Reading & Writing - Journal of the Literacy Association of South Africa
ISSN: (Online) 2308-1422, (Print) 2079-8245

Critical digital literacy enacted by Indonesian first-year undergraduates: A corpus-assisted study

Check for updates

Authors:
James E. Lalira[1]
Maya P. Warouw[2]

Affiliations:
[1]Department of English Language Education, Faculty of Teachers' Training and Education, Christian University of Indonesia in Tomohon, Tomohon, Indonesia

[2]English Department, Faculty of Humanities, Sam Ratulangi University, Manado, Indonesia

Corresponding author:
James Lalira,
jameslalira@fkipukit.ac.id

Dates:
Received: 16 May 2026
Accepted: 08 Sept. 2026
Published: 15 Oct. 2026

How to cite this article:
Lalira, J.E. & Warouw, M.P., 2026, 'Critical digital literacy enacted by Indonesian first-year undergraduates: A corpus-assisted study', *Reading & Writing* 17(1), a700. https://doi.org/10.4102/rw.v17i1.700

Copyright:
© 2026. The Authors. Licensee: AOSIS. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license (https://creativecommons.org/licenses/by/4.0/).

Read online:
Scan this QR code with your smart phone or mobile device to read online.

Background: Problematic digital information increasingly shapes the literacy demands placed on first-year undergraduates, yet limited evidence explains how students demonstrate critical evaluation in written responses to unsupported claims, partial news items and distrust narratives expressing distrust of mainstream media.

Objectives: This study examined how critical digital literacy was manifested in written evaluative discourse of Indonesian first-year undergraduates and how those patterns varied across stimulus types and study programmes.

Method: A cross-sectional descriptive quantitative design was used with 70 first-year students in their first semester from seven study programmes at two Indonesian higher education institutions. Participants were recruited through classroom-based non-probability quota sampling. Each participant produced three written responses, yielding 210 texts and 41 864 tokens. Quantitative content analysis of four binary dimensions and descriptive statistics were complemented by corpus-assisted concordance analysis.

Results: Most responses were classified at the moderate Critical Digital Literacy Score level. Credibility evaluation was the most frequently observed dimension, whereas bias awareness was the least frequent. Unsupported health claims elicited more high-level responses than narratives expressing distrust of mainstream media, and higher-scoring responses contained a greater density of evaluative lexical markers.

Conclusion: Critical digital literacy was not uniform. It appeared to be a multi-layered, partial, and context-specific literacy practice, as observed in students' written evaluations and distinct from their real-time online verification behaviour.

Contribution: This study provides a literacy-oriented, corpus-assisted account of how credibility judgement, verification, reasoning, and bias recognition became visible in undergraduate writing in a Global South higher education context.

Keywords: bias awareness; corpus-assisted discourse analysis; credibility evaluation; critical digital literacy; higher education; Indonesian undergraduates; misinformation; written evaluative discourse.

Introduction

Digital environments have become the primary context in which higher education students access information, including search engines, social media feeds, messaging apps, artificial intelligence (AI) tools, and algorithmically driven news feeds. Literacy is not simply the ability to use devices or platforms. It also involves judgement, interpretation, verification and the ability to identify how digital texts situate readers. Thus, digital literacy in higher education can be defined as a social and academic practice that shapes how students read, write and evaluate knowledge claims (Farias-Gaytan, Aguaded & Ramírez-Montoya 2023; Goodfellow 2011; Ng 2012). This more expansive view has grown in importance as university students increasingly encounter unsupported health claims, partial headlines, decontextualised evidence, AI-generated learning content and narratives of distrust that challenge the authority of trusted information sources. For instance, in the context of Indonesian language learning, AI tools have been found to affect grammar learning and students' learning outcomes (Lalira et al. 2024), and in such learning environments, verification, evidence and source evaluation should receive greater attention. In addition to technical skills, critical digital literacy (CDL) needs to be addressed in terms of credibility, verification, reasoning and ideological framing (Gutiérrez-Ujaque 2024; Pangrazio 2016; Smith & Storrs 2023). This need is reinforced by Indonesian scholarship. Silvhiany, Huzaifah and Ismet (2021) reveal that evaluating online sources requires explicit support from CDL, and

Iskandar et al. (2022) indicate that digital literacy must be infused deliberately into authentic academic digital practices. The same finding was reported by Hardianti (2024), who found that students in higher education in Indonesia perceive critical literacy as important. Still, its implementation in the classroom and academic activities remains suboptimal. Explicit critical literacy awareness was also highlighted by Mbau and Sugeng (2019) to enable learners to question texts, ideas and communicative purposes in Indonesian English language teaching (ELT) contexts rather than simply processing information. In conjunction with the findings from AI-mediated learning environments (Lalira et al. 2024), these studies situate the current research in a local need for evaluative, verification-oriented and discourse-based literacy instruction. There has been a significant increase in research on digital literacy; however, it remains a relatively broad topic. Systematic reviews have associated digital literacy with competence, skills, transformation and digital thinking, while also noting that assessment practices remain heterogeneous across educational contexts (Farias-Gaytan et al. 2023; Sillat, Tammets & Laanpere 2021; Tinmaz et al. 2022). This is helpful for field mapping, but it can complicate the determination of what students do when they assess problematic digital information (PDI) in written assignments. Literacy-oriented research therefore requires evidence not only about students' perceived competence but also about how evaluative judgement is enacted in discourse. A second strand of scholarship has examined online information evaluation, source evaluation and critical online reasoning. This body of work shows that the features of the source, relevance, quality of the evidence, argumentative coherence and lateral comparison across sources (Breakstone et al. 2021; Coiro 2011; Hahnel, Eichmann & Goldhammer 2020; Marttunen, Salminen & Utriainen 2021; McGrew et al. 2018; Molerov et al. 2020; Polizzi 2020) are all involved in credibility judgements. Other research has found that social and cognitive contexts affect students' online information behaviour, even when they understand the importance of credibility (Metzger, Flanagin & Zwarun 2003). Digital access and digital confidence should not be considered evidence of CDL. Misinformation and media-literacy research reinforces this concern further. Experimental and meta-analytic studies show that targeted interventions can improve fake-news credibility assessment, misinformation discernment and resilience to misleading content (Allcott & Gentzkow 2017; Berger et al. 2025; Ferrucci & Hopp 2023; Huang, Jia & Yu 2024; Jones-Jang, Mortensen & Liu 2021; Kahne & Bowyer 2017; Lu et al. 2024). Research on misinformation also reveals that mere exposure to correct information is insufficient: Evaluative habits are needed to facilitate source comparison, evidential reasoning and resistance to plausibility-based acceptance (Kozyreva, Lewandowsky & Hertwig 2020; Lewandowsky et al. 2012; Pennycook & Rand 2021). Although significant progress has been made, three gaps remain. First, many studies focus on self-report measures, general competence mapping, or structured assessment instruments, while fewer studies focus on students' own written evaluative responses as literacy artefacts. Second, the internal structure of CDL is often treated as a general competence, rather than as a collection of related but distinct practices. Third, there is a limited understanding of the level of criticality that different types of problematic information evoke, particularly among first-year students in under-represented higher education settings. The conceptual framework of the current study seeks to fill these gaps and conceptualises CDL as a written evaluative practice made visible through four interconnected dimensions. Credibility evaluation (CE) indicates students' assessment of the credibility of a claim or source. Information verification (IV) is used to record whether they are stating procedures for checking, comparing or consulting authoritative sources. Critical reasoning (CR) determines whether their judgement is related to evidence, logic or inference. Bias awareness (BA) records whether they are aware of framing, sensationalism, one-sided presentation or attempts to guide interpretation. These dimensions are analytically distinct, but together they form a compact discourse-based model of how students write critically. The Indonesian Higher Education context is a suitable setting for analysing this model. When entering university, students are already engaged in digital information ecologies. However, they may not have received consistent academic instruction in credibility assessment, verification processes, evidence-based reasoning or bias analysis. The focus on first-year undergraduates allows for an early snapshot of an evaluative repertoire before the effects of longer disciplinary socialisation. This is analytically relevant because familiarity with digital technologies does not equate to CDL. This study aimed to identify the pattern of CDL of first-year students who were asked to respond to three types of PDI texts: an unsupported health claim, a clickbait or partial news item, and a distrust narrative about mainstream media, using corpus-assisted analysis of written responses. It combined manual coding and corpus-assisted evidence from students' evaluative lexis to demonstrate how criticality was manifested in discourse. To answer the two research questions, this study focused on the following: (1) *How was CDL enacted in Indonesian first-year undergraduates' written evaluative discourse?* and (2) *How do patterns of CDL vary across stimulus types and study programmes?*

Research methods and design

Study design

This study employed a cross-sectional descriptive quantitative design. The quantitative tabulation component comprised manual binary content coding of the four Critical Digital Literacy Score (CDLS) dimensions, descriptive statistical tabulation and frequency-based corpus measures. Concordance-oriented analysis of evaluative lexical markers in their local contexts was used as an interpretive corpus-assisted procedure. The quantitative results established the distributional patterns, while the concordance analysis showed how evaluative practices were realised linguistically (O'Brien et al. 2014; Von Elm et al. 2007). Causal interventions were not tested. This study describes how first-year undergraduate students expressed critical evaluation in a controlled environment with written responses.

Stimulus- and programme-level patterns are reported descriptively and interpreted as tendencies at the response level, not as inferential evidence of disciplinary or stimulus effects. This reporting decision is based on the premise that repeated responses from the same participant should not be treated as fully independent observations without an analytic model designed for that purpose.

Setting, participants and sampling

The study was conducted with first-semester first-year undergraduate students, who were at an early stage in the formation of university literacy. The participants were from seven study programmes: (1) English Education, (2) Economics Education, (3) Pharmacy, (4) Biology, (5) Theology, (6) Diploma in Nursing, and (7) Bachelor of Health Administration. The sampling method was non-probability sampling. Students could participate if they were first-year undergraduates in one of the participating classes, attended the data collection session, provided written informed consent and completed all three written-response tasks. Students who refused to participate or did not complete the responses were excluded from the analysis. There were 70 students in total, with 10 students from each study programme. At the institutional level, 40 participants were from Institution 1, and 30 were from Institution 2. There were no incomplete response sets among any consenting participants; therefore, all 70 participants and 210 responses were retained for analysis. As this study did not test between-group hypotheses or estimate population parameters, no a priori power calculation was performed. Instead, sample adequacy was justified analytically and pragmatically. The sample in Table 1 was suitable for the descriptive, coding-based and corpus-assisted purposes of the study as it yielded a balanced structure of responses across programmes and stimuli and a corpus of 41 864 tokens. The study did not claim statistical representativeness of Indonesian undergraduates as a population. Instead, it sought to describe written evaluative performance among an early undergraduate cohort in an analytically transparent manner.

TABLE 1: Participant profile and corpus characteristics (N = 70).

Variable	Category	Frequency (n)	%
Institution	Institution 1	40	57.1
	Institution 2	30	42.9
Study programmes	English Education	10	14.3
	Economics Education	10	14.3
	Pharmacy	10	14.3
	Biology	10	14.3
	Theology	10	14.3
	Diploma in Nursing	10	14.3
	Bachelor of Health Administration	10	14.3
Gender	Male	26	37.1
	Female	44	62.9
Daily internet use	< 2 h	3	4.3
	2–4 h	17	24.3
	5–7 h	27	38.6
	> 7 h	23	32.9
Checking information before sharing	Yes	22	31.4
	No	12	17.1
	Sometimes	36	51.4
Corpus	Written responses	210	100.0
	Tokens after cleaning	41 864	-

Note: Participant profile percentages were calculated using N = 70 as the denominator. Each participant produced three responses; therefore, the written-response total was calculated using N = 210 texts. Here, n denotes frequency. The token count represents corpus size and was not converted into a percentage.

Materials, stimulus sequence and written-response task

The materials comprised three brief stimuli that represented recurring PDI: (1) an unsupported health claim, (2) clickbait or partial news, and (3) a distrust narrative about mainstream media. The stimuli were based on the recurrent misinformation genres described in the literature and were not used as a psychometrically validated instrument. They were designed to elicit comparable written evaluations across three information contexts congruent with the four analytic dimensions of the study. The three stimuli were mapped analytically against the four dimensions of the coding framework: (1) CE, (2) IV, (3) CR and (4) BA. They were designed to be accessible to first-year students in all programmes and to make no special demands on specialist knowledge of the disciplines. The stimuli were presented in the same order to all participants: (1) unsupported health claims, (2) clickbait or partial news items, and (3) media-distrust narratives. The sequence was not counterbalanced; therefore, any differences in the stimulus level were seen as descriptive tendencies and were not reported as a causal effect of the stimulus level. One response was written for each stimulus, with a length of 150 words–250 words. The prompt required students to assess the information, provide their assessment and explain what should be done before accepting or sharing it. There was no requirement for or systematic monitoring of internet use in the task protocol. No browsing logs or observations were recorded. The task was thus given as a written evaluation exercise, not as a fact-checking exercise in real time. The study elicited written evaluative performance; it did not measure whether participants actually verified the content online in real time. The responses were originally written in Indonesian. The marker families reported in the corpus analysis are lexical items in Indonesian. Excerpts and explanatory descriptions included in the article were translated into English for publication and verified against the original responses to ensure the accuracy of the credibility judgements, verification, reasoning and bias recognition.

Corpus preparation and processing

All responses were anonymised before analysis and saved as plain text. Data cleaning included the deletion of personal information, normalisation of spacing and punctuation, and light orthographic normalisation without altering meaning. The responses were kept as individual texts and were assigned to subcorpora based on their CDLS levels in the density comparisons. Word-level tokenisation was performed. Lemmatisation was not exhaustive, but was implemented for Indonesian, a productive affixation language. Analytically

equivalent forms were classified into marker families: (1) *sumber* [source], (2) *bukti* [evidence], (3) *cek/periksa/verifikasi* [check/examine/verify], (4) *data* [data], (5) *resmi* [official] and (6) *bias/clickbait/sensasional* [bias/clickbait/sensational]. The raw frequencies were presented for the entire corpus, and the evaluative density was normalised per 1000 tokens to compare levels. In addition to frequency, markers were inspected using concordance-oriented analysis to examine their functioning in local contexts (Gillings & Mautner 2024). The corpus-assisted layer was not considered a validation of the coding framework but as interpretive support. This boundary was required because some marker families, particularly those related to sources, evidence and bias, partially overlapped with the conceptual content of the coding dimensions. A lexical approach was thus used to demonstrate the critical importance of textured language, while keeping manual coding as the primary classification method.

Coding framework and score construction

The complete written responses were the coding units. Clauses and sentences were used as evidential units, and final scores were assigned at the response level. There were four binary dimensions: (1) CE, (2) IV, (3) CR and (4) BA. The dimensions were scored as 0 if absent and 1 if present. The CDLS was used as a descriptive index, which was calculated by adding the four dimensions, resulting in a score of 0–4. The scores were categorised as Low (0–1), Moderate (2) and High (3–4). The four dimensions were treated equally, as they were considered foundational and complementary literacy practices rather than hierarchical, as shown in Table 2. This decision made the scores transparent and replicable. It also ensured that one component of performance (e.g. initial credibility judgement) did not overshadow another (e.g. verification or bias recognition). Discursive response type (DRT) was coded separately as Accepting, Questioning, Rejecting or Neutral. The DRT was not used in the CDLS calculation but was used descriptively to represent the stance.

TABLE 2: Analytic framework and score interpretation.

Dimension or category	Operational definition	Coding decision
Credibility evaluation (CE)	Explicit judgement about the credibility, trustworthiness or reliability of a claim, source or text.	0 = absent; 1 = present
Information verification (IV)	Mention of checking, comparing, consulting official sources, searching for evidence or verifying before accepting or sharing.	0 = absent; 1 = present
Critical reasoning (CR)	A traceable reason connecting judgement to evidence, logic, inconsistency, source quality or inferential support.	0 = absent; 1 = present
Bias awareness (BA)	Recognition of bias, framing, one-sided presentation, sensationalism, clickbait or attempts to steer interpretation.	0 = absent; 1 = present
CDLS	Sum of CE, IV, CR and BA.	Range = 0–4
Level	Interpretive grouping of CDLS.	Low = 0–1; Moderate = 2; High = 3–4
Discursive response type	Overall stance towards the content.	Accepting; Questioning; Rejecting; Neutral

Note: DRT was used as a descriptive stance indicator and was not included in the CDLS calculation. The four CDLS dimensions are binary coding indicators, not Likert-scale items. CDLS is reported as a descriptive additive index rather than a psychometrically validated scale. DRT is a separate nominal category and is not included in the CDLS calculation.

CDLS, Critical Digital Literacy Score; DRT, discursive response type.

Inter-coder reliability and analytic procedure

Two coders used a common codebook containing operational definitions, decision rules and coding examples. A calibration phase was conducted before the formal reliability audit to align the coders' interpretations. A 20% subset of the corpus, comprising 42 responses, was then independently double-coded. Cohen's kappa was calculated separately for the four binary CDLS dimensions. The coefficients were 0.82 for CE, 0.78 for IV, 0.76 for CR and 0.74 for BA, indicating substantial to almost perfect inter-coder agreement. Observed agreement ranged from 83.3% to 90.5%. Disagreements were resolved through discussion and re-examination of the codebook, after which consensus codes were used to calculate the final CDLS values. The reliability audit assessed the reproducibility of the four component coding decisions. It was not intended as evidence of the internal consistency or psychometric validation of CDLS as a unidimensional scale. An internal consistency coefficient such as Cronbach's alpha was therefore not treated as the appropriate reliability criterion because the four dimensions represented analytically distinct but complementary practices. Discursive response type categories were assigned after the four binary dimensions had been coded and were checked subsequently through coder discussion. Because a separate independent kappa coefficient was not calculated for DRT, the categories were used only as descriptive indicators of stance. The analysis proceeded in three stages. First, descriptive statistics summarised participant characteristics, CDLS distributions, dimension frequencies, DRT profiles and response patterns by stimulus and programme. Second, corpus analysis identified lexical-marker frequencies and normalised marker densities across CDLS levels. Third, concordance-oriented interpretation connected the lexical evidence with the manual coding results. No inferential statistical tests were conducted because the study was not designed as an experiment or a representative population survey.

Ethical considerations

Ethical clearance to conduct this study was obtained from the Research Ethics Committee, Universitas Kristen Indonesia Tomohon (No. 15/LPPM-UKIT/I/2026). Participation was voluntary, and written informed consent was obtained before data collection. Students were informed that participation was not related to course grades, academic standing, or institutional services. The procedures adhered to the institutional ethical guidelines for educational research involving human subjects. Responses were anonymised before analysis, identifiers were coded, and no personally identifiable information was retained in the analytic corpus or reported excerpts.

Results

Overall Critical Digital Literacy Score profile

The complete response corpus consisted of 210 texts, all produced by the 70 participants and all of which were analysable. Overall, the distribution of CDLS results showed that CDL most commonly occurred at a moderate level. The highest number of responses, 94 (44.8%) was for a score of 2. Low-level scores were also substantial: (1) 60 responses (28.6%) received a score of 1, and (2) 8 responses (3.8%) received a score of 0. When score 3 and score 4 were combined, a high-level performance was observed in 48 responses (22.9%). The mean CDLS was 2.05 (standard deviation [SD] = 1.10; median = 2). This distribution as shown in Table 3 indicates that the students who participated in this study were not uncritical generally. Many could identify credibility issues or say they are wary. However, relatively few responses consistently combined credibility judgement with explicit verification, reasoning and bias recognition. The predominant pattern was thus partial criticality, not absence or mastery of criticality.

Dimension level and discursive patterns

The results at the dimension level showed a clear frequency ordering. Credibility Evaluation was the most frequently observed dimension, with 146 responses (69.5%). There were 106 (50.5%) responses to CR and 101 (48.1 %) responses to IV. Bias Awareness was the lowest dimension, with only 77 responses (36.7%). This pattern indicates that while many students can recognise questionable information, fewer can describe verification procedures or recognise the framing, partiality or sensationalism of the information. The DRT results provide a complementary view of the stance (see Table 4). Questioning was the most common type of response, followed by rejection. In almost a quarter of the corpus, accepting responses remained visible. The results indicate that caution was used frequently, but not necessarily as part of a complete critical sequence. Students might question a text without necessarily explaining how to verify the information or how the rhetorical context affects meaning.

TABLE 3: Overall Critical Digital Literacy Score distribution and dimension profile (N = 210).

Panel	Category	*n*	%
CDLS	0	8	3.8
	1	60	28.6
	2	94	44.8
	3	10	4.8
	4	38	18.1
CDLS level	Low (0–1)	68	32.4
	Moderate (2)	94	44.8
	High (3–4)	48	22.9
Dimension present	Credibility evaluation	146	69.5
	Information verification	101	48.1
	Critical reasoning	106	50.5
	Bias awareness	77	36.7

Note: Mean CDLS = 2.05; SD = 1.10; median = 2.
CDLS, Critical Digital Literacy Score; SD, standard deviation.

TABLE 4: Discursive response types in the full corpus (N = 210).

Discursive response type	*n*	%	Interpretive meaning
Questioning	70	33.3	Expressed doubt or caution without necessarily rejecting the content
Rejecting	62	29.5	Dismissed the content as unreliable, biased or unsupported
Accepting	48	22.9	Treated the content as plausible or true
Neutral	30	14.3	Remained descriptive or non-committal

Note: DRT is reported descriptively and is not part of the CDLS calculation.
CDLS, Critical Digital Literacy Score; DRT, discursive response type.

Variation by stimulus type

The largest difference in the stimulus level was for the media-distrust narrative. The unsupported health claim had the highest percentage of high responses (30.0%), while the media-distrust narrative had the highest percentage of low responses (38.6%). The clickbait or partial news stimulus was dominated by moderate responses (50.0%). These patterns indicate that students were more prepared to question a visibly unsupported health claim than to challenge the credibility of mainstream media. The latter seems to be a more challenging form of CDL, as students need to be sceptical not only of media institutions but also of the alternative claims themselves. The results shown in Table 5 are reported descriptively because the fixed stimulus sequence and repeated-response design do not allow causal comparisons across stimuli.

Variation by study programmes

There was also variation at the programme level. The highest percentage of high responses was in English Education (33.3%) and Pharmacy (30.0%). Diploma in Nursing and Bachelor of Health Administration had the lowest percentage of High (13.3% each) and the highest percentage of Low (40.0% each). Biology showed the strongest concentration in the Moderate category, whereas Theology showed the most polarised distribution. These differences in Table 6 should be interpreted with caution as they are descriptive response-level distributions rather than inferential comparisons. Despite this, the balanced programme structure ensured that the patterns were not due to unequal programme counts. They refer to the potential for early disciplinary literacy experiences to be reflected in students' credibility judgements, verification procedures and bias analyses.

Corpus-assisted lexical evidence and reliability

The coding results were complemented by a linguistic layer through corpus-assisted analysis. The most frequent evaluative marker families were *sumber* [source] (162 occurrences), *bukti* [evidence] (139), *cek/periksa/verifikasi* [check/examine/verify] (127), data (88), *resmi* [official] (84), and *bias/clickbait/sensasional* [bias/clickbait/sensational] (71). These markers indicate that students often used source-, evidence-, and verification-related language, while explicit bias-related lexis was less common. An increase in the density of evaluative markers was observed at the CDLS levels. High-level responses contained 25.1 evaluative markers per 1000 tokens, moderate responses contained 15.4, and low responses

TABLE 5: Critical Digital Literacy Score levels by stimulus type.

Stimulus type	High		Moderate		Low		Total
	n	%	*n*	%	*n*	%	
Unsupported health claim	21	30.0	29	41.4	20	28.6	70
Clickbait or partial news item	14	20.0	35	50.0	21	30.0	70
Media-distrust narrative	13	18.6	30	42.9	27	38.6	70
Total	**48**	-	**94**	-	**68**	-	**210**

Note: Percentages are based on 70 responses per stimulus. The stimulus sequence was fixed; therefore, the results are descriptive tendencies rather than causal effects.

TABLE 6: Critical Digital Literacy Score levels by study programme.

Study programme	High		Moderate		Low		Total responses
	n	%	*n*	%	*n*	%	
English Education	10	33.3	12	40.0	8	26.7	30
Economics Education	8	26.7	14	46.7	8	26.7	30
Pharmacy	9	30.0	13	43.3	8	26.7	30
Biology	7	23.3	15	50.0	8	26.7	30
Theology	6	20.0	12	40.0	12	40.0	30
Diploma in Nursing	4	13.3	14	46.7	12	40.0	30
Bachelor of Health Administration	4	13.3	14	46.7	12	40.0	30
Total	**48**	-	**94**	-	**68**	-	**210**

Note: Each programme contributed 10 participants and 30 responses. Programme-level patterns are descriptive.

TABLE 7: Corpus-assisted lexical evidence and reliability audit.

Panel	Item	Value	Interpretive focus
Marker frequency	*Sumber* [source]	162	Source evaluation
	Bukti [evidence]	139	Evidence orientation
	cek/periksa/verifikasi [check/examine/verify]	127	Verification procedure
	data	88	Evidence or factual support
	Resmi [official]	84	Official or authoritative source
	bias/ clickbait/ sensasional [bias/ clickbait/sensational]	71	Bias, framing or sensationalism
Density per 1000 tokens	High CDLS	25.1; token denominator = 9562	Dense evaluative language
	Moderate CDLS	15.4; token denominator = 18 744	Partial evaluative language
	Low CDLS	7.2; token denominator = 13 558	Thin evaluative language

Note: Marker families are reported in Indonesian because responses were originally written in Indonesian. The token denominators sum to 41 864 tokens. Reliability audit: CE, kappa = 0.82, agreement = 90.5%; IV, kappa = 0.78, agreement = 88.1%; CR, kappa = 0.76, agreement = 85.7%; BA, kappa = 0.74, agreement = 83.3%.

CDLS, Critical Digital Literacy Score; CE, credibility evaluation; IV, information verification; CR, critical reasoning; BA, bias awareness.

contained 7.2, as shown in Table 7. Concordance-oriented inspection revealed that high responses were more likely to include markers in more complete argumentative sequences, while moderate responses were more likely to include generalised caution, and low responses were more likely to include plausibility, popularity or impressionistic acceptance. The reliability audit confirmed the reliability of these four binary dimensions. The kappa values varied from 0.74 to 0.82, and the observed agreement varied from 83.3% to 90.5%. These findings confirm the stability of the CDLS dimensions and that the individual DRT categories are still descriptive stance indicators.

Discussion

The findings suggest that perception of CDL among the participating Indonesian undergraduates was uneven rather than absent. The majority of responses were at the moderate level, with many students being able to identify problematic information but not consistently creating a full critical sequence. This distinction is important for higher education. It calls into question deficit narratives about students' lack of criticality and the notion that digital familiarity translates automatically into high levels of evaluative literacy. These findings align with research that conceptualises digital literacy as an educational practice of judgement and interpretation, rather than access to technologies (Goodfellow 2011; Ng 2012; Smith & Storrs 2023). The most clear-cut indication of unevenness is the dimension hierarchy. There was a high level of CE and a relatively low-level of BA. Students were more likely to say that a claim was questionable than to explain how framing or sensationalism/partial presentation influenced their interpretation. This finding refines the notion of CDL, demonstrating its internal complexity. The first threshold of criticality is awareness that information may be unreliable. At higher levels of criticality, verification procedures are needed, as well as reasoned justification and recognition of the construction of texts to persuade (Hahnel et al. 2020; Molerov et al. 2020; Pangrazio 2016). This internal stratification is pedagogically significant. Teaching students to simply 'check sources' can help them gain a basic level of awareness of credibility, but not necessarily help them to develop an awareness of bias and argumentative reasoning. The comparatively low BA results indicate that students require guidance in identifying one-sidedness, emotional steering, clickbait, selective evidence and distrust rhetoric. Such instruction should not be delivered only as abstract advice. It should be based on real texts that students struggle with, and students should be asked to write justifications, compare texts and explain the importance of certain text features. The DRT results support this interpretation. Questioning was the most prevalent stance, followed by Rejecting and Accepting, which were present in almost a quarter of the responses. This implies that not all students were gullible. However, questioning alone is not

sufficient. Students can state doubts without providing evidence, verification pathways or any indication of bias. One of the analytic contributions of the study is the difference between stance and depth of evaluation. This illustrates the importance of CDL assessment not only in considering whether a student agrees or disagrees with a claim, but also in exploring the reasoning behind the stance. Stimulus level variation also demonstrates that CDL is context sensitive. Students were more critical of unsupported health claims than of media-distrust narratives. A health claim that is not backed up by science might be somewhat familiar as an issue, as students can refer to familiar concepts such as scientific evidence, official health sources and testing. By contrast, a media-distrust narrative might seem critical and might also be accepted uncritically when it comes to other claims. Therefore, it is important for students to assess institutional credibility and the evidence presented in the distrust narrative itself. This requires a higher level of scepticism, verification and awareness of bias. These descriptive patterns are relevant to contemporary environments that may involve misinformation. Distrust narratives are not only false content but can also undermine trust in common knowledge systems. A student who rejects mainstream media wholesale while accepting alternative claims without equivalent scrutiny may appear sceptical but remain vulnerable to manipulation. The results indicate that CDL pedagogy needs to differentiate between productive scepticism and cynicism. Productive scepticism requires evidence, comparison of sources and logical consistency. Cynicism can lead to the rejection of well-supported sources and the acceptance of less well-supported ones (Lewandowsky et al. 2012; Pennycook & Rand 2021). A careful interpretation of the programme-level variation is also required. English Education and Pharmacy had higher percentages of high responses, while Diploma in Nursing and Bachelor of Health Administration had lower percentages. The differences reported in this study are descriptive patterns of responses at the item level and should not be interpreted as evidence of the causal effects of the discipline. However, these patterns suggest that early disciplinary orientations towards text, evidence and source authority may be reflected in students' evaluative repertoires. This potential reinforces the need for digital literacy instruction across the university and in a discipline-specific manner. The evidence in the corpus supports the argument that criticality is manifested in language. Evaluative marker density increased across CDLS levels. High-level responses used words more frequently in connected reasoning that were related to source, evidence, verification, officiality and bias. Low-level responses, on the other hand, were more likely to be lexically thin and based on popularity or plausibility. This finding highlights the importance of corpus-assisted analysis in literacy research, as it shows how evaluative practices manifest in students' discourse (Gillings & Mautner 2024). However, the corpus layer must be considered with caution. The marker families are not completely independent of the conceptual dimensions of the rubric, so frequency evidence should not be treated as independent validation of the coding scheme. It is interpretive, that is, it demonstrates the linguistic patterning of the coded differences in the data. This coding- and concordance-oriented analysis is thus more powerful than either approach alone. Coding revealed the presence of key evaluative dimensions and corpus analysis revealed how these dimensions were textured in language. This research contributes to the field of literacy studies by moving away from a focus on digital skills towards written evaluative performance. Critical digital literacy can be seen in the students' responses in writing credibility judgements, verification statements, reasons and bias-related observations. This is particularly relevant to literacy research because the study treats digital criticality as a literacy practice enacted through discourse. It also provides insights into the engagement of Global South students, specifically Indonesian students, in problematising digital information at the onset of their university studies. The findings support an early undergraduate curriculum that teaches four related capacities explicitly: (1) assessing credibility, (2) checking information, (3) reasoning from evidence and (4) identifying bias. These capacities should be embedded in writing, reading and disciplinary literacy activities. Students should be asked not only to determine whether information is reliable but also to explain their reasoning, identify relevant sources, compare evidence and recognise rhetorical framing. This would shift digital literacy learning from the awareness level to the procedural and discursive levels.

There are several limitations to consider when interpreting the results. The sample was easy to access in the classroom and was limited to only two institutions; therefore, the results obtained cannot be generalised statistically to all students in Indonesia. The data only included evaluative performance in the task conditions, not natural browsing behaviour. The sequence of stimuli was not counterbalanced. Response-level patterns were not completely independent, as each participant provided three responses. Descriptive, rather than inferential, comparisons were made between the programmes and the stimuli. Lastly, the corpus marker analysis provides a helpful aid to interpretation, but cannot replace close reading and coding. Future research should involve larger multi-institutional samples, counterbalanced stimulus sequences, longitudinal designs and tasks that involve written responses and observed verification behaviour.

Conclusion

This study aimed to analyse the CDL of first-year students in Indonesia using a corpus-assisted approach, examining 210 written responses to PDI. The results indicate that students' CDL was mostly moderate, with CE more frequent than IV, CR and BA. The most significant finding was that CDL is multilevel and situational. Students were able to identify questionable credibility better than bias or framing effects. They were more critical of unsupported health claims than of media-distrust narratives, suggesting that the evaluative resources needed to deal with the two types of problematic information differed. The variation at the programme level also indicated that

disciplinary literacy contexts may be reflected in how students compose critical judgements, but this needs further exploration. These results should be interpreted in light of evidence of written evaluative performance in controlled task conditions and not as direct evidence of real-time online verification behaviour. Within that boundary, this study offers a compact model for exploring the manifestation of digital criticality in undergraduate writing and a practical basis for developing early university literacy instruction focused on verification, reasoning and BA.

Acknowledgements

The authors would like to express sincere gratitude to the students and lecturers of Universitas Kristen Indonesia Tomohon and Sekolah Tinggi Ilmu Kesehatan Bethesda Tomohon for their participation and support during the implementation of this study.

During the preparation of this work, the authors used Grammarly for grammar checking, spelling consistency, punctuation review, language polishing and clarity refinement. The content was reviewed and edited by the authors, who take full responsibility for its accuracy.

Competing interests

The authors, James E. Lalira and Maya P. Warouw, declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

CRediT authorship contribution

James E. Lalira: Conceptualisation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. Maya P. Warouw: Conceptualisation, Formal analysis, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. All authors reviewed the article, contributed to the discussion of results, approved the final version for submission and publication, and take responsibility for the integrity of its findings.

Funding information

The authors received no financial support for the research, authorship, and/or publication of this article.

Data availability

The data that support the findings of this study are not openly available due to confidentiality. De-identified excerpts and analytic materials are available from the corresponding author, James E. Lalira, upon reasonable request.

Disclaimer

The views and opinions expressed in this article are those of the authors and are the product of professional research. They do not necessarily reflect the official policy or position of any affiliated institution, funder, agency, or that of the publisher. The authors are solely responsible for the results, findings, and content of this article.

References

Allcott, H. & Gentzkow, M., 2017, 'Social media and fake news in the 2016 election', *Journal of Economic Perspectives* 31(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Berger, L.M., Kerkhof, A., Mindl, F. & Münster, J., 2025, 'Debunking "fake news" on social media: Immediate and short-term effects of fact-checking and media literacy interventions', *Journal of Public Economics* 245, 105345. https://doi.org/10.1016/j.jpubeco.2025.105345

Breakstone, J., Smith, M., Wineburg, S., Rapaport, A., Carle, J., Garland, M. et al., 2021, 'Students' civic online reasoning: A national portrait', *Educational Researcher* 50(8), 505–515. https://doi.org/10.3102/0013189X211017495

Coiro, J., 2011, 'Predicting reading comprehension on the Internet: Contributions of offline reading skills, online reading skills, and prior knowledge', *Journal of Literacy Research* 43(4), 352–392. https://doi.org/10.1177/1086296X11421979

Farias-Gaytan, S., Aguaded, I. & Ramírez-Montoya, M.-S., 2023, 'Digital transformation and digital literacy in the context of complexity within higher education institutions: A systematic literature review', *Humanities and Social Sciences Communications* 10, 386. https://doi.org/10.1057/s41599-023-01875-9

Ferrucci, P. & Hopp, T., 2023, 'Let's intervene: How platforms can combine media literacy and self-efficacy to fight fake news', *Social Media + Society* 8(4), 367–389. https://doi.org/10.1177/20570473231203081

Gillings, M. & Mautner, G., 2024, 'Concordancing for CADS: Practical challenges and theoretical implications', *International Journal of Corpus Linguistics* 29(1), 34–58. https://doi.org/10.1075/ijcl.21168.gil

Goodfellow, R., 2011, 'Literacy, literacies and the digital in higher education', *Teaching in Higher Education* 16(1), 131–144. https://doi.org/10.1080/13562517.2011.544125

Gutiérrez-Ujaque, D., 2024, 'Towards a critical digital literacy and consciousness in higher education: The emancipatory role of critical digital pedagogy', *Pedagogies: An International Journal* 19(3), 337–371. https://doi.org/10.1080/1554480X.2024.2379787

Hahnel, C., Eichmann, B. & Goldhammer, F., 2020, 'Evaluation of online information in university students: Development and scaling of the screening instrument EVON', *Frontiers in Psychology* 11, 562128. https://doi.org/10.3389/fpsyg.2020.562128

Hardianti, M., 2024, 'Indonesian higher education students' perception on critical literacy', *Englisia: Journal of Language, Education, and Humanities* 11(2), 111–123. https://doi.org/10.22373/ej.v11i2.20003

Huang, G., Jia, W. & Yu, W., 2024, 'Media literacy interventions improve resilience to misinformation: A meta-analytic investigation of overall effect and moderating factors', *Communication Research* 53(7), 899–926. https://doi.org/10.1177/00936502241288103

Iskandar, I., Sumarni, S., Dewanti, R. & Asnur, M.N.A., 2022, 'Infusing digital literacy in authentic academic digital practices of English language teaching at universities', *International Journal of Language Education* 6(1), 75–90. https://doi.org/10.26858/ijole.v6i1.31574

Jones-Jang, S.M., Mortensen, T. & Liu, J., 2021, 'Does media literacy help identification of fake news? Information literacy helps, but other literacies don't', *American Behavioral Scientist* 65(2), 371–388. https://doi.org/10.1177/0002764219869406

Kahne, J. & Bowyer, B., 2017, 'Educating for democracy in a partisan age: Confronting the challenges of motivated reasoning and misinformation', *American Educational Research Journal* 54(1), 3–34. https://doi.org/10.3102/0002831216679817

Kozyreva, A., Lewandowsky, S. & Hertwig, R., 2020, 'Citizens versus the Internet: Confronting digital challenges with cognitive tools', *Psychological Science in the Public Interest* 21(3), 103–156. https://doi.org/10.1177/1529100620946707

Lalira, J.E., Pangemanan, Y.A.T., Scipio, J.E., Lumi, S., Merentek, T.C. & Tumuju, V.N., 2024, 'Evaluating the impact of AI tools on grammar mastery: A comparative study of learning outcomes', *Voices of English Language Education Society* 8(3), 701–713. https://doi.org/10.29408/veles.v8i3.27856

Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N. & Cook, J., 2012, 'Misinformation and its correction: Continued influence and successful debiasing', *Psychological Science in the Public Interest* 13(3), 106–131. https://doi.org/10.1177/1529100612451018

Lu, C., Hu, B., Bao, M.-M., Wang, C., Bi, C. & Ju, X.-D., 2024, 'Can media literacy intervention improve fake news credibility assessment? A meta-analysis', *Cyberpsychology, Behavior, and Social Networking* 27(4), 240–252. https://doi.org/10.1089/cyber.2023.0324

Marttunen, M., Salminen, T. & Utriainen, J., 2021, 'Student evaluations of the credibility and argumentation of online sources', *The Journal of Educational Research* 114(3), 294–305. https://doi.org/10.1080/00220671.2021.1929052

Mbau, A.T. & Sugeng, B., 2019, 'Critical literacy for ELT in Indonesia: What EFL teachers should be aware of', *Journal of English Language Teaching and Linguistics* 4(2), 143–156. https://doi.org/10.21462/jeltl.v4i2.255

McGrew, S., Breakstone, J., Ortega, T., Smith, M. & Wineburg, S., 2018, 'Can students evaluate online sources? Learning from assessments of civic online reasoning', *Theory & Research in Social Education* 46(2), 165–193. https://doi.org/10.1080/00933104.2017.1416320

Metzger, M.J., Flanagin, A.J. & Zwarun, L., 2003, 'College student web use, perceptions of information credibility, and verification behavior', *Computers & Education* 41(3), 271–290. https://doi.org/10.1016/S0360-1315(03)00049-6

Molerov, D., Zlatkin-Troitschanskaia, O., Nagel, M.-T., Brückner, S., Schmidt, S. & Shavelson, R.J., 2020, 'Assessing university students' critical online reasoning ability: A conceptual and assessment framework with preliminary evidence', *Frontiers in Education* 5, 577843. https://doi.org/10.3389/feduc.2020.577843

Ng, W., 2012, 'Can we teach digital natives digital literacy?' *Computers & Education* 59(3), 1065–1078. https://doi.org/10.1016/j.compedu.2012.04.016

O'Brien, B.C., Harris, I.B., Beckman, T.J., Reed, D.A. & Cook, D.A., 2014, 'Standards for reporting qualitative research: A synthesis of recommendations', *Academic Medicine* 89(9), 1245–1251. https://doi.org/10.1097/ACM.0000000000000388

Pangrazio, L., 2016, 'Reconceptualising critical digital literacy', *Discourse: Studies in the Cultural Politics of Education* 37(2), 163–174. https://doi.org/10.1080/01596306.2014.942836

Pennycook, G. & Rand, D.G., 2021, 'The psychology of fake news', *Trends in Cognitive Sciences* 25(5), 388–402. https://doi.org/10.1016/j.tics.2021.02.007

Polizzi, G., 2020, 'Digital literacy and the national curriculum for England: Learning from how the experts engage with and evaluate online content', *Computers & Education* 152, 103859. https://doi.org/10.1016/j.compedu.2020.103859

Sillat, L.H., Tammets, K. & Laanpere, M., 2021, 'Digital competence assessment methods in higher education: A systematic literature review', *Education Sciences* 11(8), 402. https://doi.org/10.3390/educsci11080402

Silvhiany, S., Huzaifah, S. & Ismet, I., 2021, 'Critical digital literacy: EFL students' ability to evaluate online sources', *Indonesian Journal of EFL and Linguistics* 6(1), 249–269. https://doi.org/10.21462/ijefl.v6i1.364

Smith, E.E. & Storrs, H., 2023, 'Digital literacies, social media, and undergraduate learning: What do students think they need to know?' *International Journal of Educational Technology in Higher Education* 20, 29. https://doi.org/10.1186/s41239-023-00398-2

Tinmaz, H., Lee, Y.-T., Fanea-Ivanovici, M. & Baber, H., 2022, 'A systematic review on digital literacy', *Smart Learning Environments* 9, 21. https://doi.org/10.1186/s40561-022-00204-y

Von Elm, E., Altman, D.G., Egger, M., Pocock, S.J., Gøtzsche, P.C. & Vandenbroucke, J.P., 2007, 'The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies', *PLoS Medicine* 4(10), e296. https://doi.org/10.1371/journal.pmed.0040296