstrates that grading was treated as a multidimensional process rather than simply reducing text length. This approach is consistent with Astutik et al. (2025), who emphasize the importance of aligning graded readers with learners' proficiency levels and controlling linguistic demands.

It also corresponds with Véliz-Campos et al. (2026), whose findings highlight the contribution of familiar vocabulary and accessible syntax to learners' reading confidence and enjoyment. Furthermore, Afdaliah et al. (2025) emphasize the importance of aligning CEFR-based e-graded readers with EFL learners' reading proficiency. Together, these studies support the use of proficiency-based progression as a foundation for designing accessible digital reading materials.

Step 3 - Development: Generative AI within Human Pedagogical Control

During Development, the blueprint functioned as a set of constraints for AI-assisted production. ChatGPT was used to generate and adapt reading texts according to proficiency level, target length, vocabulary characteristics, sentence structures, SDG theme, vocational context, and reading focus. ChatGPT was also used to generate illustrations corresponding to the texts, while Google AI Studio was used to create audio narration.

The important feature of this workflow was that AI did not determine the pedagogical specifications. The researchers first established what the materials needed to achieve and then used AI to assist with production. AI-generated content was reviewed and revised before being treated as part of the product. This approach is consistent with Arora et al. (2025), who demonstrate the potential of AI-supported content adaptation while retaining the need for human evaluation. It also responds to concerns about accuracy, contextual fit, and overdependence identified in AI-supported language-learning research (Siregar & Hz, 2025; Wahyuningtyas & Suparmi, 2026).

The development process therefore positions Generative AI as a production assistant rather than an autonomous instructional designer. This distinction matters particularly for graded reading because an apparently fluent AI-generated text may still contain vocabulary, syntax, discourse patterns, or conceptual density that exceed the intended level. Similarly, an AI-generated SDG text may be factually plausible but insufficiently localized to students' vocational lives. Human review was consequently necessary to

maintain alignment among CEFR level, SDG theme, vocational relevance, and reading purpose.

The multimodal components were developed with complementary functions. Written text provided the primary reading input, illustrations represented key content visually, and audio provided access to pronunciation and spoken language. This design is consistent with recent perspectives on multimodal communication that emphasize the integration of visual, auditory, and textual elements in meaning-making (Al-Muttairi & Al-Alusi, 2025; Guo, 2023). It also corresponds with Fathin and Mutiarani (2025), who examine the integration of text, images, audio, and video in reading instruction. In the present product, however, the three modes were deliberately limited to text, illustration, and audio to provide complementary access to meaning without unnecessarily distracting learners from the reading content.

Step 4&5 - Implementation and Expert Evaluation

Before implementation with students, the product was evaluated by an English education expert. The overall validation mean was 4.40/5.00, and the overall judgment was "Appropriate." Accuracy and relevance of SDG information, illustrations, audio, and multimodal integration each received 5/5. Material quality, language, SDG-topic alignment, reading pedagogy, comprehension questions, and technology each received 4/5.

Aspect	**Score**
Material	4
Language	4
SDG topic alignment	4
Accuracy and relevance of SDG information	5
Reading pedagogy	4
Comprehension questions	4
Illustrations	5
Audio	5
Multimodal integration	5
Technology	4

Overall mean	4.40

Table 3. Expert validation results.

The expert described the product as innovative and considered the SDG topics, illustrations, and audio useful. The recommendations were nevertheless specific: audio should ideally be integrated directly into the application rather than accessed through an external link or barcode, and future versions could include interactive games and word-level translation. These comments are important because they distinguish product appropriateness from product completeness. A high validation score does not mean that no further development is required; instead, it identifies a viable product whose usability and functionality can still be improved.

Characteristics Emerge from the Development

Five interrelated characteristics emerged from the development process. First, the texts were systematically graded. The five levels differed not only in length but also in vocabulary range, sentence complexity, content depth, and cognitive reading demands. Lower levels emphasized concrete information and basic comprehension, whereas higher levels introduced inference, analysis, interpretation, and evaluation. This multidimensional progression reflects the importance of matching reading materials with learners' proficiency levels, as emphasized by Astutik et al. (2025) and Afdaliah et al. (2025). It also corresponds with Véliz-Campos et al. (2026), who highlight the contribution of accessible vocabulary and syntax to adolescent EFL learners' reading confidence and enjoyment. The findings therefore suggest that effective grading involves balancing linguistic accessibility with progressive reading demands.

Second, the texts were vocationally contextualized through SDG content. SDG 8 connected reading to work, employment, and entrepreneurship; SDG 9 connected it to innovation, technology, and infrastructure; and SDG 12 connected it to consumption, production, and waste. The topics were localized through familiar settings such as villages,

schools, markets, businesses, workplaces, and communities. This localization responds to the needs-analysis finding that work-related SDGs were highly relevant to vocational learners (M = 4.50) and supports the argument that sustainability content should be connected with students' immediate and future contexts (Azizah et al., 2025; Mambu, 2025).

Third, the materials were multimodal rather than text-only. Each reading experience combined written text with illustration and audio. The modes were intended to perform complementary functions: text provided linguistic input, illustrations provided contextual or conceptual cues, and audio provided access to pronunciation and spoken language. The expert's 5/5 ratings for illustrations, audio, and multimodal integration support the coherence of this design. This characteristic is consistent with recent studies that emphasize the integration of visual, auditory, and textual resources in language learning (Al-Muttairi & Al-Alusi, 2025; Guo, 2023). It also corresponds with Fathin and Mutiarani (2025), who highlight the relevance of multimodal resources in reading instruction. Importantly, the present study did not include media simply to make the materials technologically rich. Instead, each mode was assigned a specific instructional function.

Fourth, the materials were supported by Generative AI but remained human-designed. The defining feature was not simply the presence of AI-generated content; it was the position of AI within the development sequence. Needs analysis established the requirements, the blueprint established constraints, AI supported production, and human review determined final suitability. This workflow is compatible with emerging research that treats GenAI as a potentially powerful development aid while emphasizing human judgment and critical oversight (Arora et al., 2025; Brierley & Ross, 2025; Morita et al., 2025).

Fifth, the product combined reading content with structured questions. The 45 questions were organized around content, language, and comprehension, with increasing cognitive demands across levels. This feature responded directly to teachers' requests for text-based exercises and ensured that the product was more than a digital collection of stories. Materials-

development research emphasizes that content, learning purposes, and activities should be designed as an integrated whole (Tomlinson, 2023). The resulting product is therefore better understood as a structured multimodal reading series than as a collection of AI-generated texts.

Taken together, these characteristics show how the different design decisions reinforced one another. Grading addressed linguistic accessibility; SDGs supplied meaningful content; multimodality provided complementary routes to meaning; structured questions connected texts to reading purposes; and Generative AI increased production support without replacing human pedagogical judgment. This coherence is the principal product-level contribution of the development.

Students Perception on the Use of the SDG-Based Multimodal Texts

Sixty-one students completed the post-use questionnaire. The overall mean was 3.91/5.00, indicating generally positive perceptions. The highest-rated statement concerned audio helping students understand SDG information (M = 4.13; 86.9% agreed or strongly agreed). This was followed by the text helping students understand the SDG issue (M = 4.02; 83.6%) and illustrations making SDG information easier to understand (M = 3.98; 83.6%).

No.	Evaluation indicator	Mean	Agree/Strongly Agree
1	Text difficulty was appropriate for students' English ability	3.82	70.5%
2	Text helped students understand the SDG issue	4.02	83.6%
3	Language helped students understand SDG information	3.92	78.7%
4	Illustrations helped students understand the SDG issue	3.92	78.7%
5	Illustrations made SDG information easier to understand	3.98	83.6%

6	Illustrations and text complemented each other	3.95	86.9%
7	Audio helped students understand SDG information	4.13	86.9%
8	Audio speed was comfortable	3.59	60.7%
9	Audio helped students understand English pronunciation	3.90	77.0%
10	Text, illustrations, and audio helped students understand SDG issues better	3.92	80.3%
11	Graded reader increased students' awareness of SDG issues	3.85	82.0%
12	Graded reader encouraged SDG-supportive behavior	3.93	82.0%
Overall		3.91	—

Table 4. Students' perceptions of the SDG-based multimodal graded readers.

The positive response to audio is particularly noteworthy because audio and pronunciation support were identified as needs during Analysis. The student evaluation therefore provides convergence across the ADDIE stages: a feature requested by teachers was subsequently perceived as useful by students. The result is consistent with the intended multimodal function of audio, which adds an auditory representation of the language while the written text remains the principal reading input.

Students also responded positively to the relationship between text and illustration. The statement that illustrations and text complemented each other received 86.9% positive responses. This finding is important because the value of multimodality depends on how different modes contribute to meaning. Rather than functioning as unrelated additions, illustrations were designed to provide visual support for the information presented in the texts. The students' responses suggest that they generally perceived the visual and written modes as complementary. This interpretation is consistent with Fathin

and Mutiarani (2025), who examine the use of multiple modes in reading instruction, and with Al-Muttairi and Al-Alusi (2025), who discuss the integration of visual, auditory, and textual elements in language learning.

The combined use of text, illustration, and audio received M = 3.92, with 80.3% of students giving positive responses. This finding supports the product's multimodal design at the level of user perception. It also corresponds with recent EFL research suggesting that multimedia and interactive resources can support engagement and reading when they are linked to instructional purposes (Daulay & Dewi, 2025; Nur et al., 2025). Importantly, the positive perception should not be attributed to the mere presence of technology; it is more plausibly related to the complementary functions assigned to each mode.

The student responses concerning SDG awareness and supportive behavior were also positive. The awareness item received M = 3.85, with 82.0% positive responses, while the item concerning encouragement of SDG-supportive behavior received M = 3.93, also with 82.0% positive responses. These results should be interpreted carefully. They indicate students' perceptions of what the materials contributed, not objective changes in awareness or behavior. The study did not include pre- and post-measures of sustainability knowledge or behavior, so causal claims are not warranted.

The lowest-rated item was audio speed (M = 3.59), with only 60.7% positive responses. Open-ended feedback reinforced this finding: some students considered the audio too fast, while others reported slow loading, lag, system errors, and less responsive sections of the application. Thus, the evaluation reveals a useful distinction between pedagogical value and technological usability. Students could perceive audio as helpful while simultaneously experiencing its delivery speed and technical access as barriers.

The technical findings also converge with the expert recommendations for direct audio playback and improved functionality. Nur et al. (2025) similarly emphasize that the contribution of technology depends on usability, instructional structure, and the connection between digital tools

and learning goals. For the present product, the clearest refinement priorities are therefore adjustable or slower audio, direct in-application audio playback, improved loading and responsiveness, and stronger SDG-oriented activities.

The open-ended responses produced three recurring themes: English learning, multimodal support, and technical usability. First, students reported learning vocabulary, expressions, grammar, reading skills, and pronunciation. Several students noted that the graded levels helped them recognize the relationship between text difficulty and their proficiency. Some also reported learning about SDGs and gaining broader awareness of sustainability-related issues. These responses are consistent with recent research highlighting the importance of accessible vocabulary, comprehensible syntax, and appropriate proficiency levels in graded reading (Astutik et al., 2025; Véliz-Campos et al., 2026). They also suggest that students recognized the relevance of the graded structure to their reading experience.

Second, students generally described the combination of text, illustrations, and audio as helpful. Illustrations were perceived as making the content easier to visualize, while audio helped students with pronunciation and provided an alternative route to the material when reading was difficult. This supports the interpretation that the three modes served complementary functions rather than simply duplicating the same information.

Third, technical usability was the dominant weakness in the qualitative responses. Students mentioned slow loading, lag, system errors, and sections that were difficult to press or use. These comments are important because they show that material quality in a digital environment has at least two dimensions: pedagogical quality and technical accessibility. A strong instructional design cannot fully realize its potential if learners encounter unnecessary barriers when accessing its components.

Implications for Vocational English Material Development

The findings suggest that developing vocational English materials requires attention to learners' needs, appropriate reading levels, relevant SDG topics, and multimodal support. Teachers and material developers can use

Generative AI to produce reading materials by providing clear instructions about CEFR levels, vocabulary, text length, vocational contexts, and learning objectives. However, AI-generated materials still require human review to ensure their quality and suitability. Schools should also provide adequate technological facilities and support teachers' digital skills to ensure that digital materials are accessible and easy to use.

Limitations and Future Research

This study involved only four English teachers and 61 students from one vocational school, which limits the generalizability of the findings. Student evaluation was also based on self-reported perceptions rather than direct measurements of learning outcomes. In addition, the implementation period was relatively limited. Future research should involve more schools and vocational programs and use pre- and post-tests or longitudinal designs to examine the effects of the materials on reading comprehension, vocabulary, motivation, and SDG-related learning. Further studies could also explore adjustable audio, interactive features, and AI-supported text adaptation.

Conclusion

his study developed SDG-based multimodal texts for graded readers through the ADDIE model in an Indonesian vocational school. The product consisted of 15 texts across five proficiency levels, covering SDG 8, SDG 9, and SDG 12, with supporting illustrations, audio narration, and reading questions. Generative AI supported the production process under human pedagogical supervision. Expert validation reached 4.40/5.00, while students' overall perceptions were positive (M = 3.91/5.00). Although audio speed and technical usability require further improvement, the findings indicate that the materials were considered appropriate and useful. Overall, the study demonstrates the potential of integrating vocational contexts, SDG themes, multimodality, and Generative AI into English reading material development. Further research is needed to investigate actual learning outcomes and the use of the materials in other educational contexts.

Acknowledgment

This research was funded by the Hibah Internal Penelitian Pemula of the Institute for Research and Community Service (LPPM), Universitas Singaperbangsa Karawang. The authors gratefully acknowledge the financial support provided for the conduct of this research.

References

Afdaliah, N., Nurchalis, N. F., Aliyu, M. M., & Jakob, J. C. (2025). Aligning CEFR E-Graded Readers with EFL Learners' Reading Proficiency: A Mixed-Methods Study from West Sulawesi, Indonesia. IJORER, 6(4), 1124–1135. https://doi.org/10.46245/ijorer.v6i4.935

Albay, M. (2017). The benefits of graded reading. International Journal of Social Sciences & Educational Studies, 3(4), 177–180. https://doi.org/10.23918/ijsses.v3i4p177

Aradi, C. (2021). English L2 thematic vocabulary acquisition through graded readers: A single-participant study. EduLingua, 7(1), 25–45. https://doi.org/10.14232/edulingua.2021.1.2

Arora, J., Elgort, I., & Zhao, J. (2025). Content adaptation for language learning: A hybrid AI approach. EuroCALL Review, 32(2), 169–182. https://doi.org/10.4995/eurocall.2025.23898

Azizah, N. L., Mustofa, A., & Retnaningdyah, P. (2025). Sustainable development in English language teaching: Prevailing trends, challenges, and outcomes. Education and Human Development Journal, 10(1), 1–16. https://doi.org/10.33086/ehdj.v10i1.6763

Branch, R. M. (2009). Instructional design: The ADDIE approach. Springer. https://doi.org/10.1007/978-0-387-09506-6

Brierley, M., & Ross, G. (2025). AI-enhanced extensive reading: Empowering learners as content creators. In Advancing CALL: New research agendas—EUROCALL 2025 short papers. https://doi.org/10.4995/EuroCALL2025.2025.21274

Buck, R. H., Abu-Hassan, J., Jimenez, M., & AlDarwish, S. (2024). Implementing SDGs in a pre-sessional English classroom. Open Scholar Teaching and Learning, 3(1), 96–106.

Daulay, E., & Dewi, U. (2025). Reading comprehension through multimodal literacy approach for English education. ETERNAL: English Teaching Journal, 16(2), 397–404. https://doi.org/10.26877/eternal.v16i2.1910

Farida, A., Yusuf, A., Kurniawan, E., Syafrudin, T., Baladina, A., Faidah, D. N., & Ahmad, F. (2025). Designing e-comic for English learning based on needs analysis: A mixed-methods study of vocational high school students. Linguistics and English Language Teaching Journal, 13(2), 350–362.

Kara, B. (2019). The effect of graded readers on reading comprehension skill of EFL students. Language Teaching and Educational Research, 2(2), 160–172. https://doi.org/10.35207/later.554716

Kress, G. (2020). Multimodality: A social semiotic approach to contemporary communication (2nd ed.). Routledge.

Mambu, J. E. (2023). Embedding Sustainable Development Goals into critical English language teaching and learning. Critical Inquiry in Language Studies, 20(1), 46–76. https://doi.org/10.1080/15427587.2022.2099863

Mambu, J. E. (2025). Corpus-informed and critical approaches to addressing Sustainable Development Goals (SDGs). LLT Journal: A Journal on Language and Language Teaching, 28(2), 587–613. https://doi.org/10.24071/llt.v28i2.12023

Mayer, R. E., & Fiorella, L. (2021). Introduction to multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), The Cambridge handbook of multimedia learning (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781108894333.003

Morita, R., Watanabe, K., Zhou, J., Dengel, A., & Ishimaru, S. (2025). GenAIReading: Augmenting human cognition with interactive digital textbooks using large language models and image generation models. In

Proceedings of the Augmented Humans International Conference 2025 (pp. 289–301). ACM. https://doi.org/10.1145/3745900.3746066

Nation, P. (2015). Principles guiding vocabulary learning through extensive reading. Reading in a Foreign Language, 27(1), 136–145.

Nation, I. S. P., & Waring, R. (2019). Teaching extensive reading in another language. Routledge. https://doi.org/10.4324/9780367809256

Nur, D. R., Rohman, A., Geroda, G. B., Pane, W. S., Rachman, D., Pasaribu, I. F., Auliasari, Y., Hiping, M. G. A. K. Y., & Rohmah, A. (2025). Multimedia and interactive tools foster engagement in English as a foreign language critical reading. Academia Open, 10(2). https://doi.org/10.21070/acopen.10.2025.11118

Octavia, H., Widiati, U., & Irawati, E. (2019). Vocational students' perceptions of Mobile Assisted Language Learning (MALL) materials. Jurnal Pendidikan: Teori, Penelitian, dan Pengembangan, 4(10), 1347–1351. https://doi.org/10.17977/jptpp.v4i10.12842

Reiber-Kuijpers, M., Kral, M., & Meijer, P. (2021). Digital reading in a second or foreign language: A systematic literature review. Computers & Education, 163, 104115. https://doi.org/10.1016/j.compedu.2020.104115

Robb, T. N., & Ewert, D. (2024). Classroom-based extensive reading: A review of recent research. Language Teaching, 57(3), 295–324. https://doi.org/10.1017/S0261444823000319

Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. The Modern Language Journal, 95(1), 26–43. https://doi.org/10.1111/j.1540-4781.2011.01146.x

Siregar, A. R. F., & Hz, B. I. R. (2025). Undergraduate students' perceptions on utilizing DeepSeek AI as a tool in writing journal articles. Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 12(2), 713–729. https://doi.org/10.22219/celtic.v12i2.41375

Suk, N. (2016). The effects of extensive reading on reading comprehension, reading rate, and vocabulary acquisition. Reading Research Quarterly, 52(1), 73–89. https://doi.org/10.1002/rrq.152

Sulaiman, M. A. B. A., Khallaf, N. M., & Abugohar, M. A. (2025). Integrating UN SDGs in ELT: How far tertiary EFL teachers in Oman are successful. World Journal of English Language, 15(8), 319–331. https://doi.org/10.5430/wjel.v15n8p319

Syaifuddin, S., Asi, N., & Karani, E. (2022). English reading materials in vocational high school: A need analysis. Jurnal Bahasa dan Sastra, 10(1). https://doi.org/10.24036/jbs.v10i1.116662

Tomlinson, B. (2023). Developing materials for language teaching. Bloomsbury Academic.

United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development.

Walsh, M. (2015). Multimodal literacy: Researching classroom practice. In J. Rowsell & K. Pahl (Eds.), The Routledge handbook of literacy studies. Routledge.

Wahyuningtyas, R. T., & Suparmi, S. (2026). Exploring the role of ChatGPT voice conversation feature in improving EFL students' speaking self-efficacy. Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics, 13(1), 195–212. https://doi.org/10.22219/celtic.v13i1.44287

Wang, D. (2019). A preliminary study of employment-oriented English teaching model in secondary vocational schools. In Advances in Social Science, Education and Humanities Research (Vol. 286, pp. 634–638). Atlantis Press.

Little, D. (2006). The Common European Framework of Reference for Languages: Content, purpose, origin, reception and impact. *Language Teaching*, 39, 167–190. https://doi.org/10.1017/S0261444806003557

Li, G., & Xu, H. (2024). CEFR-informed learning, teaching and assessment: A practical guide [Book review]. *Studies in Second Language Learning and Teaching*, 14(2), 365–370. https://doi.org/10.14746/ssllt.32669

Nurhidaya, Badeni, & Kasmaini. (2026). Aligning English instruction with workplace communication: A needs analysis of vocational high school students. *GHAITSA: Islamic Education Journal,* 7(2). https://doi.org/10.62159/ghaitsa.v7i2.2961

Astutik, Y. P., Hartono, R., Saleh, M., & Widiati, U. (2025). Development and validation of English graded readers' e-books based on folklores for extensive reading classes for university students. English Review: Journal of English Education, 13(3), 1043–1056. https://doi.org/10.25134/erjee.v13i3.109

Véliz-Campos, M., Castel Pérez, G. A., & Veliz, L. (2026). Graded readers as scaffolded stories: Self-efficacy, enjoyment, and narrative engagement in adolescent EFL reading. Reading & Writing Quarterly. https://doi.org/10.1080/10573569.2026.2735881

Leitch, T. (2023). Decodable readers versus leveled text [Master's research project, Minnesota State University Moorhead]. RED: A Repository of Digital Collections. https://red.mnstate.edu/thesis/868

Nabung, A. (2023). Improving EFL learning outcomes by interactive multimedia use and learning style. *Premise: Journal of English Education and Applied Linguistics*, 12(1), 256–274. https://doi.org/10.24127/pj.v12i1.6223

Al-Muttairi, F. Z. S., & Al-Alusi, A. H. S. H. (2025). Multimodal communication in ESL learning: Examining the integration of visual, auditory, and textual elements in digital media with a focus on quality education (SDG 4). *Journal of Lifestyle and SDGs Review*, 5(3), e04773. https://doi.org/10.47172/2965-730X.SDGsReview.v5.n03.pe04773

Fathin, S. F. F., & Mutiarani. (2025). The role of multimodal approach in enhancing reading comprehension skills in junior high school students.

Jurnal Riset Rumpun Ilmu Bahasa, 4(3), 603–612. https://doi.org/10.55606/jurribah.v4i3.7717

Guo, X. (2023). Multimodality in language education: Implications of a multimodal affective perspective in foreign language teaching. *Frontiers in Psychology*, 14. https://doi.org/10.3389/fpsyg.2023.1283625

Nurhidaya, et al. (2026). Aligning English instruction with workplace communication: A needs analysis of vocational high school students. *GHAITSA: Islamic Education Journal*, 7(2). https://doi.org/10.62159/ghaitsa.v7i2.2961