Development and Validation of the Atma AI Ethics in Academic Writing Scale (AAEWS)

Angelica Levina Setyawan [1], Vanessa Layono [2], Maria Theresa Rizano [3], Theodoure Noah Kristan Sompie [4], Regina Sugito [5], Marcha Joan Fibi Lumban Tobing [6], Christiany Suwartono [7] *

[1]Fakultas Psikologi, Universitas Katolik Indonesia Atma Jaya, Jakarta, Indonesia
[2]Center for the Study of Sustainable Community, Tangerang, Indonesia
*christiany.suwartono@atmajaya.ac.id

ABSTRACT

The integration of artificial intelligence (AI) in education presents new opportunities and ethical challenges, especially concerning academic integrity in student papers. Misuse of AI can result in issues such as plagiarism, data privacy violations, and theoretical reference hallucinations. To address this, a standardized measurement tool is needed to evaluate students' ethical perceptions on using AI for academic writing. This study aims to develop and test the psychometric properties of a measurement tool to assess students' ethical perceptions of AI in writing scientific papers. Ethical Perception in this test tool is defined as awareness and commitment to the responsible use of AI in research and scientific writing, which includes student's responsibility for the accuracy and integrity of results, substantial author contributions, and transparency and recognition of AI use with five dimensions: reporting standards, data access and retention, originality and acknowledgment of sources, citations, then the use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork. With convenience sampling, the sample size in this study for the item analysis was 34 participants and for validity and reliability testing was 120 participants. The initial number of items was developed from 60 to 42. This measurement tool has an omega reliability of 0.918, and its five dimensions have alpha reliability of 0.335-0.742. For validation, a comparative measurement tool from Meiriza et al.'s (2024) research served as external validity, yielding $r(118) = 0.288$, $p < 0.001$. In conclusion, the results of this psychometric testing indicate that the AAEWS is valid according to external validity, correlation with the other test that had similar construct and reliable. Therefore, based on external validity and reliability, the AAEWS can confidently measure students' ethical perceptions of the use of AI in scientific writing.

Keywords: *AI ethics, perception of AI ethics, scientific writing, student assignments, written work.*

Introduction

The fast development of the digitalization era has given birth to many facets of improvements in technology, one of them being Artificial Intelligence (AI). AI has become one of the most prominent innovations because it has given a huge impact in many aspects, including the education sector, especially in higher education (Oktavianus et al., 2023). Education is a deliberate and systematic process that aims to create learning environments and experiences that enable students to actively develop their potential (Rahman, 2022). In the context of education, AI offers various learning opportunities to support personalized learning, which refers to a learning system that can be automatically adjusted in accordance with students' needs and abilities (Oktavianus et al., 2023). On the other hand, the use of AI in the education sector also creates complex ethical challenges. Some emerging ethical issues include violation of data privacy, lack of transparency in AI usage, and a decrease in students' creativity (Baskara, 2023, as cited in Dabis & Csáki, 2024). Therefore, higher education institutions are required to build a strong ethical framework for AI usage that upholds fairness, accountability, and respect towards academic work (Dabis & Csáki, 2024).

AI has two important principles in academic use, which are human agency and human authorship. These principles emphasize the importance of human involvement and responsibilities within the academic process even when being assisted by AI, and that the use of AI must always be based on conscious human decisions that prioritize ethical values. The use of ethics in AI also must pay attention to principles such as human autonomy (human-centeredness), prevention from harm (safety and security), fairness (free from bias and discrimination), and openness (transparency, traceability, auditability) (Papagiannidis et al., 2025).

One of the most prominent ethical issues arising from the use of AI is the possibility of plagiarism and academic misconduct. Some AI users fail to provide clear attribution towards AI-generated text, obscuring who the real author is. This makes it difficult to assess students' ability in formulating original ideas, which can end up potentially jeopardising academic integrity (Subaveerapandiyan et al., 2025). Another concern is overreliance on AI, which causes students to be less inclined in analytical and critical thinking (Yeo, 2023, as cited in Hutson, 2024).

Various research shows that dependence on AI can negatively impact students' cognitive function. Research done by Kosmuna et al. (2025) states that the excessive dependence on generative AI (GenAI) when working on their assignments causes the loss of cognitive ability. Another research by Ainaya and Saragih (2025) strengthens that the continuous use of AI reduces student independence in exploring ideas. This was also conveyed by Ulfah (2024), which states that excessive dependence on AI leads to decreasing analytical and creative abilities that are essential for academic success.

A 2024 survey by the Ministry of Communication and Information shows that 87% of Indonesian students use AI for assignments. Chegg and Goodstats states that Indonesia is a country with the highest usage of AI among university students, compared to countries such as Malaysia, Spain, and Saudi Arabia (Yonathan, 2025). Tirto (Hartanto & Rohmah, 2024) found 86.21% of students use AI, with 56% being university students, mainly for summarizing articles, creating journals, writing essays, and drafting scientific paper.This high usage raises academic integrity risk, especially as Meho (2025) identified high integrity risks in 13 leading Indonesian universities out of 1,500 institutions surveyed.This condition highlights the urgent need for a valid measuring instrument to understand students' ethical perceptions as a basis for developing appropriate institutional guidelines and policies. This need not only arises from the issues of plagiarism but also on the

importance of privacy, ethical responsibility, as well as potential dependency of technology from university students (Fadhillah & Lestari, 2025; Judijanto & Harsya, 2025).

This instrument measures university students' ethical perceptions of AI use in academics, focusing on awareness and commitment to responsible AI use, accuracy, integrity, contribution, transparency, and acknoledgement (Mann et al., 2024; Elsevier, 2025). It is comprised of five dimensions: reporting standards (accurate reporting), data access and retention (data availability for review), originality and acknowledgement of sources (ensures originality), citations (use of relevant references), and use of generative AI and AI-assissted technologies in scientific writing and in figures, images, and artwork (transparent use of AI). The measurements of these dimensions within the ethical perception towards AI can be utilized to develop institutional policies in education regarding the use of AI in academic writing, to encourage academic integrity and the ethical use of AI.

Methods

The participants involved were 3rd to 7th semester undergraduate students at X University with ages ranging between 18-22 years old. Data was collected twice with convinience sampling technique, labeled as tryout data and field data. The tryout data collection, the participants' ages ranged from 18 to 21 years (M = 7.25, SD = 8.96) while the field data collection, the participants ranged from 18 to 22 years old with two participants aged over 22 years (M = 20, SD = 18.12).

Table 1. Participants' Demographic Data

Demographic	**Tryout** (n = 34)		**Field** (n = 120)	
	f	**%**	**f**	**%**
Using GenAI				
Yes	29	85.3%	120	96.8%
No	5	14.7%	4	3.2%
Gender				
Female	22	75.9%	82	68.3%
Male	7	24.1%	38	31.7%
Study Program				
Psychology	26	89.6%	75	62.5%
Management	2	6.8%	6	5%
Communication	1	3.4%	9	7.4%
Accounting			16	13.3%
Law			7	5.8%
Medical Science			2	1.7%
Biotechnology			3	2.5%
Industrial Engineering			1	0.8%
English Language Studies			1	0.8%

Instrument

The Atma AI Ethics in Academic Writing Scale (AAEWS) initially consisted of 60 items, and was condensed to 42 after conceptual screening and preliminary psychometric testing. These items were then grouped into five main dimensions (Mann et al., 2024; Elsevier, 2024): (a) Reporting Standards, (b) Data Access and Retention, (c) Originality and Acknowledgement of Sources, (d) Citations, and (e) Use of Generative AI and AI-Assisted Technologies in Scientific Writing as well as in Figures, Images, and Artwork. Each item is presented as a statement, and respondents are asked to choose the option that best describes themselves. The response options use a 6-point Likert scale, where 1 means "never," 2 means "rarely," 3 means "sometimes," 4 means "often," 5 means "almost always," and 6 means "always." All items are favorable items. The higher the score obtained, the higher an individual's awareness of ethical AI use in producing academic writing. The scale will be administered online using the Google Forms platform. The scoring of the instrument uses a cumulative scoring method to measure the level of perception regarding ethical AI use and its five dimensions.

External Comparative Measuring Instrument

The external test used for validity (AUP) used in this study is a scale designed to measure the use of AI on students' learning interest developed by Meiriza et al. (2024). The Cronbach's alpha reliability results were 0.803 for the AI usage variable and 0.808 for the learning interest variable. Researchers selected the learning interest scale as the AUP because it measures the same construct. The AUP was included in the same Google Forms questionnaire alongside the AAEWS instrument. The AUP consists of 15 items with a Likert-scale response format ranging from 1 to 4, where (1) Strongly disagree, (2) Disagree, (3) Agree, and (4) Strongly agree.

Procedure

Developing a measurement instrument begins with identifying a phenomenon to be used. After identifying the phenomenon, the next step is to develop a theoretical framework and establish a background. One of the key elements in developing a theoretical framework is determining the constructs and dimensions to be measured, along with their indicators. Researchers used dimensions from Elsevier to measure students' ethical awareness and attitudes towards the use of AI. There are five dimensions: (a) Reporting standards, (b) Data access and retention, (c) Originality and Acknowledgement of sources, (d) Citations, (e) Use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork. After that researchers determined the characteristics of the participants to be studied, namely students in semesters 3-7 at University X who had used GenAI.

Consideration was given to the items and participants' characteristics to be targeted. Researchers began designing the tryout by creating 60 items based on the selected dimensions. These items then underwent a process of refinement and revision as well as expert judgement. All 60 items were entered into Google Forms and given to the participants, generating a "tryout dataset" over a week that contained 34 "clean" responses. Using that information, researchers proceed to the next step, which was the creation of the final measuring instrument. Using tryout data, researchers performed items

and dimensions reliability testing with Corrected Item-Total Correlation (CITC). After a repeated data cleaning process, items that were classified as appropriate were re-entered for use for the final data distribution.

After the CITC process was complete, researchers randomized the order of the items and entered them into Google Forms along with an external test (AUP). Finally, researchers distributed it to participants who met the criteria. The target is 120 clean responses, followed by validity and reliability tests. After all the data was collected, the researchers conducted normality, validity, and reliability tests again to determine the results of the measurement tool. These tests include the Shapiro-Wilk normality test, validity was conducted on each dimension using Cronbach's alpha, and validity for all the items using McDonald's omega. Last, researchers wrote and reported the overall results of the study.

Data Analysis Technique

Researchers used measures of central tendency to provide an overview of the distribution of scores per dimension and construct of the test instrument. First, they performed item analysis using item-rest correlation. From the tryout data, the researcher obtained the 42 best items out of the 60 items developed. Researchers used the Shapiro-Wilk normality test to determine if the data was normally distributed or not. After determining that the data is normally distributed, the researcher tested its validity using Pearson correlation. Additionally, the researchers also tested the reliability of each dimension using Cronbach's alpha technique. The total reliability calculation was performed using McDonald's omega technique.

The item analysis was conducted in two ways: qualitative item analysis through expert judgment and quantitative item analysis using Corrected Item-Total Correlation (CITC). Expert judgment was performed by one measurement expert to assess the clarity of language, relevance, and simplicity of the items. The researcher then conducted the item analysis process using the Corrected Item-Total Correlation (CITC) technique, where each dimension was tested and if an item did not meet the CITC criteria ≤ 0.30 (Nunnally & Bernstein, 1994) were eliminated. Items in each dimension that did not meet the criteria were eliminated, starting with those that had the lowest CITC values.

After that, the second round of data collection was conducted. For the external validity testing, the researcher used the Shapiro–Wilk normality test and Pearson's correlation validity test after collecting the data using the JASP application.

To determine whether the measurement tool data distribution is normal or not, the Shapiro-Wilk normality test is performed. The results of the Shapiro-Wilk normality test will be categorized as significant and normally distributed if the p-value is greater than 0.05 (Field, 2009). If the data is normally distributed, then a validity test is performed using Pearson's correlation. The correlation will be significant if the p-value is below 0.05 and the correlation is positive (Cohen & Swerdlik, 2018). Once the target number of respondents has been reached, the researcher performs another normality test using the Shapiro-Wilk test. Then, a validity test is conducted between the AAEWS test and the external test (AUP) using Pearson's correlation method. Reliability testing for each dimension was performed using Cronbach's alpha, while the reliability of the total score across all dimensions was assessed using McDonald's omega. A minimum value of 0.60 was used as the criterion indicating acceptable reliability.

Results

Item Analysis

Based on the analysis, the AAEWS test, which started with 60 items, was eventually narrowed down to 42 items. A total of 18 items were eliminated due to low Corrected Item-Total Correlation (CITC) results (≤ 0.30). The researchers carried out the item analysis process using the CITC technique in which every item in each dimension was tested and if an item did not meet the CITC criteria of ≤ 0.30, the item would be discarded. After running through the CITC process several times, the researchers discarded 18 items; nine from the reporting standards dimension, two from the data access and retention dimension, one from the originality dimension, none from the citations dimension, and six from the use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork dimension. When reviewing the participant responses, the 18 items that were discarded shared a similarity in which the data that was too homogeneous and thus made the items repetitive. The final AAEWS test consists of 42 items that measure the scale of the ethical use of AI in the creation of student written work. Eventhough, in the field data, there still several items that below the CITC criteria of ≤ 0.30, we consider still using them due to the important content of these items to measure its dimension and the reliability coeficients still sufficient to be consider as a reliable measure (see Table 2.).

Reliability

To test the reliability, a sample of 120 participants was obtained and used. The results of the reliability analysis using the Cronbach's alpha technique produced alpha coefficients for all five dimensions; the reporting standards dimension with a value of 0.335, the data access and retention dimension with 0.705, the originality and acknowledgment of sources dimension with 0.742, the citations dimension with 0.735, and the use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork dimension with 0.696. The researchers also tested the holistic reliability by using McDonald's omega on the total of all items which resulted in a value of 0.918. Table 2 presents the reliability results per dimension as well as the total of all items.

Table 2. Reliability per dimension and total of all items

Dimension	Initial Number of Items	Range Item-rest Correlation (try out)	Final Number of Items	Range Item-rest Correlation (field)	Reliability Coefficient (field)	M	SD	SEM
Reporting Standards	12	0.023 – 0.608	3	0.015 – 0.295	0.335	15.98	2.012	1.641
Data Access and Retention	12	-0.087 – 0.666	10	0.034 – 0.590	0.705	43.63	5.977	3.246
Originality and Acknowledgement of Sources	12	0.210 – 0.768	11	-0.143 – 0.716	0.742	48.83	7.770	3.947
Citations	12	0.304 – 0.776	12	0.046 – 0.541	0.735	57.18	7.086	3.648
UAI*	14	0.219 – 0.745	8	0.311 – 0.586	0.696	27.88	4.919	2.712
AAEWS	60	0.120 – 0.710	42	-0.221 – 0.719	0.918	193.5	23.91	6.848

Note. UAI* = use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork

Validity

The Shapiro-Wilk normality test resulted in W(120) = 0.981, p = 0.094 > 0.05. Knowing this, the data can be categorized as a normal distribution due to p > 0.05 (Cohen & Swerdlik, 2018). As a result, the data proceeded to be tested using the Pearson's correlation method which produced r = 0.288. The positive r value result indicates the AAEWS test as being a valid test. In addition, with r(118) = 0.288, p < 0.001, r2 = 0.083, the AAEWS test has the same variance as the AUP with a value of 8.3%. The remaining 91.7% is influenced by other factors not examined in this study. The relatively low coefficient of determination value indicates that the ethical use of AI only explains a small portion of the variation in student writing. However, the AAEWS test measures the same construct as the AUP regarding the use of AI in academic contexts.

Table 3 . Matrix of validity between dimensions and tests

Variable	AUP	AIEAS	DA	RS	O	C	UAI
1. AUP	—						
2. AAEWS	0.288**	—					
3. DA	0.260**	0.864***	—				
4. RS	-0.039	0.508***	0.458***	—			
5. O	0.349***	0.899***	0.699***	0.358***	—		
6. C	0.260**	0.904***	0.708***	0.410***	0.724**	—	
7. UAI	0.173	0.880***	0.675***	0.344***	0.752**	0.783**	—

Notes*:* AUP = external validity test; AAEWS = Atma AI Ethics in Academic Writing Scale; DA = data access and retention; RS = reporting standards; O = originality; C = citations; UAI = use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork

Discussions

According to Griethuijsen et al. (2014), a good reliability value is at least 0.60. As seen in Table 2, the reliability result using Cronbach's alpha shows that four of the five dimensions have high reliability (α = 0.696-0.742) because the alpha coefficient is more than 0.60. Only the reporting standards dimension has a relatively low alpha coefficient (α = 0.335) since it is below the limit according to Griethuijsen et al. (2014), rendering it unreliable. As for the reliability of the construct, it resulted in an omega coefficient of 0.918, indicating that the AAEWS test is reliable.

The reporting standards dimension is unreliable due to several factors. According to Kaplan and Saccuzzo (2018), the reliability of a test increases as the number of items increases. This dimension only has a total of three items, in contrast to other dimensions which have at least five items. During the item analysis process, the highest CITC result for the reporting standards dimension was 0.295. This value does not meet the CITC

criteria (≤ 0.30) based on Nunnally and Bernstein (1994). However, beyond psychometric considerations, the low CITC that lead to low reliability of the Reporting Standards dimension also have theoretical implications for understanding AI ethical perceptions in academic writing. We have impression that the students may not experience reporting practices as an area of ethical dilemma, but rather as a universally accepted academic rule. The three items in this dimension are: "I make sure when using GenAI to double-check the spelling included in my report", "I avoid using references whose accuracy is questionable from GenAI results", and "I make sure to correct GenAI results that are found to be inaccurate." These items do not conceptually raise 'pro-contra' issues. The phrases contained in the items, for example, "I make sure…" and "I avoid…" are too normative.

Social desirability bias also has the potential to cause low reliability in the reporting standards dimension. Social desirability is an individual's tendency to provide answers that appear positive and align with social expectations (Paulhus, 2002, in Oktapialdi, 2018). Items in the reporting standards dimension are normative, hence respondents choose answers that are considered in line with relevant social norms. This has the potential to cause responses to become homogenous. This condition not only reduces reliability but can also affect construct validity, as the responses given do not fully reflect the respondent's actual behaviour or perceptions.

For further research, an internal evaluation of the measuring instrument could be conducted, including improving the number of items, item construction, and avoiding overly normative items on dimensions with low reliability. This in-depth study is expected to improve the accuracy of measuring the construct of AI ethical perceptions. Our findings contribute to the conceptualization of AI ethics in education by indicating that not all ethical dimensions operate at the same psychological level. Some dimensions such as originality, data access, or citation practices may involve evaluative judgment and ethical reasoning. However, reporting standards may reflect internalized academic norms. Therefore, to conceptualize AI ethical perception, we need both attitudinal components (evaluations) and normative compliance components (rule-following tendencies).The AAEWS test in this study was compared with the AEPS (Artificial Intelligence Ethics Perception Scale) developed by Saatci (2025) as a reference for evaluating the quality of the measuring instrument. Conceptually, AAEWS focuses on students' ethical perceptions regarding the use of AI in scientific writing, encompassing five dimensions measured in AAEWS, which are Reporting Standards, Data Access and Retention, Originality and Acknowledgement, Citations, and Use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork. In contrast, AEPS measures perceptions of AI ethics in a much broader context, involving macro ethical dimensions, such as transparency, accountability, privacy, fairness, and human oversight. Knowing this, AEPS can be used for discriminant validity because the use of AAEWS focuses specifically on the education sector.

The AEPS study demonstrated stronger psychometric standards. Saatci (2025) used a large, cross-national sample and comprehensively applied EFA and CFA, resulting in a stable factor structure and high reliability for each dimension (α = 0.85-0.94). Meanwhile,

the AAEWS is in the early stages of measuring instrument development, using CITC analysis. This indicates that the AAEWS still requires strengthening its structural analysis to prove its internal validity, similar to the AEPS. According to Marengo et al. (2025), the validity techniques that can be used are factor analysis. The factor analyses used are Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). Marengo et al. (2025) explain that EFA is used to explore the factor structure underlying statement items and identify whether items are grouped consistently according to the theoretical construct to be measured. EFA is needed to explore and identify which items form a dimension. Then, CFA was conducted to validate the factor structure obtained from EFA (Marengo et al., 2025). Both analyses served not only to strengthen construct validity but also to refine items and improve the accuracy of construct representation in the measurement instrument.

Conclusion

The Atma AI Ethics in Academic Writing Scale (AAEWS) is a measuring tool that aims to measure students' ethical perceptions in writing. With the data distribution carried out with $n = 120$, the Shapiro-Wilk normality test showed $p = 0.094 > 0.05$, meaning that the data can be categorised as a normal distribution. Knowing this, the validity test used the Pearson correlation technique with results showing $r(118) = 0.288$, $p < 0.001$, $r^2 = 0.083$, indicating that the AAEWS is valid. In addition, the reliability test shows that the AAEWS is reliable due to the consistency in the results of data analysis and the omega coefficient of 0.918. Thus, the Atma AI Ethics in Academic Writing Scale (AAEWS) can be declared as valid and reliable and is suitable for use to measure students' ethical perceptions in writing, developing responsible AI literacy curricula, and evaluating educational interventions related to the ethics of using AI in the academic world.

References

Ainaya, J. R., & Saragih, L. S. (2025). Pengaruh penggunaan Artificial Intelligence (AI) dan kemandirian belajar terhadap kreativitas mahasiswa program studi pendidikan bisnis stambuk 2022 fakultas ekonomi Universitas Negeri Medan. *Khatulistiwa: Jurnal Pendidikan Dan Sosial Humaniora*, *5*(3), 767–776. https://doi.org/10.55606/khatulistiwa.v5i3.7226

Cohen, R. J., & Swerdlik, M. E. (2018). *Psychological testing and assessment: An introduction to tests and measurement* (9th ed.). New York, Ny: Mcgraw-Hill Education.

Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-03526-z

Elsevier. (2025). Publishing ethics | Elsevier policy. www.elsevier.com. https://www.elsevier.com/about/policies-and-standards/publishing-ethics#4-duties-of-authors

Fadhillah, R., & Lestari, B. D. (2025). Penggunaan AI pada mahasiswa psikologi dalam meningkatkan kesehatan mental. *Jurnal Empati, 13*(4), 280–290.

Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London: Sage.

Hartanto, A. Y., & Rohmah, F. N. (2024, May 31). *Makin marak siswa pakai AI untuk mengerjakan tugas*. Retrieved from https://tirto.id/penggunaan-ai-di-dunia-pendidikan-makin-marak-dan-merata-gZax

Hutson, J. (2024). Rethinking plagiarism in the era of generative AI. *Journal of Intelligent Communication*, *4*(1), 20–31. https://doi.org/10.54963/jic.v4i1.220

Kaplan, R. M., & Saccuzzo, D. P. (2018). *Psychological testing: Principles, applications and issues* (9th ed.). Cengage Learning.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. *ResearchGate*. https://doi.org/10.48550/arXiv.2506.08872

Mann, P., Vazirani, A. A., Aboy, M., Earp, B. D., Minssen, T., Glenn, C. I., & Savulescu, J. (2024). Guidelines for ethical use and acknowledgement of large language models in academic writing. *Nature Machine Intelligence*, *6*, 1–3. https://doi.org/10.1038/s42256-024-00922-7

Marengo, A., Karaoglan-Yilmaz, F. G., Yılmaz, R., & Ceylan, M. (2025). Development and validation of generative artificial intelligence attitude scale for students. *Frontiers in Computer Science*, *7*. https://doi.org/10.3389/fcomp.2025.1528455

Meiriza, M. S., Sembiring, G. B., Sitorus, M., Wardana, V., & Sakinah, N. (2024). Pengaruh penggunaan AI terhadap minat belajar di kalangan mahasiswa: Studi kasus pada generasi Z. *AR-RUMMAN: Journal of Education and Learning Evaluation*, *1*(2), 319–327. https://doi.org/10.57235/arrumman.v1i2.4019

Meho, L. I. (2025). *Gaming the metrics? Bibliometric anomalies and the integrity crisis in global university rankings*. https://doi.org/10.1101/2025.05.09.653229

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). Mcgraw-Hill.

Oktapialdi, R., Medianta Tarigan, & M. Ariez Musthofa. (2018). Pengembangan skala social desirability. *JURNAL PSIKOLOGI INSIGHT*, *2*(1), 33–42. https://doi.org/10.17509/insight.v2i1.11921

Oktavianus, A. J. E., Naibaho, L., & Rantung, D. A. (2023). Pemanfaatan artificial intelligence pada pembelajaran dan asesmen di era digitalisasi | JURNAL KRIDATAMA SAINS DAN TEKNOLOGI. *Jurnal.umnu.ac.id*, *5(2)*, 473–486.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, *34*(2). https://doi.org/10.1016/j.jsis.2024.101885

Rahman, A., Munandar, S. A., Fitriani, A., Karlina, Y., & Yumriani. (2022). Pengertian pendidikan, ilmu pendidikan dan unsur-unsur pendidikan. *Al Urwatul Wutsqa: Kajian Pendidikan Islam*, *2*(1).

Recker, J., Chatterjee, S., Sundermeier, J., & Tarafdar, M. (2025). Digital responsibility: Current perspectives and future directions. *Journal of the Association for Information Systems*, *26*(5), 1222–1238. https://doi.org/10.17705/1jais.00966

Saatci, E. Y. (2025). AI and ethics: Scale development for measuring ethical perceptions of artificial intelligence across sectors and countries. *International Journal of Economic Behavior and Organization, 13*(1), 35–50. https://doi.org/10.11648/j.ijebo.20251301.14

Subaveerapandiyan, A., Kalbande, D., & Ahmad, N. (2025). Perceptions of effectiveness and ethical use of AI tools in academic writing: A study among phd scholars in India. *Information Development*, *41*(3), 728–746. https://doi.org/10.1177/02666669251314840

Ulfah, M. (2024). Dampak ketergantungan pada artificial intelligence terhadap kemampuan analitis dan kreatif mahasiswa. *VOX EDUKASI: Jurnal Ilmiah Ilmu Pendidikan*, *15*(1), 120–130.

Yonatan, A. Z. (2025, October 2). *Benarkah mahasiswa RI ketagihan AI? GoodStats*. Retrieved from https://goodstats.id/article/benarkah-mahasiswa-ri-ketagihan-ai-Y2Mf0

Appendix

Appendix A

Full Atma AI Ethics in Writing Scale (42 Questions)

Please respond to each question using the following scale:

1= Never, 2 = Rarely, 3 = Sometimes, 4 = Often, 5 = Almost Always, 6 = Always

1) I recheck the work when using GenAI to double check the spelling that I enter my reports
2) I am committed to not abusing GenAI
3) I avoid using references of questionable accuracy from GenAI results.
4) I save all the results that GenAI gives me.
5) I confirm that the literature references provided by GenAI come from verified sources.
6) I copy all the writing produced by GenAI.
7) I understand the consequences of violating the rules for using AI in report writing.
8) I understand the regulations set by the university regarding research.
9) If unsure, I will ask further about the policies set by the university regarding the use of GenAI in writing
10) I explain to the interviewees how their data would be used in processing with GenAI.
11) I report the results of data analysis processed by GenAI.
12) I ask GenAI to paraphrase the writing on my report.
13) I follow the university's established restrictions regarding the use of GenAI.
14) I avoid listing GenAI as author or co-author.
15) I try to present more of my own writing rather than quoting everything from GenAI.
16) I make sure that the data I use in writing with GenAI is original data.
17) I avoid unauthorized use of personal information in all writing using GenAI.
18) I feel responsible for ensuring that the sources of quotations obtained from GenAI are included in the bibliography.
19) When creating verbatim using GenAI, I ensure that the source was aware of and agreed to this in the informed consent.
20) When requesting a quote from GenAI, I make sure that the quote is relevant to my research.
21) When I want to process data with GenAI, I make sure that this is stated in the informed consent.
22) I respect the sources' decision to refuse permission for the use of GenAI in processing their data.
23) I ensure that the work written by GenAI does not resemble the work of others.
24) I ensure that the literature references provided by GenAI are relevant to my research.
25) I document every step of data processing done by GenAI.
26) I reread the policies set by the university regarding the use of AI.
27) I avoid excessively citing the literature references obtained from GenAI.
28) I ensure that the literature references provided by GenAI are relevant to the time.
29) When asking GenAI to provide a quote, I ensure that the quote is accurate.
30) I double check the quote GenAI provided by looking up the original source.

31) I only use quotes from GenAI that are truly necessary for my research.
32) I make sure to double-check the results from GenAI to avoid any factual errors.
33) I combine writings from GenAI with my own thoughts.
34) When asking GenAI for ideas, I ask for the source of the idea.
35) I feel responsible for citing the sources I got from GenAI.
36) I ensure that I do not enter personal data of the source person into GenAI.
37) I state the use of GenAI in journal searches.
38) I include GenAI that helped improve the writing in the acknowledgement section of the report.
39) I avoid listing sensitive information to GenAI.
40) If processing data with GenAI, I include it in the research methods section.
41) I censor brand names before processing the data using GenAI
42) I made sure to correct the GenAI results that were found to be inaccurate

Indonesian Translation of Atma AI Ethics in Academic Writing Scale

1 = Tidak Pernah, 2 = Jarang, 3 = Kadang-Kadang, 4 = Sering, 5 = Hampir Selalu, 6 = Selalu

1) Saya memastikan saat menggunakan GenAI untuk memeriksa kembali ejaan yang saya masukkan ke dalam laporan.
2) Saya berkomitmen untuk tidak menyalahgunakan GenAI.
3) Saya menghindari menggunakan referensi yang diragukan keakuratannya dari hasil GenAI.
4) Saya menyimpan semua hasil yang diberikan GenAI
5) Saya memastikan bahwa rujukan pustaka yang diberikan oleh GenAI berasal dari sumber yang terverifikasi.
6) Saya menyalin semua tulisan yang dihasilkan oleh GenAI.
7) Saya memahami konsekuensinya jika melanggar aturan penggunaan AI dalam penulisan laporan.
8) Saya mengerti peraturan yang ditetapkan oleh universitas mengenai penelitian.
9) Jika belum yakin, saya akan bertanya lebih lanjut mengenai kebijakan yang ditetapkan universitas mengenai penggunaan GenAI dalam penulisan.
10) Saya menjelaskan kepada narasumber bagaimana data mereka akan digunakan dalam pengolahan dengan GenAI.
11) Saya melaporkan hasil analisis data yang diproses oleh GenAI.
12) Saya meminta GenAI untuk memparafrase penulisan pada laporan.
13) Saya mengikuti batasan yang ditetapkan universitas terkait penggunaan GenAI.
14) Saya menghindari mencantumkan GenAI sebagai author atau co-author.
15) Saya berusaha menyajikan tulisan sendiri lebih banyak daripada mengutip semua dari GenAI.
16) Saya menjamin bahwa data yang saya gunakan dalam penulisan dengan GenAI adalah data asli.
17) Saya menghindari penggunaan informasi pribadi tanpa izin dalam semua penulisan yang menggunakan GenAI.
18) Saya merasa bertanggung jawab untuk memastikan sumber kutipan yang didapat dari GenAI ke dalam daftar pustaka.
19) Ketika membuat verbatim menggunakan GenAI, saya memastikan dalam informed consent bahwa narasumber mengetahui dan menyetujui hal tersebut.
20) Ketika meminta kutipan dari GenAI, saya harus memastikan bahwa kutipan tersebut relevan dengan penelitian saya.

21) Ketika ingin mengolah data dengan GenAI, saya memastikan hal tersebut tercantum dalam informed consent.
22) Saya mendokumentasikan setiap langkah pengolahan data yang dilakukan GenAI.
23) Saya memastikan bahwa karya yang ditulis GenAI tidak menyerupai karya orang lain.
24) Saya memastikan bahwa rujukan pustaka yang diberikan oleh GenAI relevan dengan penelitian saya.
25) Saya menyatakan penggunaan GenAI dalam pencarian jurnal.
26) Saya menghargai keputusan narasumber jika menolak memberi izin untuk penggunaan GenAI dalam pengolahan data mereka.
27) Saya harus menghindari mengutip rujukan pustaka yang didapat dari GenAI secara berlebihan.
28) Jika mengolah data dengan GenAI, saya mencantumkannya pada bagian metode penelitian
29) Ketika meminta GenAI untuk memberikan kutipan, saya memastikan bahwa kutipan tersebut akurat.
30) Saya mencantumkan GenAI yang membantu meningkatkan penulisan pada bagian acknowledgement pada laporan.
31) Saya hanya menggunakan kutipan dari GenAI yang sungguh-sungguh dibutuhkan dengan penelitian saya.
32) Saya memastikan mengecek ulang hasil dari GenAI untuk menghindari adanya faktual eror.
33) Saya menggabungkan tulisan dari GenAI dengan pemikiran sendiri.
34) Saya mengerti peraturan yang ditetapkan oleh universitas mengenai penelitian.
35) Saya merasa bertanggung jawab untuk mengutip sumber yang saya dapat dari GenAI.
36) Saya memastikan bahwa data pribadi narasumber tidak saya masukkan ke GenAI.
37) Saya membaca kembali kebijakan yang ditetapkan oleh universitas terkait penggunaan AI.
38) Saya harus memeriksa kembali kutipan yang diberikan GenAI dengan mencari sumber aslinya.
39) Saya melaporkan jika menemukan potensi pelanggaran privasi saat menggunakan GenAI dalam pengolahan data.
40) Saya memastikan bahwa rujukan pustaka yang diberikan oleh GenAI bersifat relevan dengan waktu.
41) Jika belum yakin, saya akan bertanya lebih lanjut mengenai kebijakan yang ditetapkan universitas mengenai penggunaan GenAI dalam penulisan.
42) Saya memastikan mengoreksi hasil GenAI yang ditemukan kurang tepat.

Appendix B

Dimensions and Items before Tryout

DIMENSIONS	DEFINITION	CODE	ITEM
Reporting standards	Author of research articles should present an accurate account of the work performed as well as an objective discussion of its significance. Articles should contain sufficient detail and references to permit others to replicate the work. Fraudulent or knowingly inaccurate statements constitute unethical behavior and are unacceptable.	RS_1	I double-checked the results from GenAI before using them in writing the report results.
		RS_2	Even though I use GenAI, I ensure that the report results remain objective.
		RS_3	I made sure to correct the GenAI results that were found to be inaccurate.
		RS_4	I recheck the work when using GenAI to double check the spelling that I enter my reports
		RS_5	I ensure that the reports I write with the help of GenAI still reflect my analytical understanding.
		RS_6	Even though it comes from GenAI, I avoid using inaccurate information.
		RS_7	I ask GenAI to create a bibliography
		RS_8	I adjust the format of the bibliography generated by GenAI to comply with institutional guidelines.
		RS_9	I double-checked the references GenAI made to ensure there were no fictitious references.
		RS_10	I use GenAI only to find additional references.
		RS_11	I make sure not to only copy/paste the references provided by GenAI.
		RS_12	I avoid using references of questionable accuracy from GenAI results.
Data access and retention	Authors may be asked to provide research data supporting their research articles for editorial review or to comply with journals' open data requirements. Authors are expected to provide public access to such data.	DA_1	I understand the regulations set by the university regarding research.
		DA_2	I am committed to not abusing GenAI
		DA_3	I reread the policies set by the university regarding the use of AI.
		DA_4	I follow the university's established restrictions regarding the use of GenAI.
		DA_5	If unsure, I will ask further about the policies set by the university regarding the use of GenAI in writing

		DA_6	I understand the consequences of violating the rules for using AI in report writing.
		DA_7	I report the results of data analysis processed by GenAI.
		DA_8	I make sure that the data I use in writing with GenAI is original data.
		DA_9	I save all the results that GenAI gives me.
		DA_10	I ensure that the data processed with GenAI is compatible to the topic of my report.
		DA_11	I make sure to sort the data results that GenAI provided.
		DA_12	I document every step of data processing done by GenAI.
Originality and acknowledgement of sources	Authors must ensure that they write an original work and if they have used the work and/or quotations of others, that this has been done appropriately and permission has been obtained where necessary. Proper acknowledgment of the work of others must always be given. Authors must cite publications that have influenced the reported work. Information obtained privately (such as in conversations, correspondence, or discussions with third parties) must not be used or reported without the express written permission of the source. Plagiarism in all forms is unethical and unacceptable behaviour. Plagiarism can take the following forms: a) Claiming someone else's work as your own b) Copying or paraphrasing a large portion of someone else's work (without attribution) c) Claiming research results conducted by someone else	O_1	I ensure that the work written by GenAI does not resemble the work of others.
		O_2	I paraphrase the writing produced by GenAI.
		O_3	I ask GenAI to paraphrase the writing on my report.
		O_4	I copy all the writing produced by GenAI.
		O_5	I avoid listing GenAI as author or co-author.
		O_6	I combine writings from GenAI with my own thoughts.
		O_7	I avoid unauthorized use of personal information in all writing using GenAI.
		O_8	I ensure that I do not enter personal data of the source person into GenAI.
		O_9	When creating verbatim using GenAI, I ensure that the source was aware of and agreed to this in the informed consent.
		O_10	When I want to process data with GenAI, I make sure that this is stated in the informed consent.
		O_11	I respect the sources' decision to refuse permission for the use of GenAI in processing their data.
		O_12	I explain to the interviewees how their data would be used in processing with GenAI.
Citations		C_1	I ensure that the literature references provided by GenAI are relevant to my research.

	Authors should cite relevant, current, and verified literature to support the claims made in their articles. Authors should: a) Ensure that statements in their articles that depend on external sources are accurately quoted. b) Ensure that all citations in their articles can be verified through a permanent Digital Object Identifier or other permanent identifier. c) Avoid excessive and inappropriate citations of their own work. d) Avoid excessive and inappropriate citations of the work of other authors or institutions and refrain from making agreements to cite the work of other authors or groups of authors. e) Avoid excessive citations to support a single statement in their article.	C_2	I confirm that the literature references provided by GenAI come from verified sources.
		C_3	I ensure that the literature references provided by GenAI are relevant to the time.
		C_4	When asking GenAI to provide a quote, I ensure that the quote is accurate.
		C_5	When requesting a quote from GenAI, I make sure that the quote is relevant to my research.
		C_6	I double check the quote GenAI provided by looking up the original source.
		C_7	I avoid excessively citing the literature references obtained from GenAI.
		C_8	I only use quotes from GenAI that are truly necessary for my research.
		C_9	I try to present more of my own writing rather than quoting everything from GenAI.
		C_10	When asking GenAI for ideas, I ask for the source of the idea.
		C_11	I feel responsible for citing the sources I got from GenAI.
		C_12	I feel responsible for ensuring that the sources of quotations obtained from GenAI are included in the bibliography.
Use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork	a) Authors are responsible for the content of their work and include accountability for: - Carefully reviewing and verifying the accuracy, completeness, and neutrality of all AI-generated outputs (including checking sources because AI-generated references can be incorrect or fabricated). - Editing and adapting all material thoroughly to ensure the article represents the author's authentic and original contribution and reflects their own analysis, interpretation, insights, and ideas. - Ensuring that the use of any tools or sources, AI-based or otherwise, is clearly and transparently explained to readers (for the use	UAI_1	I re-read the literature references provided by GenAI.
		UAI_2	I ensure that the results from GenAI match my work.
		UAI_3	I double-check the transcript results provided by GenAI.
		UAI_4	I double-check the data processed by GenAI.
		UAI_5	I make sure to double check the results from GenAI to avoid potential bias.
		UAI_6	I state the use of GenAI in journal searches.
		UAI_7	I make sure to double-check the results from GenAI to avoid any factual errors.
		UAI_8	I explain the use of GenAI in detail in my writing.
		UAI_9	I include GenAI that helped improve the writing in the acknowledgement section of the report.

	of AI tools, Elsevier requires a disclosure statement upon submission). - Ensuring that manuscripts are developed in a manner that protects data privacy, intellectual property rights, and other rights by checking the terms and conditions of each AI tool used. b) Elsevier does not permit the use of generative AI or AI-assisted tools to create or modify images in its articles. This includes enhancing, obscuring, moving, removing, or adding specific features in images or figures. c) The only exception is if the use of AI or AI-assisted tools is part of the research design or research methods (such as AI-assisted imaging approaches to generate or interpret underlying research data, for example in the field of biomedical imaging).	UAI_10	If processing data with GenAI, I include it in the research methods section.
		UAI_11	I censor the source person's name before processing the data using GenAI.
		UAI_12	I censor brand names before processing the data using GenAI
		UAI_13	I avoid listing sensitive information to GenAI.
		UAI_14	I report if I find a potential privacy violation when using GenAI in data processing.

Appendix C

Dimensions, Indicators, and Items after Tryout

DIMENSIONS	DEFINITION	CODE	ITEMS
Reporting standards	Author of research articles should present an accurate account of the work performed as well as an objective discussion of its significance. Articles should contain sufficient detail and references to permit others to replicate the work. Fraudulent or knowingly inaccurate statements constitute unethical behavior and are unacceptable.	RS_3	I made sure to correct the GenAI results that were found to be inaccurate
		RS_4	I recheck the work when using GenAI to double check the spelling that I enter my reports
		RS_12	I avoid using references of questionable accuracy from GenAI results.
Data access and retention	Authors may be asked to provide research data supporting their research articles for editorial review or to comply with journals' open data requirements. Authors are expected to provide public access to such data.	DA_1	I understand the regulations set by the university regarding research.
		DA_2	I am committed to not abusing GenAI
		DA_3	I reread the policies set by the university regarding the use of AI.
		DA_4	I follow the university's established restrictions regarding the use of GenAI.
		DA_5	If unsure, I will ask further about the policies set by the university regarding the use of GenAI in writing
		DA_6	I understand the consequences of violating the rules for using AI in report writing.
		DA_7	I report the results of data analysis processed by GenAI.
		DA_8	I make sure that the data I use in writing with GenAI is original data.
		DA_9	I save all the results that GenAI gives me.
		DA_12	I document every step of data processing done by GenAI.
Originality and acknowledgement of sources	Authors must ensure that they write an original work and if they have used the work and/or quotations of others, that this has been done appropriately and permission has been obtained where necessary. Proper acknowledgment of the work of others must always be given. Authors must cite	O_1	I ensure that the work written by GenAI does not resemble the work of others.
		O_3	I ask GenAI to paraphrase the writing on my report.
		O_4	I copy all the writing produced by GenAI.
		O_5	I avoid listing GenAI as author or co-author.

	publications that have influenced the reported work. Information obtained privately (such as in conversations, correspondence, or discussions with third parties) must not be used or reported without the express written permission of the source. Plagiarism in all forms is unethical and unacceptable behaviour. Plagiarism can take the following forms: a) Claiming someone else's work as your own b) Copying or paraphrasing a large portion of someone else's work (without attribution) c) Claiming research results conducted by someone else	O_6	I combine writings from GenAI with my own thoughts.
		O_7	I avoid unauthorized use of personal information in all writing using GenAI.
		O_8	I ensure that I do not enter personal data of the source person into GenAI.
		O_9	When creating verbatim using GenAI, I ensure that the source was aware of and agreed to this in the informed consent.
		O_10	When I want to process data with GenAI, I make sure that this is stated in the informed consent.
		O_11	I respect the sources' decision to refuse permission for the use of GenAI in processing their data.
		O_12	I explain to the interviewees how their data would be used in processing with GenAI.
Citations	Authors should cite relevant, current, and verified literature to support the claims made in their articles. Authors should: a) Ensure that statements in their articles that depend on external sources are accurately quoted. b) Ensure that all citations in their articles can be verified through a permanent Digital Object Identifier or other permanent identifier. c) Avoid excessive and inappropriate citations of their own work. d) Avoid excessive and inappropriate citations of the work of other authors or institutions and refrain from making agreements to cite the work of other authors or groups of authors. e) Avoid excessive citations to support a single statement in their article.	C_1	I ensure that the literature references provided by GenAI are relevant to my research.
		C_2	I confirm that the literature references provided by GenAI come from verified sources.
		C_3	I ensure that the literature references provided by GenAI are relevant to the time.
		C_4	When asking GenAI to provide a quote, I ensure that the quote is accurate.
		C_5	When requesting a quote from GenAI, I make sure that the quote is relevant to my research.
		C_6	I double check the quote GenAI provided by looking up the original source.
		C_7	I avoid excessively citing the literature references obtained from GenAI.
		C_8	I only use quotes from GenAI that are truly necessary for my research.
		C_9	I try to present more of my own writing rather than quoting everything from GenAI.
		C_10	When asking GenAI for ideas, I ask for the source of the idea.
		C_11	I feel responsible for citing the sources I got from GenAI.

		C_12	I feel responsible for ensuring that the sources of quotations obtained from GenAI are included in the bibliography.
Use of generative AI and AI-assisted technologies in scientific writing and in figures, images, and artwork	a) Authors are responsible for the content of their work and include accountability for: - Carefully reviewing and verifying the accuracy, completeness, and neutrality of all AI-generated outputs (including checking sources because AI-generated references can be incorrect or fabricated). - Editing and adapting all material thoroughly to ensure the article represents the author's authentic and original contribution and reflects their own analysis, interpretation, insights, and ideas. - Ensuring that the use of any tools or sources, AI-based or otherwise, is clearly and transparently explained to readers (for the use of AI tools, Elsevier requires a disclosure statement upon submission). - Ensuring that manuscripts are developed in a manner that protects data privacy, intellectual property rights, and other rights by checking the terms and conditions of each AI tool used. b) Elsevier does not permit the use of generative AI or AI-assisted tools to create or modify images in its articles. This includes enhancing, obscuring, moving, removing, or adding specific features in images or figures. c) The only exception is if the use of AI or AI-assisted tools is part of the research design or research methods (such as AI-assisted imaging approaches to generate or interpret underlying research data, for example in the field of biomedical imaging).	UAI_6	I state the use of GenAI in journal searches.
		UAI_7	I make sure to double-check the results from GenAI to avoid any factual errors.
		UAI_9	I explain the use of GenAI in detail in my writing.
		UAI_10	I include GenAI that helped improve the writing in the acknowledgement section of the report.
		UAI_11	If processing data with GenAI, I include it in the research methods section.
		UAI_12	I censor the source person’s name before processing the data using GenAI.
		UAI_13	I censor brand names before processing the data using GenAI
		UAI_14	I avoid listing sensitive information to GenAI.