Vol. 7, No. 1, 2026, pp. 764-772
DOI: https://doi.org/10.29210/992650

Contents lists available at Elora Center

Lentera Negeri

Journal homepage: https://lentera.eloracenter.org/lentera

The mosaik model: an expert-validated model integrating toulmin argumentation and multimodal literacy for vocational writing instruction

Regina Fitrie*), Andoyo Sastromiharjo, Khaerudin Kurniawan
Indonesian Language Education Study Program, Postgraduate School, Universitas Pendidikan Indonesia, Indonesia

Article Info

Article history:
Received Jul 04th, 2026
Revised Jul 04th, 2026
Accepted Jul 13th, 2026

Keyword:
Learning model,
Multimodal text,
Digital literacy,
Argumentative writing,
Vocational high school

ABSTRACT

This study develops and validates the MOSAIK Model, a five-phase digital-literacy-based learning model for argumentative multimodal writing among vocational high school students. Named after the Indonesian word for "mosaic," the model integrates verbal, visual, and digital elements into coherent, evidence-based arguments. Using the ADDIE framework, the study covered the analysis, design, development, and implementation stages, with data collected through writing tests, questionnaires, expert validation, and classroom observations. The analysis of 373 students showed a generally good baseline writing ability (M = 14.60, 73.0%), although evidence-based argumentation was the weakest component. This indicated the need for a structured instructional model rather than remedial teaching. A digital-literacy self-efficacy questionnaire demonstrated good reliability (Cronbach's α = 0.849). Expert validation of the MOSAIK instructional products yielded a mean feasibility score of 86.5% ("Highly Feasible"), with inter-rater agreement ranging from weak to good (ICC = 0.32–0.66). The model was implemented in limited (n = 67) and broad (n = 306) trials. The broad trial revealed significant differences in implementation across study programs, $F(2, 303) = 20.44$, $p < .001$, $\eta^2 = 0.12$, while multimodal production remained the weakest phase. Overall, the findings indicate that the MOSAIK Model is a feasible and structurally valid instructional model, although further refinement is needed to strengthen multimodal production support and standardize group assessment procedures.

Corresponding Author:
Regina Fitrie,
reginafitrie@upi.edu

Introduction

The development of information and communication technology in the twenty-first century has transformed the way people communicate and produce meaning, so that communication no longer relies solely on verbal text but has evolved into multimodal communication that combines written text, images, infographics, audio, video, and animation (Al Fajri, 2018). Multimodal text is understood as text that constructs meaning through the integrated combination of various semiotic modes (Kress & van Leeuwen, 2001), in line with the multiliteracies framework, which asserts that literacy is no longer confined to conventional printed text (New London Group, 1996). A bibliometric study of international publications between 1997 and 2023 revealed that research on multimodal discourse has grown rapidly since 2012, with a primary focus on education, communication, and digital media (Liu et al., 2024). Cross-national research further confirms this trend: studies on digital multimodal composing in writing instruction have developed rapidly and have been shown to expand students' meaning-making skills beyond verbal text alone (Jiang, 2017; Dahlström, 2022).

In Indonesia, this demand is addressed through the Merdeka Curriculum, which establishes the ability to view, present, and produce multimodal texts as a mandatory competency within the Indonesian Language Learning Outcomes for Phases E and F (Badan Standar, Kurikulum, dan Asesmen Pendidikan, 2025). This competency is strategically relevant for vocational high school students in particular, as it relates directly to workplace demands—ranging from laboratory analysis reports for the Analisis Kimia (Chemical Analysis) program, technical documentation and digital product presentations for Rekayasa Perangkat Lunak (Software Engineering) and Teknik Komputer dan Jaringan (Computer Networking), to argumentative visual products for Desain Komunikasi Visual (Visual Communication Design). Developing this competency also cannot be separated from the ability to read critically across multiple sources (syntopical reading)—the ability to read and synthesize various sources into a coherent argument (Adler, 1972)—and the ability to construct evidence-based arguments using the structure of claim, evidence, and warrant (Toulmin, 1958).

Nevertheless, the gap between these policy demands and classroom realities remains wide. The results of the 2022 Programme for International Student Assessment placed Indonesia at a concerning level in reading literacy and critical writing (OECD, 2023), while the 2022 National Assessment revealed that most secondary school students still struggle to compose critically reasoned texts that integrate information from multiple sources (Kemendikbudristek, 2022). Available teaching materials generally still rely on conventional text-based modules (Sahidah & Kirana, 2021), while teachers' ability to design digital teaching materials that systematically facilitate the stages of multimodal writing remains limited (Yuniarti et al., 2026). Teachers tend to interpret technology integration merely as the use of digital devices for conventional instructional purposes, rather than as an effort to develop a new genre of writing grounded in multiliteracies (Howell et al., 2017; Rahma et al., 2023). The use of generative artificial intelligence in the writing process further introduces new challenges regarding originality and academic integrity that must be anticipated pedagogically (Pratiwi et al., 2024).

A number of prior studies have made contributions, although these remain partial. In a formative experiment conducted over 11 weeks in high school English classrooms, an intervention combining claim-evidence-warrant argument construction with digital multimodal tools broadened students' understanding of argumentation, as evidenced through classroom observations, interviews, and analysis of student-produced products (Howell et al., 2017). Digital multimodal tools can improve the organizational quality of argumentative writing among foreign language learners, with a large effect size (Hutasuhut et al., 2025). A genre-based approach combined with digital storytelling effectively improves students' multimodal literacy (Nur Fajriah et al., 2026), while a multiliteracies approach has been shown to improve digital, visual, and critical literacy (Fadli, 2025). Internationally, drawing on interviews with teachers assessing students' digital video documentaries in an English-for-science course, a process-based assessment model comprising four stages—pre-design, design, sharing, and reflecting—has been proposed for digital multimodal composing in second language writing instruction (Hafner & Ho, 2020); digital multimodal composing has been found to increase English learning investment among ethnic minority students (Jiang, Yang, & Yu, 2020); digital storytelling has been shown to facilitate language and digital literacy learning among students from refugee backgrounds (Kendrick et al., 2022); and inquiry-based multimodal writing on climate change issues has been found to improve conceptual understanding and critical awareness among English as a foreign language learners (Silvhiany et al., 2026). However, these studies were generally conducted in general senior high school or foreign language learning contexts, focused on only a single aspect—multimodality, argumentation, or digital literacy—in isolation, and had not yet produced a complete instructional model with accompanying teaching materials that was systematically developed, expert-validated, and directly trialed in vocational high school classrooms across diverse study programs.

Based on this gap, the present study develops the MOSAIK Model, a digital literacy-based learning model for writing argumentative multimodal texts for vocational high school students, using the ADDIE research and development framework (Branch, 2010). The name MOSAIK—the Indonesian word for "mosaic"—was chosen to capture the model's core principle: that a coherent, evidence-based multimodal argument is assembled from carefully arranged verbal, visual, and digital pieces, much as a mosaic is composed of individual fragments unified into a meaningful whole. A preliminary needs assessment conducted for this study (N = 373 students across three vocational study programs) found that students' baseline argumentative multimodal writing ability was, on average, already at a "Good" level (73.0%). However, performance was uneven across the five core sub-skills that make up this competency, with reasoning and evidence-based argumentation construction comparatively the weakest (70.6%), and no single program or class showing a consistent profile. This suggests that the central problem facing vocational writing instruction is not primarily a uniform skills deficit, but the absence of a structured, program-agnostic pedagogical model capable of scaffolding argumentative and multimodal competence consistently across classes with heterogeneous

starting profiles. Addressing this gap in instructional consistency and scalability, rather than remediating a presumed deficit, is the central rationale for developing the MOSAIK Model. The research questions addressed in this study are: (1) how consistent is the learning-needs profile for argumentative multimodal text writing across vocational study programs, as measured through students' baseline writing performance, self-reported digital literacy, and teacher-reported instructional practice; (2) what five-phase syntax, teaching materials, and assessment instruments can be designed to address the identified needs; (3) to what extent do expert validators rate the resulting MOSAIK Model as feasible, and how consistent is that rating across validators; and (4) when implemented under real classroom conditions across vocational study programs, what is the quality of the MOSAIK Model's implementation, and which factors differentiate full from partial implementation. Specifically, this study aims to: (a) profile the learning-needs consistency for argumentative multimodal writing across vocational study programs; (b) design and expert-validate the MOSAIK Model's five-phase syntax and accompanying teaching materials; (c) measure the model's feasibility through inter-validator-verified expert judgment; and (d) evaluate the quality and consistency of the model's implementation across heterogeneous classroom conditions, as a concrete contribution to Indonesian language writing instruction responsive to the literacy demands of the twenty-first century. Theoretically, the development of this model draws on the integration of five frameworks: social semiotics, which views text as a combination of modes of meaning representation and positions writing as one design choice among several available semiotic resources (Kress & van Leeuwen, 2001; Bezemer & Kress, 2008; Jewitt, 2008); multiliteracies (New London Group, 1996); syntopical reading (Adler, 1972); the claim-evidence-warrant model of argumentation (Toulmin, 1958); and the instructional model quality criteria of validity, practicality, and potential positive response (Nieveen, 1999).

Method

This study employed a research and development approach using the cyclical and iterative ADDIE framework (Branch, 2010), encompassing the Analyze, Design, Development, and Implementation stages, with Evaluate carried out formatively at each stage. In the Analyze stage, data were collected from tenth- and eleventh-grade students in the Rekayasa Perangkat Lunak (Software Engineering), Teknik Komputer dan Jaringan (Computer Networking), and Analisis Kimia (Chemical Analysis) study programs at a public vocational high school in Bandung City (N = 373 for the writing performance test), as well as 27 Indonesian language teachers from various schools across Greater Bandung, using an argumentative multimodal text writing performance test scored with a five-phase rubric (orientation to issue, multi-source exploration, Toulmin argumentation, multimodal production, and peer-review/reflection; maximum score of 20), a supplementary 17-item Likert-scale questionnaire on digital literacy and writing self-efficacy administered to a subset of students (n = 176; Cronbach's $\alpha = 0.849$), and semi-structured interviews. For the Multimodal Production phase, visual and digital elements were scored as functionally integrated when accompanied by explicit verbal justification linking the element's form (e.g., color, layout, image choice) to the text's claim or evidence, and as decorative when such elements appeared without an articulated argumentative function. In the Design stage, the identified needs were formulated into a model specification comprising ABCD-format learning objectives, a five-phase syntax, a social system and principles of reaction, an assessment framework, and a support system consisting of a package of teaching materials, whose internal coherence was verified through an alignment matrix.

In the Development stage, four products—the Model Guidebook, a Deep Learning-format Teaching Module, a five-phase Student Worksheet, and the Toulmin Assessment Rubric—were validated by two expert validators in instructional design and multimodal semiotics from universities. Quantitative validation used a scaled assessment sheet with feasibility categories based on the percentage of the score relative to the maximum score (adapted from Akbar, 2013). Inter-rater agreement between the two validators, computed using the intraclass correlation coefficient (ICC), varied by instrument (Model Guidebook ICC = 0.32; Teaching Module ICC = 0.61; Student Worksheet ICC = 0.66; Assessment Rubric ICC = 1.00). In the Implementation stage, Prototype II of the MOSAIK Model was trialed across two cycles: a limited trial in one Teknik Komputer dan Jaringan (Computer Networking) class and one Rekayasa Perangkat Lunak (Software Engineering) class (n = 67) serving as an initial baseline, and a broad trial across nine classes in the Teknik Jaringan Komputer dan Telekomunikasi (Computer Network and Telecommunication Engineering), Rekayasa Perangkat Lunak (Software Engineering), and Analisis Kimia (Chemical Analysis) programs at a different public vocational high school in Bandung City (n = 306). Syntax implementation was measured through structured observation and document analysis of worksheets and the multimodal text products generated by students at each phase.

Quantitative data—writing performance test scores and expert validation scores—were analyzed using descriptive statistics in the form of means and percentages. Expert validation scores were classified using feasibility categories adapted from Akbar (2013): Highly Feasible (≥85%), Feasible (70–84.9%), Fairly Feasible (55–69.9%), and Not Feasible (<55%). Writing performance and implementation scores were classified using the same percentage-based bands applied to sub-skill category: Excellent (≥85%), Good (70–84.9%), Fair (55–69.9%), and Poor (<55%). Qualitative data from interviews, validator input, and implementation documentation were analyzed through data reduction, thematic categorization, data display, and conclusion drawing, following the interactive analysis model (Miles & Huberman, 1994). Data trustworthiness was maintained through source and method triangulation: every finding at the Implementation stage was explicitly traced back to the findings of the Analyze stage, the design decisions of the Design stage, and the validator input from the Development stage, so that the chain of evidence from field needs to implementation outcomes could be coherently accounted for.

Results and Discussions

Learning Needs Profile for Multimodal Text Writing

The argumentative multimodal text writing performance test administered to 373 students showed a mean total score of 14.60 out of a maximum score of 20 (73.0%), placing performance in the Good category. Scores were relatively consistent across the five rubric phases (70.6%–75.4%), with the Toulmin Argumentation phase comparatively the weakest (70.6%). Table 1 presents the mean scores per assessment aspect by study program.

Table 1. Mean Scores of the Initial Argumentative Multimodal Text Writing Test

Assessment Aspect	RPL	TKJ	KA	Mean	Category
Issue Orientation (Phase 1)	2.5	2.8	3.3	3.0	Good
Multi-Source Exploration (Phase 2)	3.1	2.8	3.0	2.9	Good
Toulmin Argumentation (Phase 3)	3.0	3.0	2.7	2.8	Good
Multimodal Production (Phase 4)	3.6	2.6	2.9	2.9	Good
Peer-review & Reflection (Phase 5)	2.4	2.1	3.4	2.9	Good

Note: RPL = Rekayasa Perangkat Lunak (Software Engineering); TKJ = Teknik Komputer dan Jaringan (Computer Networking); AK = Analisis Kimia (Chemical Analysis).

Scores were relatively consistent across study programs, with Chemical Analysis students scoring comparatively higher on Issue Orientation and Peer-review & Reflection, Software Engineering students scoring higher on Multimodal Production, and Computer Networking students showing a flatter profile across phases. The relatively weakest phase overall was Toulmin Argumentation (70.6%), suggesting that constructing explicit logical links between claims and evidence is the sub-skill most in need of consistent scaffolding, even though no phase fell into a genuinely low category. The supplementary questionnaire reinforced this picture: students reported high levels of daily digital consumption for learning purposes, with generally acceptable self-reported literacy and writing self-efficacy ($\alpha = 0.849$), though responses on classroom learning experience indicated that instruction remained largely theory-oriented with limited structured practice. Teachers likewise reported that, while students were broadly capable writers, constructing logical reasons and evidence and assessing multimodal texts consistently across classes remained difficult in the absence of an operational, program-agnostic rubric. This combination—generally adequate but uneven baseline ability, and the absence of a shared instructional structure—became the primary empirical foundation for the design of the MOSAIK Model. This is consistent with the view that combining verbal and visual elements into a coherent argument benefits from explicit instructional scaffolding, rather than being acquired through mere exposure to digital technology (Howell et al., 2017).

Characteristics of the MOSAIK Model

Based on these identified needs, the MOSAIK Model was designed as a learning model for writing argumentative multimodal texts for vocational high school students in the Analisis Kimia (Chemical Analysis), Rekayasa Perangkat Lunak (Software Engineering), and Teknik Komputer dan Jaringan (Computer Networking programs, grounded in the Merdeka Curriculum, Phases E–F. The name MOSAIK—Indonesian for "mosaic"—reflects the model's guiding metaphor: a well-formed multimodal

argument, like a mosaic, is built by deliberately arranging distinct verbal, visual, and digital pieces into a single, coherent, evidence-based whole. The model draws on five complementary theoretical foundations: social semiotics (Kress & van Leeuwen, 2001), multiliteracies (New London Group, 1996), syntopical reading (Adler, 1972), Toulmin's argumentation model (1958), and the principles of Deep Learning, which emphasize mindful, meaningful, and joyful learning.

The syntax of the MOSAIK Model was designed in five progressive phases, moving from guided to independent practice: (1) Issue Orientation, in which students observe a contextual multimodal trigger text relevant to their study program and formulate an argumentative question; (2) Multi-Source Exploration, in which students access, evaluate using accuracy-authority-objectivity-relevance-coverage criteria, and synthesize at least two different digital sources; (3) Argument Design, in which students construct a Toulmin argument map comprising claim, evidence, and reasoning while also planning the semiotic modes to be used; (4) Multimodal Text Production, in which students produce a text that functionally integrates verbal, visual, and digital elements using a digital platform; and (5) Peer Review and Revision, in which students provide rubric-based feedback and revise their texts accompanied by metacognitive reflection. Each phase is equipped with an integrated framework of diagnostic, formative, and summative assessment together with peer and self-assessment, and is supported by four teaching-material components: the Model Guidebook, a Deep Learning-format Teaching Module, a five-phase Student Worksheet, and a six-aspect Toulmin Assessment Rubric. Internal alignment verification showed that all phases maintained full coherence among learning objectives, student activities, and assessment instruments, so that no gap existed between what was taught, how it was taught, and how it was assessed.

Validity of the MOSAIK Model Products

The four MOSAIK Model products were validated by two academic experts in instructional design and multimodal semiotics. Table 2 presents a summary of the quantitative validation results.

Table 2. Summary of Expert Validation Results for the MOSAIK Model Products

Product	Max Score	V1	V2	Mean (%)	Category
MOSAIK Model Guidebook	84	73	66	82.7%	Highly Feasible
Deep Learning Teaching Module	76	65	69	88.2%	Highly Feasible
Five-Phase Student Worksheet	68	59	59	86.8%	Highly Feasible
Toulmin Assessment Rubric*	32	28	28	87.5%	Highly Feasible

Note: Overall, the mean validation score across all products reached 86.5%, in the Highly Feasible category, based on the average of the two validators' ratings. Inter-rater agreement (ICC) ranged from 0.32 (Model Guidebook) to 1.00 (Assessment Rubric), with moderate agreement for the Teaching Module (ICC = 0.61) and Student Worksheet (ICC = 0.66), indicating that expert consensus was strongest for concrete, checklist-style instruments and weakest for the more conceptual rationale instrument.

Triangulation of input from the two validators produced consistent revision themes—accessibility of digital technical guidance, cross-program contextualization, strengthening of the empirical and theoretical basis, and continuity across material components—all of which were addressed before the products entered the implementation stage. This high and consistent level of feasibility is in line with the instructional model validity criteria, which require alignment between the theoretical foundation (content validity) and the interconnection among model components (construct validity) (Nieveen, 1999).

Implementation Feasibility and Quality of the MOSAIK Model

Implementation of Prototype II was carried out across two cycles. The limited trial was conducted in one Teknik Komputer dan Jaringan (Computer Networking) class and one Rekayasa Perangkat Lunak (Software Engineering) class (n = 67 students), serving as an initial implementation baseline prior to the broader trial. Students in both classes completed the five-phase syntax individually, producing argumentative multimodal texts scored with the same five-phase rubric used in the Analyze stage. The mean total score in the limited trial was 12.04 out of 20 (60.2%), lower than the corresponding baseline scores for these two classes, indicating that applying the model's full five-phase structure for the first time introduced its own procedural demands distinct from students' pre-existing writing ability.Findings from the limited trial were followed up through additive-corrective revisions to the teaching materials prior to the broad trial, including: changing the generic reflection format into explicit reflective statements that encourage the articulation of reasoning; adding a guiding column for justifying evidence-claim links on the argument map; changing the teacher approval mechanism from a simple checkmark to mandatory written feedback as a quality-control gate; and adjusting the multimodal integration assessment rubric to be relevant to the characteristics of each study program.The broad trial was then conducted across nine classes in the Teknik Jaringan Komputer dan

Telekomunikasi (TJKT), Rekayasa Perangkat Lunak (RPL), and Analisis Kimia (AK) programs at a different public vocational high school in Bandung City (n = 306). A one-way ANOVA showed a statistically significant difference in implementation scores between study programs, $F(2, 303) = 20.44$, $p < .001$, $\eta^2 = 0.12$; Tukey HSD post-hoc comparisons indicated RPL scored significantly higher than both AK and TKJ, and AK scored significantly higher than TKJ (all pairwise $p < .001$). Across all broad-trial classes, the Multimodal Production phase was consistently the weakest component (M = 2.28/4, 57.0%), regardless of study program, while Issue Orientation and Multi-Source Exploration were consistently the strongest (89.8% and 91.3%, respectively). Table 3 summarizes the implementation of the five-phase syntax by class.

Table 3. Matrix of Five-Phase Syntax Implementation by Class in the Broad Trial

Class	Implementation	Differentiating Factor
XI TKJ 2	Full — all five phases completed	Highest-performing TKJ class; adequate time allocation and active teacher guidance
XI TKJ 3	Partial — Phase 5 (Peer-review & Reflection) left incomplete in most groups	Several groups submitted highly similar or identical responses on shared phases, consistent with limited independent group work
XI RPL 2	Full — best overall performance in the broad trial	Strongest argumentation quality across the broad trial; one group's submitted digital product links could not be verified and were excluded from qualitative analysis
XI AK 1–6	Full — consistently completed across all six classes	Most uniform implementation across classes; scored significantly higher than TKJ but below RPL

This pattern shows that full implementation was most consistently achieved in classes with a high degree of alignment between the study program's competency profile and the model's multimodal demands, whereas classes that experienced time constraints during the transition from source exploration to argument design tended to show partial implementation. In terms of product quality, a number of groups achieved high performance along different dimensions: the use of nationally recognized academic sources and three-pronged reasoning among the Rekayasa Perangkat Lunak (RPL) groups, which produced the strongest argumentation quality across the entire broad trial; contextualization of vocational issues linking national issues to professional identity; peer review that used argumentation terminology appropriately, comparable to the ability of novice reviewers; and topic originality in the Analisis Kimia (AK) program, which successfully connected social concepts with quantitative analytical data. The achievements of these groups exceeded the initial expectations identified during the needs-analysis stage.

On the other hand, cross-class analysis identified six systemic obstacles that consistently appeared in more than one class: difficulty completing the argument-design phase within the available time allocation; the teacher-approval mechanism, which in most cases was not implemented substantively because it consisted only of an administrative checkmark without substantive feedback; a small number of groups copying example trigger questions instead of formulating their own; source-credibility evaluation that was procedural rather than substantively justified; reasoning (warrant) that was rarely articulated explicitly even when claims and evidence were already available; and inconsistent collection of digital product links. Triangulation between the limited and broad trials revealed a distinct pattern from the needs-analysis stage: although baseline ability across the five sub-skills was relatively uniform, with no pronounced multimodal weakness, multimodal production became the most persistent difficulty once students actively applied the model, indicating that this weakness is specific to the demands of model implementation rather than a pre-existing deficit. This pattern was most substantially resolved in classes with strong visual or digital competency profiles, whereas the weakness in evidence-based reasoning—already the relatively weakest sub-skill at baseline—remained a challenge in most classes, showing that reasoning competency requires more explicit scaffolding and a longer time allocation compared with multimodal-integration competency.

Discussion: The Extent to Which the MOSAIK Model Meets Instructional Model Quality Criteria

With reference to three instructional model quality criteria—validity, practicality, and potential positive effect (Nieveen, 1999)—the findings of this study show that the MOSAIK Model strongly meets the validity criterion, as evidenced by consistently Highly Feasible expert validation scores (86.5%) and full alignment among the theoretical foundation, learning objectives, activities, and assessment instruments in the alignment matrix. The practicality criterion was met conditionally: the model proved capable of full implementation in contexts with a high degree of vocational-profile fit, as shown by full implementation in

the limited trial and in most Rekayasa Perangkat Lunak (RPL) classes in the broad trial, but encountered obstacles in classes with limited time allocation and underdeveloped process quality-control mechanisms. This finding is consistent with the view that a model's practicality is not absolute but depends on actual implementation conditions in the field (Nieveen, 1999)—meaning that the implementation variation found does not indicate a design failure, but rather confirms the need for contextual adjustment, precisely as the model was adaptively designed to allow for from the Design stage onward.

This finding also reinforces the argument that weak support from teaching materials contributes to low independent thinking and poor argumentation quality among secondary school students (Syaifudin Nakrowi et al., 2024)—in this study, valid and systematic material support was shown to be capable of raising argumentation quality among groups with optimal implementation, while also showing that reasoning (warrant) competency is the aspect that is consistently most difficult to achieve even when scaffolding has been provided—a pattern consistent with the finding that improvements in multimodal literacy are not automatically followed by improvements in argumentative reasoning quality within a short period (Nur Fajriah et al., 2026). Meanwhile, the success in overcoming the visual-integration weakness in classes with strong visual or digital competency profiles reinforces the finding that a multiliteracies approach that is systematically integrated—rather than simply added as a supplement—can improve visual and digital literacy simultaneously (Fadli, 2025). The positive-response criterion cannot yet be fully concluded in this study, as summative response questionnaire data from students and teachers following implementation had not yet been analyzed at this reporting stage; early indications in the form of high engagement and product quality exceeding expectations in a number of groups provide a positive signal, but stronger confirmation requires analysis of summative evaluation data in the next stage of the research. This limitation is in line with the research agenda asserting that process-based assessment methods for digital multimodal composing remain an underdeveloped area compared with the rapid pace of its pedagogical implementation (Hafner, 2024)—a gap that is also evident in this study and that needs to be filled through the development of summative evaluation instruments in the next stage of the research.

Conclusions

This study produced the MOSAIK Model, a digital literacy-based learning model for writing argumentative multimodal texts for vocational high schools, systematically developed through the stages of needs analysis, design, development, and implementation within the ADDIE framework. The analysis stage found that students' baseline ability to write argumentative multimodal texts was already at a Good level overall (73.0%), though relatively uneven across sub-skills, with evidence-based reasoning (Toulmin argumentation) comparatively the weakest component. This indicated that the primary need was not remediation of a uniform deficit, but a structured, replicable model capable of ensuring consistent writing quality across heterogeneous vocational study programs. The MOSAIK Model, designed with a five-phase syntax and validated by two experts, achieved an overall Highly Feasible category (mean 86.5%), though inter-rater consensus varied by instrument (ICC 0.32–1.00), being weakest for the conceptual rationale instrument. Implementation across two trial cycles showed that the model is adaptive to the context of different study programs: full implementation and the highest argumentation quality were achieved in classes with a high degree of vocational-profile fit; the multimodal-production phase, however, emerged as the most persistent challenge during active implementation—a weakness not evident at baseline—while evidence-based reasoning remained a secondary challenge requiring more explicit scaffolding and a longer time allocation. The six systemic obstacles identified—particularly in the transition from the exploration phase to argument design and in the underdeveloped teacher quality-control mechanism—constitute a concrete agenda for model refinement.

Beyond answering the research questions, this study offers four concrete innovations that distinguish the MOSAIK Model from prior work. First, it is one of the few learning models to systematically fuse Toulmin's claim-evidence-warrant argumentation structure with functional multimodal semiotic production into a single, expert-validated, five-phase syntax purpose-built and classroom-tested for vocational rather than general secondary education. Second, the study identifies and operationalizes a "teacher quality-control gate"—replacing a merely administrative approval checkmark with mandatory written feedback—as a concrete, replicable mechanism for sustaining argumentation quality during peer-supported text production; this mechanism emerged empirically through the additive-corrective revision between the limited and broad trials, rather than being assumed a priori. Third, the findings establish an empirically grounded diagnostic distinction between the two literacy weaknesses targeted by the model: multimodal-integration weaknesses

proved responsive to explicit instructional scaffolding within a single implementation cycle, whereas evidence-based reasoning weaknesses persisted across contexts and require longer, more explicit scaffolding—a distinction with direct implications for how future multimodal argumentative writing models should allocate instructional time. Fourth, the model demonstrates a context-adaptive design strategy in which a single core syntax produced consistently structured multimodal argumentative texts across vocational study programs while still yielding measurable differences in argumentation quality between programs (RPL > AK > TKJ), offering a transferable design principle for developing instructional models intended for heterogeneous vocational contexts rather than a single uniform classroom setting.

These results demonstrate that the MOSAIK Model is worth considering as an alternative learning model for writing multimodal texts in vocational high schools, with the caveat that refinements to the reasoning-scaffolding mechanism and the process quality-control gate are needed before wider dissemination. Further research is needed to analyze summative evaluation data in the form of teacher and student responses following implementation, in order to fully address all three instructional model quality criteria.

References

Adler, M. J. (1972). How to read a book (Vol. 1). Simon & Schuster.

Akbar, S. (2013). Instrumen perangkat pembelajaran [Learning-tool instruments]. PT Remaja Rosdakarya.

Al Fajri, T. A. (2018). Pentingnya penggunaan pendekatan multimodal dalam pembelajaran [The importance of using a multimodal approach in learning]. Waskita: Jurnal Pendidikan Nilai dan Pembangunan Karakter, 2(1). https://doi.org/10.21776/ub.waskita.2018.002.01.5

Badan Standar, Kurikulum, dan Asesmen Pendidikan. (2025). Mata pelajaran Bahasa Indonesia fase A–F dan fase F tingkat lanjut [Indonesian language subject, phases A–F and advanced phase F]. Kementerian Pendidikan Dasar dan Menengah.

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. Written Communication, 25(2), 166–195. https://doi.org/10.1177/0741088307313177

Branch, R. M. (2010). Instructional design: The ADDIE approach. Springer US. https://doi.org/10.1007/978-0-387-09506-6

Dahlström, H. (2022). Students as digital multimodal text designers: A study of resources, affordances, and experiences. British Journal of Educational Technology, 53(2), 391–407. https://doi.org/10.1111/bjet.13171

Fadli, B. M. (2025). Implementasi multiliterasi melalui pendekatan ekologi sastra dalam pembelajaran Bahasa Indonesia di era digital [Implementing multiliteracies through a literary ecology approach in Indonesian language learning in the digital era]. Jurnal Pendidikan Bahasa dan Sastra, 4. http://ejournal.mandalanursa.org/index.php/JUPBAS/index

Hafner, C. A. (2024). Digital multimodal composing in L2 classrooms: A research agenda. Language Teaching. Advance online publication. https://doi.org/10.1017/S0261444824000107

Hafner, C. A., & Ho, W. Y. J. (2020). Assessing digital multimodal composing in second language writing: Towards a process-based model. Journal of Second Language Writing, 47, 100710. https://doi.org/10.1016/j.jslw.2020.100710

Howell, E., Butler, T., & Reinking, D. (2017). Integrating multimodal arguments into high school writing instruction. Journal of Literacy Research, 49(2), 181–209. https://doi.org/10.1177/1086296X17700456

Hutasuhut, M. L., Damanik, S. F., Tambunan, A. R. S., Andayani, W., & Setiasari, W. (2025). Digital scrapbook as a multimodal tool for teaching and learning writing argumentative text in Indonesian EFL classrooms. Advances in Social Science, Education and Humanities Research, 44–51. https://doi.org/10.2991/978-2-38476-501-0_6

Jewitt, C. (2008). Multimodality and literacy in school classrooms. Review of Research in Education, 32(1), 241–267. https://doi.org/10.3102/0091732X07310586

Jiang, L. (2017). The affordances of digital multimodal composing for EFL learning. ELT Journal, 71(4), 413–422. https://doi.org/10.1093/elt/ccw098

Jiang, L., Yang, M., & Yu, S. (2020). Chinese ethnic minority students' investment in English learning empowered by digital multimodal composing. TESOL Quarterly, 54(4), 954–979. https://doi.org/10.1002/tesq.566

Kemendikbudristek. (2022). Laporan hasil Asesmen Nasional 2022 [Report on the results of the 2022 National Assessment]. Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.

Kendrick, M., Early, M., Michalovich, A., & Mangat, M. (2022). Digital storytelling with youth from refugee backgrounds: Possibilities for language and digital literacy learning. TESOL Quarterly, 56(3), 961–984. https://doi.org/10.1002/tesq.3146

Kress, G., & van Leeuwen, T. (2001). Multimodal discourse: The modes and media of contemporary communication. Arnold.

Liu, H., Liu, L., & Li, H. (2024). Multimodal discourse studies in the international academic community (1997–2023): A bibliometric analysis. SAGE Open, 14(4). https://doi.org/10.1177/21582440241305454

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook (2nd ed.). SAGE Publications.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. Harvard Educational Review, 66(1), 60–92.

Nieveen, N. (1999). Prototyping to reach product quality. In J. van den Akker, R. M. Branch, K. Gustafson, N. Nieveen, & T. Plomp (Eds.), Design approaches and tools in education and training (pp. 125–135). Kluwer Academic Publishers.

Nur Fajriah, Y., Dewi, S., Nurjamin, L. R., Nurjamin, A., & Nastya, A. (2026). Transforming narrative text instruction in senior high school: Integrating digital storytelling and the SFL-GBA approach. Qalamuna: Jurnal Pendidikan, Sosial, dan Agama, 18(1), 115–126. https://doi.org/10.37680/qalamuna.v18i1.8251

OECD. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing.

Pratiwi, N. K., Yulianto, B., Mintowati, M., & Sodiq, S. (2024). Optimizing ChatGPT for scientific writing in the digital age. LITERA, 23(2), 224–233. https://doi.org/10.21831/ltr.v23i2.72979

Rahma, F. A., Harjono, H. S., & Sulistyo, U. (2023). Problematika pemanfaatan media pembelajaran berbasis digital [Problems in the use of digital-based learning media]. Jurnal Basicedu, 7(1), 603–611. https://doi.org/10.31004/basicedu.v7i1.4653

Sahidah, N., & Kirana, T. (2021). [Article title]. Jurnal Education and Development, 9(1), 370.

Silvhiany, S., Nurhaliza, S., Inderawati, R., & Trilestari, K. (2026). Climate change in EFL: Inquiry-based multimodal writing. Studies in English Language and Education, 13(1), 171–188. https://doi.org/10.24815/siele.v13i1.152

Syaifudin Nakrowi, Z., Ansori, D. S., Mulyati, Y., & Setyaningsih, Y. (2024). Kemampuan menulis argumentasi siswa: Analisis profil dan problematika pembelajaran [Students' argumentative writing ability: A profile analysis and instructional problems]. Bahasa dan Sastra, 10(1). https://e-journal.my.id/onoma

Toulmin, S. E. (1958). The uses of argument. Cambridge University Press.

Yuniarti, L., Yanti, Z. P., Gusriani, A., Noprina, W., Handayani, D. F., Srinengsi, S., & Almeyda, A. (2026). Pendampingan guru dalam penyusunan e-modul keterampilan menulis berbasis deep learning [Mentoring teachers in developing deep learning-based e-modules for writing skills]. JGEN: Jurnal Pengabdian kepada Masyarakat, 4. https://doi.org/10.60126/jgen.v4i1.1498