ePublishing

ΕΘΝΙΚΟ ΚΕΝΤΡΟ ΤΕΚΜΗΡΙΩΣΗΣ
NATIONAL DOCUMENTATION CENTRE

Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση

Τόμ. 1 (2025)

14ο Συνέδριο ΕΤΠΕ «ΤΠΕ στην Εκπαίδευση»

Teachers' Evaluation Skills on AI-generated Data Utilized in Educational Environmental Research

Maria Christoforaki, Anastasia Georgiou, Dimitris Rorris, Evangelia Mavrikaki, Apostolia Galani

doi: [10.12681/cetpe.9474](https://doi.org/10.12681/cetpe.9474)

Βιβλιογραφική αναφορά:

Christoforaki, M., Georgiou, A., Rorris, D., Mavrikaki, E., & Galani, A. (2026). Teachers’ Evaluation Skills on AI-generated Data Utilized in Educational Environmental Research. *Συνέδρια της Ελληνικής Επιστημονικής Ένωσης Τεχνολογιών Πληροφορίας & Επικοινωνιών στην Εκπαίδευση*, *1*, 229–238. https://doi.org/10.12681/cetpe.9474

Teachers' Evaluation Skills on AI-generated Data Utilized in Educational Environmental Research

Maria Christoforaki, Anastasia Georgiou, Dimitris Rorris, Evangelia Mavrikaki, Apostolia Galani
mchrst@primedu.uoa.gr, anastasg@primedu.uoa.gr, dimror@primedu.uoa.gr, emavrikaki@primedu.uoa.gr, ligalani@primedu.uoa.gr
Department of Pedagogy and Primary Education, National and Kapodistrian University of Athens

Abstract

The growing integration of Artificial Intelligence (AI) in education offers both opportunities and challenges, particularly concerning the evaluation of the validity and reliability of AI-generated data. This study examines in-service primary and secondary teachers' attitudes and skills in critically assessing textual information produced by freely accessible AI tools such as ChatGPT, DeepSeek, and Gemini, with a specific focus on urban heatwaves as climate change-related issue. Employing a quantitative methodology, 279 teachers were surveyed using a questionnaire grounded in the CRAAP framework to assess their skills in evaluating AI-generated content. Findings reveal that some educators are concerned about students relying on AI tools for environmental research, worrying that such use may reinforce misconceptions, particularly in the context of socio-environmental issues such as urban heatwaves. Furthermore, several teachers reported difficulties in effectively assessing the validity and reliability of AI-generated information across all dimensions of the CRAAP framework. These findings highlight the need for targeted training programs to enhance teachers' digital literacy and critical evaluation skills in the context of emerging AI technologies.

Keywords: AI-generated data, digital literacy, evaluation, skills

Introduction

The integration of Artificial Intelligence (AI) in education offers transformative potential, yet it also presents challenges, particularly in evaluating the validity and reliability of AI-generated data (Jemetz et al., 2025). This is especially critical in environmental education, where accurate information is essential for addressing complex socio-environmental issues like urban heatwaves. Urban heatwaves exemplify such challenges, as they represent an intensifying threat to cities, driven by the escalating climate crisis (Cheng et al., 2018). Their growing frequency demands a nuanced understanding of contributing factors such as urban form, construction materials, vegetation, and socio-economic conditions. As educators and students increasingly turn to AI tools for information, the ability to critically evaluate data becomes essential. In this context, understanding urban resilience is not only a matter of scientific literacy, but also a foundation for fostering responsible citizenship in an era of accelerating climate change.

Recent studies emphasize the necessity of equipping educators with AI literacy competencies to navigate and critically assess AI-generated content, highlighting the importance of teachers developing critical thinking to effectively integrate AI tools into their teaching (Oates & Johnson, 2025). Despite these insights, a significant gap remains in research focusing on the intersection of AI tool utilization and environmental education. This study aims to bridge this gap by enhancing the skills of primary and secondary education teachers in evaluating data generated by AI free-access tools concerning the example issue of urban heatwaves. By focusing on developing skills in recognizing misinformation and integrating

AI-generated data into educational projects, educators will be empowered to navigate the complexities of AI in environmental contexts effectively. Thus, our project's research questions are:

(a) Which are primary and secondary school teachers' attitudes and perspectives on evaluating data (text information) generated by AI tools regarding urban heatwaves as a result of climate crisis?

(b) To what extent can primary and secondary education teachers effectively evaluate the validity and reliability of AI-generated data regarding the example issue of urban heatwaves?

(c) How do teachers' self-reported attitudes correlate with their evaluative skills of AI generated data about urban heatwaves as a climate change related issue?

(d) How do certain demographic factors (e.g. age, gender, teaching experience) influence primary and secondary education teachers' (i) attitudes and perspective in AI-generated data on urban heatwaves as a climate change related issue and (ii) evaluating skills to critically assess and utilize AI-generated data on urban heatwaves and climate change in their teaching?

Materials and methods

Sample

The target population in our research was in-service primary and secondary education teachers from Greece (N = 279), as described in Figure 1. As for secondary education, we included teachers from various subject areas rather than limiting participation to science teachers, aligning with the interdisciplinary nature of environmental education and reflecting our aim to promote AI literacy and critical data evaluation skills across the broader teaching workforce.

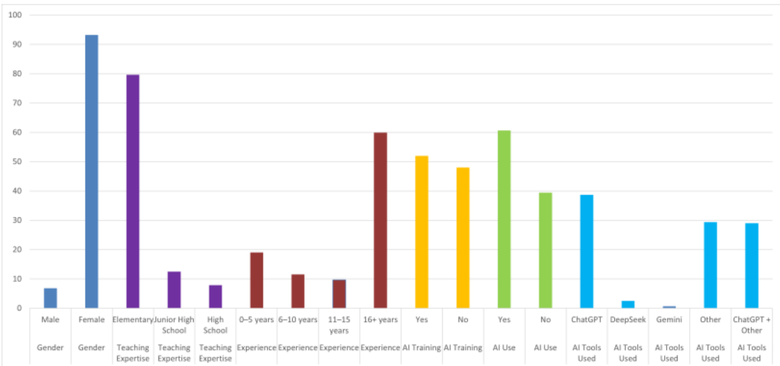


Figure 1. Description of the sample

Tools

We adopted a quantitative research approach and designed a digital questionnaire to elicit teachers' primary attitudes and skills in evaluating the validity and reliability of AI generated data regarding urban heatwaves, when using AI tools like ChatGPT, DeepSeek, and Gemini. The items of the questionnaire were developed by: (a) reviewing relevant literature (Kong et al., 2024), (b) results of the ERASMUS KA2+ Project "HEATWAVES AWARENESS THROUGH ONLINE LEARNING" (https://heatwaves-project.eu/), proposing CRAAP test as a framework for assessing the validity and reliability of data.

The CRAAP test, an evaluative framework encompassing Currency, Relevance, Authority, Accuracy, and Purpose, served as the primary tool for assessing the quality of AI-generated environmental data (Meriam Library, n.d.). The CRAAP test provides a systematic approach for critically analyzing information sources, emphasizing not only factual accuracy, but also the context, authorship, and intended use of the data. In the context of this research, the CRAAP test was adapted to guide teachers in scrutinizing AI-generated content about urban heatwaves (Figure 2).

Dimension	1(Very Low)	2(Low)	3(Moderate)	4(High)	5(Very High)
Currency	The content is outdated or undated, lacks any temporal context, and may contain broken or irrelevant links.	Dated content (3+ years old); minimal context indicating when data was generated.	Moderately current (1-3 years); some evidence of update or temporal reference.	Content is recent (within 1-2 years); the update information is clear and verifiable.	Very recent (within 12 months); all data is up-to-date with no broken links; reflects current climate science and policy.
Relevance	Content is off-topic or irrelevant to the urban heatwave theme; lacks instructional value.	Only loosely connected to urban heatwaves or climate education; too general or superficial.	Content is moderately relevant; it provides some value for teaching but lacks depth or clarity.	Supports specific teaching goals on urban heatwaves; offers structured educational value.	Highly aligned with the curriculum; provides depth, examples, and clear connections to learning outcomes.
Authority	No identifiable author or organization; unverified or questionable domain (e.g., personal blog).	The source lacks clear qualifications or affiliations and has limited evidence of expertise.	The author or source has some credibility, possibly linked to general educational or climate resources.	Authored by a recognized educator, academic, or organization; credible domain (.org, .edu).	Authored or endorsed by a recognized authority (e.g., university, government agency, peer-reviewed body); domain is .edu or .gov.
Accuracy	Content includes factual errors, misinformation, or grammar/spelling mistakes; lacks data support.	Some inaccuracies or vague/unsubstantiated claims; minimal referencing.	Mostly accurate; may lack citations but shows a reasonable level of correctness.	Factually sound with few errors; some evidence of data or expert sources.	Entirely accurate, well-researched, clearly cited, error-free, and aligned with established scientific understanding.
Purpose	Content is biased, promotional, or persuasive; it lacks transparency about intent.	Slightly biased or includes promotional language; unclear if intended for education.	Mostly informative but may include subtle bias or unbalanced perspectives.	Primarily informative and objective; minimal bias; supports learning.	Clear educational purpose; balanced, unbiased, and written to support critical thinking and research.

Figure 2. Evaluation framework adapted from the CRAAP criteria for analyzing AI-generated educational materials on urban heatwaves and climate-related topics (Meriam Library, n.d.)

The questionnaire consisted of 23 items, 7 of which were about the participant's profile (age, gender, teaching experience etc.), 6 items were 5-point Likert items (1 = totally disagree to 5 = totally agree), regarding teachers' attitudes and perspectives on evaluating data generated by AI tools regarding urban heatwaves as a result of climate crisis, and 10 5-point

Likert Items (1 = Very Low to 5 = Very High), based on the CRAAP test, assessing its distinct aspects separately for each item. Table 1 shows the research instrument's subscales and reliability, all above acceptance levels according to Nunnally (1978).

Table 1. The questionnaire's subscales

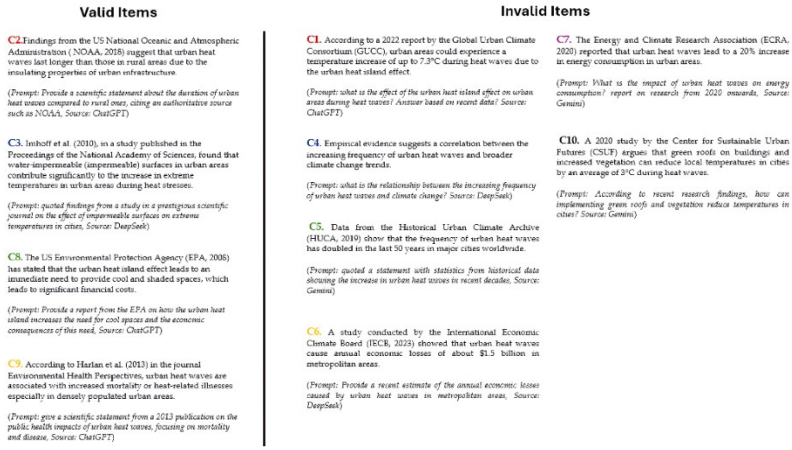
Subscale	Cronbach's Alpha	Items
AI Attitude (A_I_A)	.822	6
CRAAP AI Skills (C_A_S)	.990	50*
Total	**.986**	**56**

*10 items multiplied with the 5 aspects of CRAAP test for each one

Based on the data we gathered, we created two composite variables, Attitude about Artificial Intelligence (A_I_A) and CRAAP Artificial Intelligence Skills (C_A_S), by estimating the means of the participant's answers to each subscale's questions.

Attitude about Artificial Intelligence (A_I_A): This subscale assessed teachers' self-reported attitudes and perspectives on utilizing AI generated data in their teaching regarding environmental research concentrated on the issue of urban heatwaves (B1-B6). Example items included statements such as "*B3. I am concerned that the use of AI tools by students in my class to inform them about environmental issues such as urban heatwaves may reinforce their misconceptions*".

CRAAP Artificial Intelligence Skills (C_A_S): This subscale concentrated on teachers' understanding of evaluating the validity and reliability of AI-generated data regarding urban heatwaves as a climate change related issue (C1-C10). The items included were closed-ended, while their structure included both a prompt and an answer generated by an AI tool. Teachers were expected to assess the validity and reliability of the information, referencing the aspects of CRAAP test such as "*C1. According to a 2022 report by the Global Urban Climate Consortium (GUCC), urban areas can experience temperature increases of up to 7.3°C during heatwaves due to the urban heat island phenomenon. (Prompt: What is the impact of the urban heat island phenomenon on urban areas during heatwaves? Answer based on recent data. Source: ChatGPT)*". The included items represented either valid AI-generated data, or invalid as described in Figure 3. It is important to clarify that there were not negatively phrased items demanding reverse coding to ensure consistency.



Valid Items

C2. Findings from the US National Oceanic and Atmospheric Administration (NOAA, 2018) suggest that urban heat waves last longer than those in rural areas due to the insulating properties of urban infrastructure.

(Prompt: Provide a scientific statement about the duration of urban heat waves compared to rural ones, citing an authoritative source such as NOAA, Source: ChatGPT)

C3. Imhoff et al. (2010), in a study published in the Proceedings of the National Academy of Sciences, found that water-impermeable (impermeable) surfaces in urban areas contribute significantly to the increase in extreme temperatures in urban areas during heat stresses.

(Prompt: quoted findings from a study in a prestigious scientific journal on the effect of impermeable surfaces on extreme temperatures in cities, Source: DeepSeek)

C8. The US Environmental Protection Agency (EPA, 2006) has stated that the urban heat island effect leads to an immediate need to provide cool and shaded spaces, which leads to significant financial costs.

(Prompt: Provide a report from the EPA on how the urban heat island increases the need for cool spaces and the economic consequences of this need, Source: ChatGPT)

C9. According to Harlan et al. (2013) in the journal Environmental Health Perspectives, urban heat waves are associated with increased mortality or heat-related illnesses especially in densely populated urban areas.

(Prompt: give a scientific statement from a 2013 publication on the public health impacts of urban heat waves, focusing on mortality and disease, Source: ChatGPT)

Invalid Items

C1. According to a 2022 report by the Global Urban Climate Consortium (GUCC), urban areas could experience a temperature increase of up to 7.3°C during heat waves due to the urban heat island effect.

(Prompt: what is the effect of the urban heat island effect on urban areas during heat waves? Answer based on recent data? Source: ChatGPT)

C4. Empirical evidence suggests a correlation between the increasing frequency of urban heat waves and broader climate change trends.

(Prompt: what is the relationship between the increasing frequency of urban heat waves and climate change? Source: DeepSeek)

C5. Data from the Historical Urban Climate Archive (HUCA, 2019) show that the frequency of urban heat waves has doubled in the last 50 years in major cities worldwide.

(Prompt: quoted a statement with statistics from historical data showing the increase in urban heat waves in recent decades, Source: Gemini)

C6. A study conducted by the International Economic Climate Board (IECB, 2023) showed that urban heat waves cause annual economic losses of about $1.5 billion in metropolitan areas.

(Prompt: Provide a recent estimate of the annual economic losses caused by urban heat waves in metropolitan areas, Source: DeepSeek)

C7. The Energy and Climate Research Association (ECRA, 2020) reported that urban heat waves lead to a 20% increase in energy consumption in urban areas.

(Prompt: What is the impact of urban heat waves on energy consumption? report on research from 2020 onwards, Source: Gemini)

C10. A 2020 study by the Center for Sustainable Urban Futures (CSUF) argues that green roofs on buildings and increased vegetation can reduce local temperatures in cities by an average of 3°C during heat waves.

(Prompt: According to recent research findings, how can implementing green roofs and vegetation reduce temperatures in cities? Source: Gemini)

Figure 3. Item Description of C_A_S Variable

To assess the construct validity of the CRAAP-based items evaluating teachers' skills in assessing AI-generated data, exploratory factor analysis (EFA) was conducted separately for valid and invalid content (Table 2). Sampling adequacy was confirmed via the Kaiser-Meyer-Olkin (KMO) measure, which yielded excellent values of .954 for valid and .965 for invalid items, both exceeding the 0.90 threshold. Bartlett's Test of Sphericity confirmed significant deviations from the identity matrix for both categories (valid: $\chi^2(210) = 7444.350$, $p < .001$; invalid: $\chi^2(435) = 10907.394$, $p < .001$), supporting the factorability of the data.

Table 2. Exploratory Factor Analysis across all items of C_A_S Variable

Test Component	Valid Items	Invalid Items
Kaiser-Meyer-Olkin Measure	.954	.965
Bartlett's Test of Sphericity		
Approx. Chi-Square	7444.350	10907.394
Degrees of Freedom (df)	210	435
p	**.000**	**.000**

Results

Teachers' self-reported attitudes and perspectives on AI-generated data on urban heatwaves

Most teachers (55.2%) reported a neutral stance on the reliability of AI-generated information (B1), while 34.8% agreed that such information is reliable. Regarding the verification of sources cited by AI tools (B2), 39.8% indicated that they actively verify source validity and currency. Concerns about the potential reinforcement of students' misconceptions through AI tools (B3) were evident, with 55.6% expressing neutrality. Regarding the ability to distinguish between AI-generated content and data from scientific sources (B4), 28% reported agreement with experiencing difficulty in this area. Moreover, many teachers demonstrated uncertainty or lack of confidence in verifying the validity of AI-generated statistics and visual data (B5), with 39.1% neutral and 39.7% indicating disagreement. Similarly, responses to B6 showed limited confidence in teaching students to assess AI-generated environmental data, with 39.4% agree neither with nor disagree with this statement (Figure 4).

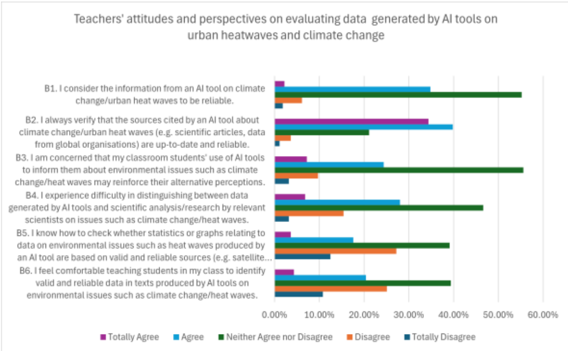


Figure 4. Teachers' answers regarding their self-reported attitudes and perspectives on evaluating AI generated data

Teachers' evaluating skills of AI generated data on urban heatwaves based on the CRAAP test

Analysis of responses across the CRAAP dimensions revealed consistent patterns of moderate to favorable evaluations, reflecting teachers' critical engagement with both valid and less reliable AI-generated environmental data (Figure 5).

Currency: Most responses were clustered in the moderate (35.8%-42.3%) and high (19.4%-25.1%) ratings, with a smaller proportion receiving very high ratings (8.6%-13.3%). Notably, valid items such as C2 and C9 received slightly elevated ratings, indicating that educators were more attuned to the temporal relevance of credible content. In contrast, invalid items (e.g., C1, C5) elicited a higher frequency of neutral or low responses.

Relevance: A similar distribution was observed, with moderate ratings ranging from 37.6% to 44.1% and high ratings from 19.7% to 26.9%. Valid items, particularly C2 and C3, received a greater proportion of high and very high relevance evaluations (up to 26.9%), suggesting that educators found these items to be more appropriately aligned with educational goals and contexts.

Authority: Most responses fell within the moderate ratings (40.5%-47.2%), followed by high (17.6%-26.9%), while low or very low ratings remained below 7.5%. Valid items such as C2 and C9 tended to receive slightly higher authority scores, whereas invalid items (e.g., C6, C7) were more frequently rated as moderate or low.

Accuracy: Valid items, such as C9, received the highest combined proportions of high and very high accuracy ratings (up to 25.4%), reflecting participants' stronger recognition of factual consistency in credible content. Conversely, invalid items (e.g., C4, C6) were more commonly rated as moderate or low.

Purpose: Evaluations of purpose were most frequently moderate (41.2%-44.4%), with high (19.0%-25.1%) and very high (7.5%-10.8%) ratings also present. Valid items such as C8 and C9 scored slightly higher in terms of transparency of purpose, though overall differences across items were relatively small.

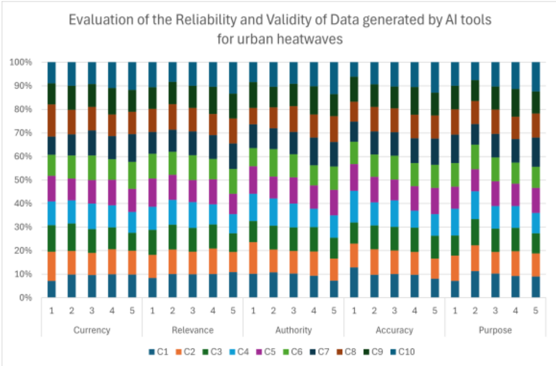


Figure 5. Teachers' answers regarding their evaluation on AI generated data about urban heatwaves based on CRAAP

Correlation between attitude and evaluative skills

To assess the relationship between teachers' attitudes toward AI-generated data (A_I_A) and their evaluative skills as measured by the CRAAP framework (C_A_S), a Pearson correlation analysis was conducted (Table 3). The results indicated a statistically significant, but weak positive correlation between A_I_A and C_A_S scores ($r = 0.206$, $p = .001$, $n = 267$), suggesting that teachers with more favorable attitudes toward the use of AI-generated data also tended to demonstrate slightly higher self-reported skills in evaluating the validity and reliability of such data.

Table 3. Correlation of teachers' attitude and skills on evaluating AI-generated data on urban heatwaves

		A_I_A	C_A_S
A_I_A	Pearson Correlation	1	.206**
	Sig. (2-tailed)		0.001
	N	279	267
C_A_S	Pearson Correlation	.206**	1
	Sig. (2-tailed)	0.001	
	N	**267**	**267**

Correlation of teachers' attitudes and evaluation skills on AI-generated data about urban heatwaves to the demographic variables

The analysis revealed that none of the demographic variables significantly predicted differences in the A_I_A or C_A_S scores. The model for A_I_A was statistically significant ($F = 1.352$, $p = .043$), indicating some explanatory variance, though no individual demographic factor reached significance (Table 4). In contrast, the C_A_S model was non-significant (F = 1.008, $p = .484$), suggesting no systematic influence. Previous AI training approached marginal significance for C_A_S ($F = 2.363$, $p = .127$), implying a weak, but non-robust association. Similarly, familiarity with specific AI tools showed a non-significant trend for A_I_A ($F =$

1.701, p = .096), suggesting a subtle influence. However, teaching experience, education level, and active AI use in teaching had no significant impact.

Table 4. Tests of Between-Subjects Effects for Demographic Variables

Source		Type III Sum of Squares	*df*	*Mean Square*	*F*	*p*
Corrected Model	A_I_A	30.423	144	0.211	1.352	0.043
	C_A_S	98.073	144	0.681	1.008	0.484
Intercept	A_I_A	552.121	1	552.121	0.000	0.000
	C_A_S	481.917	1	481.917	713.082	0.000
Gender	A_I_A	0.023	1	0.023	0.145	0.704
	C_A_S	0.252	1	0.252	0.373	0.543
Age	A_I_A	1.113	4	0.278	1.780	0.137
	C_A_S	1.823	4	0.456	0.674	0.611
Teaching Expertise	A_I_A	0.032	2	0.016	0.103	0.903
	C_A_S	0.241	2	0.120	0.178	0.837
Teaching Experience	A_I_A	0.389	3	0.130	0.830	0.480
	C_A_S	1.151	3	0.384	0.568	0.637
Previous AI Training	A_I_A	0.002	1	0.002	0.014	0.907
	C_A_S	1.597	1	1.597	2.363	0.127
Use of AI tools in teaching	A_I_A	0.040	1	0.040	0.253	0.616
	C_A_S	0.489	1	0.489	0.724	0.397
Used AI tools	A_I_A	2.393	9	0.266	1.701	0.096
	C_A_S	3.965	9	0.441	0.652	0.751

Discussion

The results of this study provide significant insights into the attitudes and evaluative skill of primary and secondary in-service teachers of AI-generated data (text information) on urban heatwaves, revealing a spectrum of perspectives and varying levels of proficiency in engaging with AI-driven environmental data.

A significant portion of teachers maintained a neutral stance on the reliability of AI-generated information, with 55.2% neither agreeing nor disagreeing with the assertion that such data is reliable. A smaller, yet still substantial proportion (34.8%) expressed agreement with its reliability, suggesting a prevailing caution among educators regarding the trustworthiness of AI output. This cautious outlook reflects broader concerns in the literature, where educators frequently question the accuracy and potential biases inherent in AI tools (Zawacki-Richter et al., 2019).

Teachers' apprehensions about AI tools reinforcing student misconceptions were also evident, with 55.6% expressing neutrality and 24.4% agreeing with the notion that AI may inadvertently perpetuate misunderstandings. This aligns with prior research highlighting teachers' concerns about the uncritical use of AI in classrooms, particularly when the data pertains to complex socio-environmental issues such as climate change and urban heatwaves

(Selwyn, 2019). These findings point to an urgent need for training programs that equip teachers not only with the technical skills to use AI tools, but also with the critical thinking abilities required to evaluate the data provided by these tools, especially in the context of urban heatwaves, where the interpretation of data can have significant implications for both education and policy.

Furthermore, a notable proportion of teachers reported difficulty distinguishing between AI-generated content and scientifically verified data, with 28% expressing challenges in this area. This difficulty reveals a critical gap in teacher education, as the ability to evaluate the authenticity of data is essential for effective teaching, particularly in fields that require nuanced data analysis such as environmental research (Wineburg & McGrew, 2019). While AI technologies are becoming more prevalent in educational settings, many teachers remain unprepared to navigate the complexities of AI-generated content, presenting a significant barrier to the successful integration of AI tools in the classroom.

Teachers evaluated AI-generated content using the five CRAAP criteria with most responses clustering within the moderate to high range, suggesting a generally positive, yet cautious, stance toward the perceived quality of AI-generated data. Notably, valid items consistently received higher ratings across all dimensions, particularly in accuracy, relevance, and currency, indicating that educators were able to identify and favor more credible and pedagogically appropriate content. In contrast, invalid items attracted more moderate or low ratings, especially regarding authority and accuracy, reflecting a degree of critical engagement with less reliable or outdated information. These findings align with existing literature indicating that educators tend to exhibit greater confidence in assessing relevance and accuracy, while demonstrating more uncertainty when evaluating authority and purpose, especially in the context of emerging technologies such as AI (Nguyen et al., 2021). This underscores a persistent gap in teachers' critical data literacy, which is particularly problematic in educational domains addressing complex socio-environmental issues, such as urban heatwaves. In such contexts, the ability to assess the credibility and accuracy of information is essential not only for effective instruction, but also for cultivating students' decision-making and argumentation skills (Lombardi et al., 2021).

The study identified a statistically significant, but weak positive correlation between teachers' attitudes toward AI-generated data and their self-reported evaluative skills. Educators with more favorable attitudes demonstrated slightly higher confidence in assessing data reliability and validity. However, the modest strength of this relationship suggests that evaluative competence is influenced by additional factors. This aligns with prior research indicating that while attitudes may support technology adoption, practical experience and targeted training are more strongly associated with effective integration in educational contexts (Ottenbreit-Leftwich et al., 2018).

Interestingly, demographic factors such as gender, age, teaching experience, prior AI training, and the use of AI tools in teaching did not significantly impact teachers' attitudes or evaluative skills regarding AI-generated data. Additionally, while the use of AI tools in teaching showed a non-significant trend toward influencing attitudes, this relationship did not reach statistical significance, indicating that familiarity with AI tools may shape teachers' views on their utility, though not necessarily their evaluation skills.

The findings highlight the need for training programs that both familiarize teachers with AI tools and develop the data literacy skills required to assess the validity and reliability of AI-generated data. As AI use in environmental education grows, teachers must be equipped to critically evaluate this data and make informed classroom decisions.

Limitations

The small sample of secondary education teachers limits the generalizability of our findings to the wider in-service teacher population in Greece. Additionally, self-reported data may introduce biases, such as social desirability and recall errors. While demographic factors like age and gender were considered, other relevant variables were not included. Future research could address these limitations by using a larger, more representative sample and including additional variables.

References

Bingjie L. (2021). In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human-AI interaction. *Journal of Computer-Mediated Communication, 26*(6), 384-402. https://doi.org/10.1093/jcmc/zmab013

Cheng, J., Xu, Z., Bambrick, H., Su, H., Tong, S., & Hu, W. (2018). Heatwave and elderly mortality: An evaluation of death burden and health costs considering short-term mortality displacement. *Environment International, 115*, 334-342. https://doi.org/10.1016/j.envint.2018.03.041

Jemetz, M., Dolezal, D., & Motschnig, R. (2025). Secondary teachers' self-perceived AI competences in relation to renowned European digital competence frameworks. In Z. Pluhár, & B. Gaál (Eds.), *Informatics in schools. Innovative approaches to computer science teaching and learning. ISSEP 2024* (Lecture Notes in Computer Science, vol. 15228, pp. 3-15). Springer. https://doi.org/10.1007/978-3-031-73474-8_1

Kong, S. C., Yang, Y., & Hou, C. (2024). Examining teachers' behavioural intention of using generative artificial intelligence tools for teaching and learning based on the extended technology acceptance model. *Computers & Education: Artificial Intelligence, 7*, 100328. https://doi.org/10.1016/j.caeai.2024.100328

Lombardi, D., & Bailey, J. M. (2023). Science learning and teaching. In *Handbook of educational psychology* (pp. 531-552). Routledge.

Meriam Library, California State University, Chico. (n.d). *Evaluating information - Applying the CRAAP test*. https://guides.lib.uchicago.edu/ld.php?content_id=77584250

Nguyen, A., Ngo, H. N., Hong, Y., Kovanović, V., Gašević, D., & Dawson, S. (2023). Ethical principles for artificial intelligence in education. Education and Information Technologies, 28, 4221-4241. https://doi.org/10.1007/s10639-022-11316-w

Nunnally, J. C. (1978). An overview of psychological measurement. In *Clinical diagnosis of mental disorders: A handbook* (pp. 97-146). Springer. https://doi.org/10.1007/978-1-4684-2490-4_4

Oates, A., & Johnson, D. (2025). ChatGPT in the classroom: Evaluating its role in fostering critical evaluation skills. *International Journal of Artificial Intelligence in Education, 2025*, 1-32. https://doi.org/10.1007/s40593-024-00452-8

Ottenbreit-Leftwich, A., Liao, J. Y. C., Sadik, O., & Ertmer, P. (2018). Evolution of teachers' technology integration knowledge, beliefs, and practices: How can we support beginning teachers' use of technology? *Journal of Research on Technology in Education, 50*(4), 282-304. https://doi.org/10.1080/15391523.2018.1487350

Selwyn, N., Hillman, T., Eynon, R., Ferreira, G., Knox, J., Macgilchrist, F., & Sancho-Gil, J. M. (2019). What's next for Ed-Tech? Critical hopes and concerns for the 2020s. *Learning, Media and Technology, 45*(1), 1-6. https://doi.org/10.1080/17439884.2020.1694945

Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: Reading less and learning more when evaluating digital information. *Teachers College Record, 121*(11), 1-40. https://doi.org/10.1177/016146811912101102

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - Where are the educators? International Journal of Educational Technology in Higher Education, *16*, 39. https://doi.org/10.1186/s41239-019-0171-0