Kognisi: Jurnal Ilmu Keguruan
ISSN (online): 2987-0240
Volume 4 Issue 1 Year 2026
https://afeksi.id/journal3/index.php/kognisi/index

Academic Paper-Writing Performance of First-Year Early Childhood Islamic Education Students

Wina Santyani, Khamim Zarkasih Putro, Siti Zubaedah
Universitas Islam Negeri Sunan Kalijaga, Yogyakarta, Indonesia
Corresponding author: winasantyani26@gmail.com
DOI: https://doi.org/10.59698/kognisi.v4i1.374

Abstract
Academic writing is a foundational practice through which university students learn to organize knowledge, construct arguments, use evidence, and participate in disciplinary communities. This study examined the quality of academic paper writing produced by first-semester students in the Early Childhood Islamic Education (PIAUD) program at Universitas Islam Negeri Sunan Kalijaga, Indonesia. A descriptive document-analysis design was used, combining quantitative scoring with qualitative interpretation. The unit of analysis comprised 14 group papers submitted at the end of an Indonesian Language course in the 2024/2025 academic year: seven papers from Class A and seven from Class C. Each paper was evaluated on five dimensions reported in the source dataset: language use, thematic relevance, organization, presentation/neatness, and references. Descriptive statistics were recalculated from the paper-level scores. The overall mean was 81.07 (SD = 2.30), with Class A averaging 80.57 and Class C 81.57. Ten of the 14 papers (71.43%) were classified as good and four (28.57%) as fair; all papers met the study's minimum passing criterion of 75. Presentation/neatness was the strongest dimension (M = 85.00), whereas organization (M = 78.21) and language use (M = 78.57) were the weakest. Qualitative observations in the original dataset further indicated recurring difficulties in academic conventions, including punctuation and spelling, systematic organization, formatting consistency, citation practice, and limited use of scholarly sources. The findings suggest that first-year students can produce acceptable academic papers, but their performance remains transitional rather than fully proficient. Explicit genre instruction, transparent rubrics, scaffolded source-use practice, reference-management training, and iterative feedback are recommended. Because the analyzed texts were group products and the original scoring instrument lacked reported inter-rater reliability, the findings should be interpreted as a profile of paper quality rather than a direct measure of individual writing competence.
Keywords: *Academic Writing, First-Year Students, Scientific Paper, Academic Literacy, Source Use, Higher Education*

INTRODUCTION

Academic writing is not simply a technical requirement of higher education; it is one of the principal means through which students learn to participate in academic communities. University assignments require students to transform reading into argument, organize evidence, make disciplinary meanings visible, and position their own claims in relation to existing scholarship. The academic literacies tradition has therefore challenged the assumption that university writing is a generic skill that students should already possess when they enter higher education. Instead, writing is understood as a socially situated practice shaped by disciplinary conventions, institutional expectations, identity, and relations of authority (Lea & Street, 1998; Lea, 2004). This perspective is particularly relevant to first-year

students, who are simultaneously adjusting to new content knowledge and learning unfamiliar ways of constructing and presenting knowledge.

The transition from secondary school to university frequently exposes a gap between students' previous writing experiences and the demands of academic discourse. First-year students may understand that a paper should be "formal" yet remain uncertain about what counts as a defensible argument, how sections should be related, how evidence should be selected, or how source material should be integrated. Studies of first-year academic literacy indicate that students benefit when expectations are made explicit and writing is embedded in disciplinary learning rather than treated as a remedial service located outside the curriculum (Khumalo & Reddy, 2021; Roald et al., 2021; Wingate, 2012). Howitt et al. (2024) likewise show that learning academic writing involves a process of becoming: students must reinterpret prior habits while developing new rhetorical and epistemic practices.

These challenges are also visible in Indonesian higher education. Research involving Indonesian university students has documented problems with organization, grammar, vocabulary, academic conventions, and confidence in scholarly writing (Azizah & Budiman, 2017; Nenotek et al., 2022). Studies focusing on Indonesian students' paper-writing performance similarly report persistent weaknesses in systematic organization, technical formatting, source selection, and referencing (Septafi, 2021; Syazali & Erfan, 2022; Widodo et al., 2020). These problems are pedagogically important because writing a course paper is often one of the earliest sustained academic genres encountered by undergraduates. Although a course paper is less complex than a thesis or journal article, it requires students to coordinate multiple practices at once: understanding a topic, locating literature, synthesizing sources, producing coherent prose, following an expected structure, and documenting references accurately.

Writing from sources is especially demanding for novice writers. Effective source use requires more than inserting citations. Writers must read selectively, understand the relationship between a source and their own argument, paraphrase or summarize without distorting meaning, and signal how cited material functions rhetorically. Research on source-based writing shows that novice writers frequently rely on surface-level borrowing or sentence-by-sentence reproduction because they have not yet developed the conceptual and linguistic resources needed for synthesis (Cumming et al., 2016; Keck, 2006; Pecorari, 2003; Shi, 2004). Even when citations are technically present, their rhetorical use may remain limited to attribution rather than evaluation, comparison, or synthesis (Lee et al., 2018; Mansourizadeh & Ahmad, 2011; Petrić, 2007). This distinction matters because a bibliography containing several sources does not necessarily demonstrate academic literacy if the sources are not meaningfully integrated into the paper.

The quality of paper writing is also influenced by students' understanding of the genre itself. Genre-based pedagogy argues that academic texts are patterned responses to recurring communicative purposes and that these patterns can be taught explicitly (Hyland, 2003). When novice students receive models, guided analysis, and opportunities to practice the rhetorical moves expected in a genre, they are better positioned to understand why a paper

contains particular sections and how those sections work together. Wingate (2012) proposes combining genre-based instruction with academic literacies principles so that students learn textual conventions without reducing writing to rigid templates. This balance is important: a template may support beginners by making expectations visible, but students also need to understand the reasoning behind the structure if they are to transfer learning to new assignments.

Assessment practices can either clarify or obscure those expectations. Analytic rubrics can improve transparency by identifying dimensions such as language, organization, argument, evidence, and referencing. Reviews of rubric use in higher education show that rubrics can support consistency and learning when criteria are sufficiently explicit and aligned with the construct being assessed (Jönsson & Svingby, 2007; Reddy & Andrade, 2010). However, scores alone are not enough. Feedback must help students understand the gap between current and desired performance and create opportunities for revision (Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006). In writing instruction, formative feedback and peer review can support revision, provided students are helped to interpret and use feedback productively (Carless & Boud, 2018; Cho & MacArthur, 2010; Huisman et al., 2018; Wingate, 2010).

The source manuscript for the present study analyzed academic papers written by first-semester students in the Early Childhood Islamic Education (Pendidikan Islam Anak Usia Dini, PIAUD) program at UIN Sunan Kalijaga. It reported results for two classes using five assessment criteria: language use, thematic relevance, organization, presentation/neatness, and references. The original draft offered a useful descriptive picture but contained three limitations that reduced its publication readiness. First, it alternated between describing the design as qualitative and quantitative, even though numerical scoring was central to the analysis. Second, several percentages in Class A were arithmetically inconsistent with the reported frequencies. Third, the discussion relied heavily on general statements and a small reference base, limiting engagement with international research on academic literacies, genre, source use, assessment, and feedback.

This revised study addresses those issues without altering the underlying dataset. The numerical results were recalculated directly from the 14 paper-level score profiles, while the qualitative observations reported in the source manuscript were retained as contextual evidence. The study is framed as a descriptive document analysis rather than an assessment of individual students because each analyzed text was a group paper. This distinction avoids attributing a shared product to an individual writer and creates a more defensible interpretation of the evidence.

Accordingly, the study addresses three research questions: (1) What is the overall quality profile of the 14 academic papers produced by first-semester PIAUD students? (2) Which assessed dimensions constitute relative strengths and weaknesses? and (3) What pedagogical implications follow from the observed profile when interpreted in relation to contemporary research on academic writing and academic literacy? By answering these questions, the study contributes a context-specific baseline for developing first-year writing support in an

Indonesian Islamic higher-education setting while also illustrating how simple course-assignment data can be used more cautiously for instructional evaluation.

METHODS

Research design

This study employed a descriptive document-analysis design with quantitative scoring and qualitative interpretation. The primary evidence consisted of academic papers that had already been produced as course assignments. The numerical scores were used to summarize the quality profile of the papers, while qualitative observations reported in the source manuscript were used to interpret recurring writing problems. This formulation is more consistent with the available evidence than describing the study as purely qualitative because the principal findings are based on criterion scores, frequencies, percentages, and means.

Context and unit of analysis

The study was conducted in the Early Childhood Islamic Education (PIAUD) program at Universitas Islam Negeri Sunan Kalijaga, Yogyakarta, during the odd semester of the 2024/2025 academic year. The source manuscript identifies first-semester students as the study population and specifies two participating classes, PIAUD A and PIAUD C. The analyzed corpus comprised 14 academic group papers submitted at the end of the Indonesian Language course: seven papers from Class A and seven papers from Class C. Because the papers were written collaboratively, the unit of analysis in this revised article is the paper rather than the individual student. Consequently, the results describe the quality of submitted group papers and should not be interpreted as direct estimates of each student's independent writing ability.

Assessment dimensions and scoring

The source dataset evaluated each paper on five dimensions: (1) language use, including grammar and academic language conventions; (2) relevance to the assigned theme; (3) systematic organization; (4) presentation/neatness; and (5) references. Each criterion was recorded on a 0-100 scale. The total score for a paper was calculated as the arithmetic mean of the five criterion scores, which reproduces the totals shown in the source tables. Overall scores were then interpreted using the categories specified in the original instrument: 90-100 = excellent, 80-89 = good, 70-79 = fair, 50-69 = poor, and 0-49 = very poor. The original study also defined a score of 75 as the minimum passing threshold.

The scoring framework has a practical advantage because it makes visible several dimensions that novice writers need to coordinate. At the same time, the source manuscript does not report analytic descriptors for different score points within each criterion, formal content-validation procedures, rater training, the number of raters, or inter-rater reliability. These omissions limit the strength of measurement claims. Contemporary assessment research emphasizes that rubrics require clear construct definitions and evidence of scoring consistency if numerical differences are to be interpreted precisely (Jönsson & Svingby, 2007; Reddy & Andrade, 2010). For this reason, the present article treats the scores as descriptive classroom evidence rather than as psychometrically validated measures.

Data analysis

The paper-level scores were recalculated to verify all descriptive results. For each class, the analysis generated the mean, standard deviation, minimum, and maximum overall score. Frequencies and percentages were calculated for the performance categories. The two classes were also combined to produce corpus-level descriptive statistics. Mean scores were computed for each of the five assessment dimensions to identify relative strengths and weaknesses across the 14 papers. No inferential test was conducted to compare classes. With only seven group papers per class, and without evidence that the papers represented independent randomly sampled observations, a significance test would add an appearance of precision not supported by the design.

The qualitative component drew on recurring issues explicitly described in the original findings and discussion: spelling and punctuation problems; use of non-standard or informal wording; inconsistent italicization of foreign terms; incomplete coverage of the assigned topic; uncertainty about paper structure; inconsistent margins, spacing, and fonts; weak citation conventions; and a small number of references in several papers. These observations were organized around the five assessment dimensions and interpreted using research on academic literacies, genre pedagogy, source use, writing self-regulation, rubrics, and feedback.

Ethical and interpretive considerations

The source manuscript does not provide details regarding formal ethical approval, informed consent, or anonymization procedures. The present revision therefore does not add claims about ethical clearance that are absent from the evidence. It reports only aggregated paper-level data and does not identify individual students. In addition, the revised article avoids causal language. The analysis describes performance at one point in one course and cannot determine whether the observed quality was caused by prior schooling, the Indonesian Language course, lecturer guidance, collaboration within groups, or other factors.

RESULTS

Overall performance profile

Across the 14 group papers, the overall mean score was 81.07 (SD = 2.30), with scores ranging from 76 to 84. Class A obtained a mean of 80.57 (SD = 2.94; range = 76-84), while Class C obtained a mean of 81.57 (SD = 1.51; range = 79-84). The one-point difference is descriptively small and is not interpreted as evidence that one class had superior writing competence, particularly because the analysis concerns a limited set of group products rather than independent individual performances.

Table 1. Overall paper-writing performance by class

Class	n	Mean	SD	Range	Category distribution
PIAUD A	7	80.57	2.94	76–84	Good: 4 (57.14%); Fair: 3 (42.86%)
PIAUD C	7	81.57	1.51	79–84	Good: 6 (85.71%); Fair: 1 (14.29%)

Combined	14	81.07	2.30	76–84	Good: 10 (71.43%); Fair: 4 (28.57%)

The categorical distribution also indicates generally acceptable performance. In Class A, four of seven papers (57.14%) were in the good category and three (42.86%) were in the fair category. In Class C, six of seven papers (85.71%) were good and one (14.29%) was fair. Across both classes, 10 of 14 papers (71.43%) were good and four (28.57%) were fair. No paper fell in the excellent, poor, or very poor categories. All 14 papers scored at least 75, meaning that every paper met the minimum passing threshold established in the original study.

Table 2. Mean scores by assessment dimension

Dimension	Class A	Class C	Combined mean	Combined SD
Language use	77.86	79.29	78.57	3.63
Thematic relevance	81.43	83.57	82.50	3.80
Organization	77.86	78.57	78.21	2.49
Presentation/neatness	85.71	84.29	85.00	3.40
References	80.00	82.14	81.07	3.50

Performance by assessment dimension

The dimension-level profile reveals more variation than the overall scores. Presentation/neatness was the strongest area, with a combined mean of 85.00 (SD = 3.40). Relevance to the assigned theme followed with a mean of 82.50 (SD = 3.80), and references obtained a mean of 81.07 (SD = 3.50). Language use averaged 78.57 (SD = 3.63), while organization recorded the lowest mean at 78.21 (SD = 2.49). Thus, the papers were generally more successful in surface presentation and staying on topic than in the linguistic and structural dimensions that support precise academic communication.

Class-level patterns were similar. In Class A, presentation/neatness had the highest mean (85.71), followed by thematic relevance (81.43) and references (80.00). Language use and organization both averaged 77.86. In Class C, presentation/neatness again scored highly (84.29), while thematic relevance reached 83.57 and references 82.14. Organization remained relatively low (78.57), and language use averaged 79.29. These data suggest that the principal instructional needs were shared across the two classes rather than confined to one class.

Qualitative profile of recurring problems

The source manuscript's qualitative observations help explain why the overall category "good" should not be equated with mature academic proficiency. In language use, most papers were understandable, but recurring errors involved spelling, punctuation, capitalization, non-standard word choices, and formatting of foreign terms. Such errors may appear minor individually, yet their accumulation can reduce readability and weaken the academic register of a paper.

In relation to content, the papers generally addressed the assigned themes, which is consistent with the relatively high relevance score. Nevertheless, the source manuscript notes that some discussions were too brief, leaving parts of the topic underdeveloped. The issue

was therefore not primarily topic deviation but depth and completeness. This distinction is important because a paper can remain relevant to a topic while still offering insufficient analysis or evidence.

Organization emerged as the weakest quantitative dimension and was also a major qualitative concern. Students were reported to have limited familiarity with a formal paper-writing guide and sometimes relied on examples obtained from senior students or internet searches. Such imitation may help novice writers begin a task, but when models vary in quality, it can produce inconsistent sectioning and weak relationships among ideas. Formatting problems were also reported, including uncertainty about margins, line spacing, and font conventions, even though neatness received the highest average score. This apparent tension suggests that basic visual presentation was relatively successful but not fully standardized.

Finally, referencing remained a substantive weakness despite a mean score above 80. The source manuscript reports difficulties with in-text citation, footnotes or author-date citation conventions, reference-list formatting, and limited source quantity; some groups reportedly used fewer than five references. The issue therefore concerns both technical accuracy and the breadth and integration of evidence. A paper may receive a moderate-to-good reference score for including a reference list while still requiring substantial development in source evaluation, synthesis, paraphrasing, and rhetorical citation.

DISCUSSION

Interpreting "good" performance as transitional academic literacy

The central finding is that the 14 group papers were generally acceptable according to the classroom scoring system, but the pattern is better described as transitional academic literacy than as fully developed academic writing competence. Ten papers were categorized as good and the remaining four as fair, yet none reached the excellent category. More importantly, the dimension profile shows that presentation and topic relevance were stronger than organization and language use. This pattern suggests that students could produce recognizable academic papers and comply with visible assignment expectations, while the deeper rhetorical practices of organizing argument, sustaining academic register, and integrating evidence remained less secure.

This interpretation is consistent with the academic literacies perspective. Lea and Street (1998) argue that difficulties in student writing should not automatically be understood as deficits located within students. Newcomers must learn how knowledge is constructed and communicated in particular institutional settings, often while expectations remain tacit. Lea (2004) further emphasizes that academic literacy should be designed into courses rather than separated from disciplinary learning. For first-semester PIAUD students, the challenge is therefore not merely to "write correctly" but to learn how an academic paper functions within university study: what counts as evidence, how a topic is narrowed, how sections contribute to a line of reasoning, and how the writer's voice relates to the voices of sources.

The first-year transition literature supports this view. Khumalo and Reddy (2021) found that integrated writing support helped students connect language practices with disciplinary

expectations. Roald et al. (2021) similarly show that contrasting genres can make academic conventions visible to first-year students. Howitt et al. (2024) describe academic writing development as a process in which undergraduates must unlearn assumptions and reconstruct their identity as academic writers. These studies help explain why even students who receive satisfactory scores can continue to experience uncertainty. A score in the low 80s may indicate successful task completion while masking unresolved questions about purpose, audience, argument, and evidence.

Organization as a priority for explicit genre instruction

Organization produced the lowest combined mean (78.21), making it the clearest quantitative priority. The source manuscript also reports that students lacked a clear paper-writing guide and sometimes learned through imitation of papers found online or obtained from senior students. This is precisely the kind of context in which explicit genre instruction can be productive. Hyland (2003) argues that genre pedagogy helps learners understand the relationship between communicative purpose and textual structure. The goal is not to impose a single inflexible template but to reveal recurring patterns that experienced academic writers often take for granted.

Wingate (2012) offers a useful synthesis by integrating genre pedagogy with academic literacies. Students need explicit guidance about text organization, but they also need opportunities to discuss why conventions exist and how they vary across tasks and disciplines. For an introductory PIAUD course, this could involve guided analysis of two or three model papers rather than simply distributing a template. Students might identify the purpose of an introduction, distinguish background from problem formulation, trace how headings build an argument, and compare strong and weak transitions between paragraphs. A paper template can then function as a scaffold rather than as a substitute for rhetorical understanding.

The Indonesian studies reinforce the relevance of such support. Widodo et al. (2020) found that new university students experienced difficulty in writing academic papers, while Syazali and Erfan (2022) specifically examined the use of a paper template as a guide for PGSD students. Septafi (2021) also documented challenges in university students' scientific article writing. The current findings extend that pattern to first-semester PIAUD group papers. The recurrence across programs suggests that first-year academic writing should be treated as a curricular issue rather than as an isolated problem of individual students.

Language accuracy, academic register, and writing development

Language use was the second-lowest dimension (M = 78.57). The qualitative notes identify spelling, punctuation, non-standard vocabulary, foreign-word formatting, and other editing issues. These problems are common among novice academic writers and should be addressed without reducing academic writing instruction to grammar correction. Nenotek et al. (2022) found that Indonesian university students' academic essay difficulties included language and organization, while Azizah and Budiman (2017) documented linguistic and rhetorical challenges experienced by Indonesian graduate writers preparing scholarly papers.

Although the populations and task demands differ, both studies show that control of academic language remains connected to broader writing competence.

A process-oriented response is preferable to simply penalizing errors in final submissions. Students can be taught to separate composing from editing: first developing content and structure, then revising paragraph coherence, and finally proofreading grammar, punctuation, formatting, and citation details. Feedback research provides a strong rationale for this staged approach. Hattie and Timperley (2007) emphasize that effective feedback clarifies where the learner is going, how current performance compares with the goal, and what action should follow. Nicol and Macfarlane-Dick (2006) similarly connect formative assessment with self-regulation. In academic writing, this means that comments such as "grammar needs improvement" are less useful than targeted guidance that helps students identify recurring patterns and revise them.

Wingate (2010) showed the importance of formative feedback for academic writing development, while Carless and Boud (2018) argue that students need feedback literacy: the ability to appreciate, interpret, and act on feedback. A practical implication for PIAUD is to require at least one structured revision cycle before a group paper receives its final score. Instead of treating the paper as a one-off product, the course can make visible the stages of planning, drafting, reviewing, revising, and proofreading. This is especially important for first-semester students, who may otherwise interpret a satisfactory final score as evidence that no further development is needed.

From citation mechanics to meaningful source use

The reference dimension averaged 81.07, but the qualitative evidence reveals several weaknesses: limited numbers of references, inconsistent citation practice, and uncertainty about the relationship between in-text citations and the reference list. These issues should not be treated only as formatting mistakes. Research on writing from sources demonstrates that citation is both a technical and rhetorical practice. Cumming et al. (2016) synthesize evidence showing that source-based writing requires coordinated reading, selection, transformation, and integration. Davis (2013), Doolan (2021), Thompson et al. (2013), and Wette (2017) similarly demonstrate that students develop source use gradually and continue to face challenges even after initial university experience.

For novices, paraphrasing is a particularly vulnerable point. Keck (2006) showed that L2 writers often produce paraphrases that remain close to source wording, while Shi (2004) and Pecorari (2003) illustrate how textual borrowing and patchwriting can occur even when students do not intend to violate academic norms. Although the present study did not conduct plagiarism analysis, the original manuscript's observation that students sometimes rely heavily on available examples and use few sources suggests the value of explicit source-transformation instruction. Students need guided practice in reading a short passage, identifying the core proposition, closing the source, restating the idea, checking semantic accuracy, and then linking the paraphrase to their own claim.

Citation quality also depends on rhetorical purpose. Petrić (2007) found that stronger academic writing uses citations for functions beyond simple attribution, including evaluation,

comparison, exemplification, and linking sources. Mansourizadeh and Ahmad (2011) showed differences between novice and expert writers in citation practices, and Lee et al. (2018) documented how first-year L2 writers vary in citation form, function, and stance. Ramoroka (2014) likewise emphasized source integration as a central feature of academic writing. These findings suggest that introductory writing instruction should move beyond "where to put the citation" toward questions such as: Why is this source cited here? Does it support, contrast with, define, or extend the writer's point? What does the writer want the reader to infer from the citation?

Reference-management software can support technical consistency, but it should be introduced after students understand the logic of citation. The current cohort's problems with references indicate a need for a short, practical sequence: searching scholarly databases; distinguishing scholarly from non-scholarly sources; recording bibliographic metadata; using an author-date system; creating a reference list; and checking one-to-one correspondence between citations and references. Syaputra et al. (2023) emphasize data literacy in scientific writing, which is closely related to students' ability to locate, evaluate, and responsibly use evidence. In first-year courses, these practices can be taught with a small number of carefully selected sources before students are expected to manage a larger literature base independently.

The role of rubrics and assessment transparency

The study's five criteria provide a useful starting structure, but publication-level interpretation requires acknowledging weaknesses in the instrument. The source manuscript does not report criterion descriptors, weighting rationale, scorer training, or inter-rater reliability. As a result, a score of 80 in "references" or 75 in "organization" cannot be interpreted with the same precision as a score produced by a validated analytic rubric. Jönsson and Svingby (2007) note that rubric reliability and validity depend on the clarity of criteria and scoring design, while Reddy and Andrade (2010) show that rubrics can support learning when students understand how criteria relate to quality.

For future use, each criterion should be expanded into observable descriptors across performance levels. For example, organization could distinguish between: an unclear or fragmentary structure; recognizable sections with weak internal coherence; logically sequenced sections with mostly coherent paragraphs; and an integrated structure in which sections, paragraphs, and transitions collectively advance a sustained argument. The reference criterion could separately assess source quality, citation accuracy, reference-list accuracy, and integration of sources. Such distinctions would make scores more diagnostically useful and help students understand what improvement looks like.

Rubrics are most effective when combined with exemplars and feedback. Students can apply the rubric to a sample paper before submitting their own work, discuss differences in judgments, and then use the same criteria for self-assessment. Peer feedback can further develop evaluative judgment. Cho and MacArthur (2010) found benefits of peer and expert review for revision, while Huisman et al. (2018) showed associations between peer feedback processes and academic writing performance. Carless and Boud's (2018) concept of feedback

literacy is especially relevant: students should not be passive recipients of lecturer comments but active interpreters who learn to recognize quality.

Writing self-efficacy and self-regulation as supporting conditions

The present study did not measure writing self-efficacy, so no claims can be made about students' confidence. Nevertheless, the wider literature indicates why first-year writing support should include opportunities for manageable success and self-regulation. Prat-Sala and Redford (2012) found that writing self-efficacy is related to undergraduate essay performance, and Sanders-Reio et al. (2014) showed that beliefs about writing, self-efficacy, apprehension, and performance are interconnected. Sun and Wang (2020) linked writing self-efficacy with self-regulated learning strategies among university EFL students, while Sun et al. (2021) reported a positive relationship between L2 writing self-efficacy and achievement in a meta-regression analysis. Teng et al. (2018) further conceptualized writing self-efficacy as multidimensional rather than a single generalized belief, and Mitchell et al. (2023) show in their scoping review that postsecondary writing self-efficacy has been studied across diverse writing contexts.

These findings do not mean that confidence should replace rigorous instruction. Rather, transparent criteria, staged assignments, practice with manageable source sets, and revision opportunities may help students develop a more accurate sense of control over the writing process. First-year students who are told only to "write a paper" may experience the task as an undifferentiated demand. Breaking the task into topic formulation, source search, outline, draft, source integration, revision, and editing turns academic writing into a sequence of learnable practices.

Interpreting class differences cautiously

Class C had a mean overall score one point higher than Class A and a larger proportion of papers in the good category. The difference should not be overinterpreted. Each class contributed only seven group papers, the assignment process may have differed in ways not recorded in the source manuscript, and group products are influenced by composition, division of labor, and peer dynamics. Moreover, no evidence is available about baseline writing ability or equivalence of instructional conditions. The class comparison is therefore descriptive only.

This cautious interpretation is important because educational manuscripts often convert small descriptive differences into claims about effectiveness or superiority. The evidence here supports a different conclusion: both classes display a broadly similar profile. In both, neatness and relevance are comparatively strong, while organization and language require more attention. This shared pattern is more pedagogically actionable than the small difference in class means.

Implications for first-year PIAUD curriculum design

The findings support an integrated writing-development model within the first-semester curriculum. First, students should receive a concise academic paper guide aligned with the actual assessment criteria. The guide should include annotated examples, not only formatting rules. Second, lecturers should explicitly teach the paper as a genre: its purpose, expected

macro-structure, paragraph functions, and typical evidence practices. Third, source use should be taught through short exercises in searching, evaluating, summarizing, paraphrasing, citation, and synthesis before the final paper is assigned. Fourth, a reference manager such as Zotero or Mendeley can be introduced for bibliographic organization, while emphasizing that software does not replace judgment about source relevance or accurate source integration.

Fifth, the assignment should include at least one draft-feedback-revision cycle. Feedback can be organized around the same five dimensions used for final assessment, with priorities limited to a manageable number of revisions. Sixth, students can participate in structured peer review using an analytic rubric and two sample papers. This can strengthen their capacity to notice organization, evidence, and language in their own work. Finally, the course should coordinate writing expectations with other first-year PIAUD subjects. Academic literacy develops more effectively when students repeatedly practice related conventions across meaningful disciplinary tasks rather than encountering writing only in a stand-alone language course (Khumalo & Reddy, 2021; Lea, 2004; Wingate, 2012).

Limitations

Several limitations constrain the conclusions. The corpus contained only 14 group papers from two classes in one program and one institution. The findings are therefore not generalizable to all PIAUD students or Indonesian undergraduates. Because the assignments were group products, the scores cannot reveal individual writing competence or identify how responsibilities were distributed among group members. The study also analyzed only final products; it did not observe planning, drafting, peer interaction, feedback uptake, or revision processes.

Measurement is another limitation. The five criteria are educationally plausible, but the source manuscript provides no evidence of instrument validation, analytic scoring descriptors, rater calibration, or inter-rater reliability. Future studies should use a clearly defined analytic rubric, involve at least two trained raters, and report agreement indices. If individual papers are analyzed, researchers could also examine relationships among rubric dimensions, prior writing experience, information literacy, self-efficacy, and feedback engagement.

The qualitative observations in the source manuscript were useful but not based on a systematic error-coding protocol. Future research could code linguistic errors, paragraph organization, citation functions, source types, and reference-list accuracy at a finer level. A longitudinal design would be particularly valuable: students' papers could be sampled early and late in the first year to examine development. Intervention studies could then test whether genre-based instruction, source-use workshops, rubric-guided peer review, or integrated writing support produces measurable improvement. Such designs would move beyond descriptive profiling toward stronger evidence about how academic writing develops in PIAUD and related teacher-education programs.

CONCLUSION

This study re-examined 14 academic group papers written by first-semester PIAUD students at UIN Sunan Kalijaga and produced a more methodologically cautious profile of their academic writing performance. The papers achieved an overall mean of 81.07, with 71.43% classified as good and 28.57% as fair. All papers met the minimum passing score. However, the dimension-level analysis shows that this generally positive result should not be read as evidence of advanced academic writing competence. Presentation/neatness and thematic relevance were the strongest areas, whereas organization and language use were the weakest. The qualitative evidence further indicates ongoing difficulties with academic conventions, depth of discussion, consistent structure, formatting, citation, and source use.

The main pedagogical implication is that first-year writing should be approached as an academic-literacy development process rather than as a one-time assignment. Students need explicit instruction in the paper genre, transparent analytic criteria, guided practice in paraphrasing and source integration, reference-management support, and iterative feedback followed by revision. These supports are likely to be more productive when embedded across first-year disciplinary courses rather than confined to a single Indonesian Language class.

The study also demonstrates the importance of aligning claims with the unit of analysis. Because the corpus consisted of group papers, the findings profile the quality of collective written products rather than individual student competence. Similarly, because the scoring instrument did not report validation or inter-rater reliability, the numerical results are best used diagnostically, not as high-precision measurement. Future research should analyze individual and longitudinal writing samples using validated rubrics and trained multiple raters, while examining writing processes and source-use practices in greater depth. Within these limits, the findings provide a useful baseline for strengthening academic writing instruction among new PIAUD students and for designing more systematic writing support in Indonesian teacher education.

REFERENCES

Azizah, U. A., & Budiman, A. (2017). Challenges in writing academic papers for international publication among Indonesian graduate students. *JEELS (Journal of English Education and Linguistics Studies), 4*(2), 175–197. https://doi.org/10.30762/jeels.v4i2.405

Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback. *Assessment & Evaluation in Higher Education, 43*(8), 1315–1325. https://doi.org/10.1080/02602938.2018.1463354

Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction, 20*(4), 328–338. https://doi.org/10.1016/j.learninstruc.2009.08.006

Cumming, A., Lai, C., & Cho, H. (2016). Students' writing from sources for academic purposes: A synthesis of recent research. *Journal of English for Academic Purposes, 23*, 47–58. https://doi.org/10.1016/j.jeap.2016.06.002

Davis, M. (2013). The development of source use by international postgraduate students. *Journal of English for Academic Purposes, 12*(2), 125–135. https://doi.org/10.1016/j.jeap.2012.11.008

Doolan, S. M. (2021). An exploratory analysis of source integration in post-secondary L1 and L2 source-based writing. *English for Specific Purposes, 62*, 128–141. https://doi.org/10.1016/j.esp.2021.01.003

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Howitt, S. M., Wilson, A. N., & Higgins, D. M. (2024). Unlearning, uncovering and becoming: Experiencing academic writing as part of undergraduate research. *Teaching in Higher Education, 29*(5), 1322–1337. https://doi.org/10.1080/13562517.2022.2145466

Huisman, B., Saab, N., van Driel, J., & van den Broek, P. (2018). Peer feedback on academic writing: Undergraduate students' peer feedback role, peer feedback perceptions and essay performance. *Assessment & Evaluation in Higher Education, 43*(6), 955–968. https://doi.org/10.1080/02602938.2018.1424318

Hyland, K. (2003). Genre-based pedagogies: A social response to process. *Journal of Second Language Writing, 12*(1), 17–29. https://doi.org/10.1016/S1060-3743(02)00124-8

Jönsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130–144. https://doi.org/10.1016/j.edurev.2007.05.002

Keck, C. (2006). The use of paraphrase in summary writing: A comparison of L1 and L2 writers. *Journal of Second Language Writing, 15*(4), 261–278. https://doi.org/10.1016/j.jslw.2006.09.006

Khumalo, N. P., & Reddy, S. (2021). Cross-disciplinary synergy: First-year students' experiences of learning academic writing through integrated writing support at a University of Technology. *The Journal for Transdisciplinary Research in Southern Africa, 17*(1), a1038. https://doi.org/10.4102/td.v17i1.1038

Lea, M. R. (2004). Academic literacies: A pedagogy for course design. *Studies in Higher Education, 29*(6), 739–756. https://doi.org/10.1080/0307507042000287230

Lea, M. R., & Street, B. V. (1998). Student writing in higher education: An academic literacies approach. *Studies in Higher Education, 23*(2), 157–172. https://doi.org/10.1080/03075079812331380364

Lee, J. J., Hitchcock, C., & Casal, J. E. (2018). Citation practices of L2 university students in first-year writing: Form, function, and stance. *Journal of English for Academic Purposes, 33*, 1–11. https://doi.org/10.1016/j.jeap.2018.01.001

Mansourizadeh, K., & Ahmad, U. K. (2011). Citation practices among non-native expert and novice scientific writers. *Journal of English for Academic Purposes, 10*(3), 152–161. https://doi.org/10.1016/j.jeap.2011.03.004

Mitchell, K. M., Zumbrunn, S., Berry, D. N., & Demczuk, L. (2023). Writing self-efficacy in postsecondary students: A scoping review. *Educational Psychology Review, 35*, Article 82. https://doi.org/10.1007/s10648-023-09798-2

Nenotek, S. A., Tlonaen, Z. A., & Manubulu, H. A. (2022). Exploring university students' difficulties in writing English academic essay. *Al-Ishlah: Jurnal Pendidikan, 14*(1), 909–920. https://doi.org/10.35445/alishlah.v14i1.1352

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199–218. https://doi.org/10.1080/03075070600572090

Pecorari, D. (2003). Good and original: Plagiarism and patchwriting in academic second-language writing. *Journal of Second Language Writing, 12*(4), 317–345. https://doi.org/10.1016/j.jslw.2003.08.004

Petrić, B. (2007). Rhetorical functions of citations in high- and low-rated master's theses. *Journal of English for Academic Purposes, 6*(3), 238–253. https://doi.org/10.1016/j.jeap.2007.09.002

Prat-Sala, M., & Redford, P. (2012). Writing essays: Does self-efficacy matter? The relationship between self-efficacy in reading and in writing and undergraduate students' performance in essay writing. *Educational Psychology, 32*(1), 9–20. https://doi.org/10.1080/01443410.2011.621411

Ramoroka, B. T. (2014). Integration of sources in academic writing: A corpus-based study of citation practices in essay writing in two departments at the University of Botswana. *Reading & Writing, 5*(1), a41. https://doi.org/10.4102/rw.v5i1.41

Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education, 35*(4), 435–448. https://doi.org/10.1080/02602930902862859

Roald, G. M., Wallin, P., Hybertsen, I. D., & Stenøien, J. M. (2021). Learning from contrasts: First-year students writing themselves into academic literacy. *Journal of Further and Higher Education, 45*(6), 758–770. https://doi.org/10.1080/0309877X.2020.1813264

Salam, H. B. (2020). Strategi menulis karya ilmiah model pembelajaran kolaboratif berbasis problem based learning pada mahasiswa Administrasi Bisnis Politeknik Informatika Nasional. *Jurnal Onoma: Pendidikan, Bahasa, dan Sastra, 6*(2), 693–703. https://doi.org/10.30605/onoma.v6i2.428

Sanders-Reio, J., Alexander, P. A., Reio, T. G., Jr., & Newman, I. (2014). Do students' beliefs about writing relate to their writing self-efficacy, apprehension, and performance? *Learning and Instruction, 33*, 1–11. https://doi.org/10.1016/j.learninstruc.2014.02.001

Septafi, G. (2021). Analisis kemampuan menulis artikel ilmiah mahasiswa Pendidikan Guru Sekolah Dasar angkatan 2019. *Educational Technology Journal, 1*(2), 1–16. https://doi.org/10.26740/etj.v1n2.p1-16

Shi, L. (2004). Textual borrowing in second-language writing. *Written Communication, 21*(2), 171–200. https://doi.org/10.1177/0741088303262846

Sun, T., & Wang, C. (2020). College students' writing self-efficacy and writing self-regulated learning strategies in learning English as a foreign language. *System, 90*, 102221. https://doi.org/10.1016/j.system.2020.102221

Sun, T., Wang, C., Lambert, R. G., & Liu, L. (2021). Relationship between second language English writing self-efficacy and achievement: A meta-regression analysis. *Journal of Second Language Writing, 53*, 100817. https://doi.org/10.1016/j.jslw.2021.100817

Syaputra, J., Damayanti, V. S., Anshori, D. S., & Sastromihardjo, A. (2023). Literasi data: Dalam menulis karya ilmiah di perguruan tinggi. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya, 9*(1), 204–212. https://doi.org/10.22219/kembara.v9i1.23883

Syazali, M., & Erfan, M. (2022). Keterampilan menulis mahasiswa PGSD melalui implementasi template makalah sebagai pedoman penyusunan. *Research and Development Journal of Education, 8*(1), 35–45. https://doi.org/10.30998/rdje.v8i1.10255

Teng, L. S., Sun, P. P., & Xu, L. (2018). Conceptualizing writing self-efficacy in English as a foreign language contexts: Scale validation through structural equation modeling. *TESOL Quarterly, 52*(4), 911–942. https://doi.org/10.1002/tesq.432

Thompson, C., Morton, J., & Storch, N. (2013). Where from, who, why and how? A study of the use of sources by first year L2 university students. *Journal of English for Academic Purposes, 12*(2), 99–109. https://doi.org/10.1016/j.jeap.2012.11.004

Wette, R. (2017). Source text use by undergraduate post-novice L2 writers in disciplinary assignments: Progress and ongoing challenges. *Journal of Second Language Writing, 37*, 46–58. https://doi.org/10.1016/j.jslw.2017.05.015

Widodo, A., Jaelani, A. K., Novitasari, S., Sutisna, D., & Erfan, M. (2020). Analisis kemampuan menulis makalah mahasiswa baru PGSD Universitas Mataram. *Jurnal DIDIKA: Wahana Ilmiah Pendidikan Dasar, 6*(1), 77–91. https://doi.org/10.29408/didika.v6i1.1946

Wingate, U. (2010). The impact of formative feedback on the development of academic writing. *Assessment & Evaluation in Higher Education, 35*(5), 519–533. https://doi.org/10.1080/02602930903512909

Wingate, U. (2012). Using academic literacies and genre-based models for academic writing instruction: A 'literacy' journey. *Journal of English for Academic Purposes, 11*(1), 26–37. https://doi.org/10.1016/j.jeap.2011.11.006