



Do Students Verify What AI Tells Them? A Survey of AI Use and Self-Assessed AI Competence in High School Physics

Do Thi Hong Phuong¹, Nguyen Quang Linh^{2*}

¹ Bao Yen No. 2 High School, Lao Cai Province, Vietnam

² Faculty of Physics, Thai Nguyen University of Education, Thai Nguyen, Vietnam

ABSTRACT: Generative artificial intelligence (AI) tools are increasingly used by secondary school students, yet in physics their fluent answers may contain conceptual or procedural errors. This study surveyed 480 upper-secondary students (Grades 10 - 12) in Lao Cai province, Vietnam, to describe how they use AI tools in physics learning, how they self-assess their competence in working with these tools, and what conditions and preferences they report regarding AI-integrated instruction. An online questionnaire of 18 five-point Likert items organised in three groups was administered; internal consistency was acceptable ($\alpha = .849-.914$; $\omega = .851-.915$). Students reported using AI tools occasionally ($M = 3.13$), mainly to preview lesson content and to obtain solutions to exercises, and less often for experimental design or data processing. Self-assessed competence was at the agreement level ($M = 3.64$), with 27.7% to 29.8% of students neutral on each item. At the item level, the distributions of self-assessment and reported practice did not correspond: 64.0% of students agreed that they distinguished their own work from AI-generated content, whereas 51.7% reported submitting unedited AI output at least sometimes. Most students expressed readiness to prepare before class (71.0%) and wished to be taught how to check AI answers (69.4%), while 11.5% reported insufficient devices or connectivity at home. The findings suggest that physics instruction should include explicit verification tasks and disclosure requirements, and that students' AI competence should be assessed with evidence beyond self-report.

Corresponding Author:
Nguyen Quang Linh

Published Online:
September 29, 2026

License:

This is an open access article under the CC BY 4.0 license:

<https://creativecommons.org/licenses/by/4.0/>

KEYWORDS:

AI literacy, generative AI, physics education, self-assessment, upper-secondary students, flipped learning

Cite the Article: Phuong, D.T.H, Linh, N.Q. (2026). Do Students Verify What AI Tells Them? A Survey of AI Use and Self-Assessed AI Competence in High School Physics. *International Journal of Human Research and Social Science Studies*, 3(9), 836-843. <https://doi.org/10.55677/ijhrsss/07-2026-Vol03I09>

1. INTRODUCTION

Conversational assistants built on large language models have become part of the everyday learning environment of secondary school students. Students can obtain explanations, worked solutions and experimental procedures within seconds, often outside the teacher's view. In physics, this availability raises a specific question. Large language models generate text by predicting plausible continuations from patterns in their training data, so their output is shaped by statistical plausibility rather than by physical correctness (Polverini & Gregorcic, 2024). A case analysed by Gregorcic and Pendrill (2023) illustrates the consequence: a chatbot produced a solution to a mechanics problem that was presented coherently but contained incorrect and internally inconsistent reasoning, and the coherence of the presentation made the errors harder to detect.

For students, productive use of these tools therefore depends on more than the ability to operate them. It requires the ability to judge whether an answer is consistent with physical laws, with the conditions under which those laws apply and with available data. Current frameworks of AI literacy recognise this requirement by placing evaluation and ethical use alongside knowledge and application (Long & Magerko, 2020; Ng et al., 2021). The AI competency framework for students published by UNESCO organises competence into four dimensions and three levels of progression (Miao & Shiohira, 2024), and a framework adapted to Vietnamese general education has recently been proposed (Lê et al., 2025). These frameworks define what students should be able to do, but they provide limited information about what students currently do when they use AI tools in a particular school subject.

Three gaps motivate the present study. First, a systematic review of AI literacy in K-12 education found that most studies describe learning experiences in technical domains or propose conceptual models, while studies situated in core school subjects

remain few (Casal-Otero et al., 2023). Evidence on how students use AI tools in physics, and for which learning purposes, is correspondingly limited. Second, the measurement of AI literacy relies largely on self-report questionnaires (Ng et al., 2024). Research on digital literacy has shown that students' perceived competence may diverge considerably from their actual performance (Porat et al., 2018), which suggests that self-assessed AI competence should be interpreted alongside other information about students' practices. Third, instructional designs that move part of the learning process outside the classroom, such as flipped learning, are increasingly discussed in combination with AI tools (López-Villanueva et al., 2024). Whether students have the conditions and the willingness to engage in such designs is an empirical question that should be examined before the designs are implemented.

The present study addresses these gaps through a survey of upper-secondary students in Vietnam. It forms the needs-analysis stage of a larger project on developing students' AI competence in teaching the Dynamics topic of Grade 10 physics through flipped learning. Three research questions guided the study:

RQ1. To what extent and for what purposes do upper-secondary students use AI tools in physics learning?

RQ2. How do students self-assess their competence in working with AI tools, and how do these self-assessments compare with the practices they report?

RQ3. What conditions and preferences do students report regarding AI-integrated physics instruction that relies on preparation before class?

The study contributes subject-specific descriptive evidence on AI use in physics learning, an item-level comparison between self-assessed competence and reported practice, and information on the conditions for AI-integrated flipped instruction in a Southeast Asian context.

2. THEORETICAL FRAMEWORK

2.1. AI competence of secondary school students

Long and Magerko (2020) defined AI literacy as a set of competencies that enables individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home and in the workplace. From an interdisciplinary review, they proposed 17 competencies together with design considerations for learner-centred technologies. Ng et al. (2021) synthesised the literature into four aspects: knowing and understanding AI, using and applying AI, evaluating and creating AI, and AI ethics. The UNESCO framework for students describes four dimensions, namely a human-centred mindset, the ethics of AI, AI techniques and applications, and AI system design, each at three levels labelled understand, apply and create (Miao & Shiohira, 2024). Lê et al. (2025) adapted these international references to the conditions of Vietnamese general education. Across these frameworks, critical evaluation and responsible use are treated as constituents of competence rather than as additions to it.

For the context of physics learning, the present study adopts an operational structure of four dimensions derived from these frameworks (Figure 1). Dimension D1 concerns understanding what AI tools are and how they produce answers, together with a human-centred view of which parts of a learning task remain the student's responsibility. Dimension D2 concerns using AI tools in physics learning, including formulating requests with sufficient data and conditions. Dimension D3 concerns evaluating and verifying AI output with physics knowledge and data, and is the dimension most specific to the subject. Dimension D4 concerns responsible and ethical use, including distinguishing one's own work from AI-generated content. Figure 1 also indicates how the questionnaire items described in the Method section relate to each dimension.

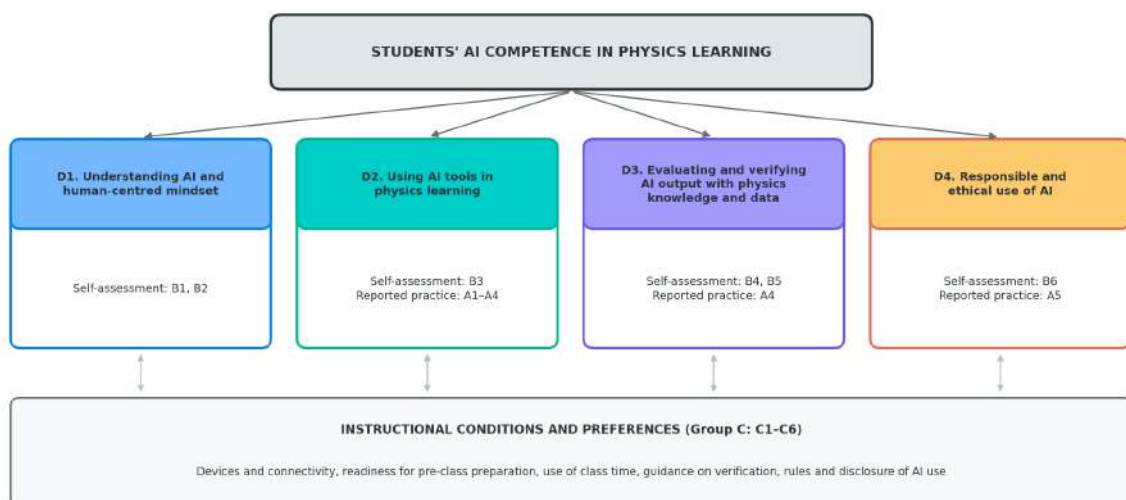


Figure 1. Conceptual framework of students' AI competence in physics learning and its relation to the questionnaire items

2.2. Generative AI in physics learning

Research on generative AI in physics education has followed two lines. The first examines the performance of chatbots on physics tasks. Polverini and Gregorcic (2024) described how large language models work, showed how the formulation of a prompt influences the quality of output on conceptual physics tasks, and discussed the implications of this knowledge for teaching. Gregorcic and Pendrill (2023) documented errors in chatbot reasoning on a mechanics problem and the difficulty of detecting them when the answer is presented fluently. The second line designs learning activities around chatbot output. Bitzenbauer (2023) proposed classroom activities in which secondary students analyse and critique chatbot answers, and reported changes in how 53 students perceived the tool after the activities. Both lines point to verification as the central requirement for learning with AI in physics. At the same time, the use of AI in schools raises ethical issues, including data privacy, bias and student dependence on the tool, which need to be addressed explicitly in K-12 settings (Akgun & Greenhow, 2021).

2.3. Self-assessment and reported practice

Self-report questionnaires are the most common instruments for measuring AI literacy. Ng et al. (2024) developed and validated a questionnaire covering affective, behavioural, cognitive and ethical components of AI literacy, which allows data collection with large samples and comparison between groups. The limitation of such instruments is that they capture students' perceptions of their own competence. In a study of 280 junior high school students, Porat et al. (2018) found that only a few perceived digital literacy skills were related to actual performance, and that students generally overestimated their competence. If a similar pattern applies to AI competence, students' agreement with statements such as checking AI answers against physical laws may not correspond to what they do when completing assignments. Comparing self-assessment items with items on reported practice within the same survey offers a first, limited indication of whether such a divergence exists.

2.4. AI-integrated instruction and preparation before class

Flipped learning moves the initial encounter with new content outside the classroom and uses class time for discussion, problem solving and experimentation. A meta-analysis of 198 studies across disciplines and educational levels reported a moderate effect of flipped instruction on student performance compared with traditional instruction, with variation across disciplines and implementations (Strelan et al., 2020). A review of the advantages and challenges of the model identified the non-completion of pre-class tasks, the workload of teachers and unequal access to devices as frequently reported difficulties (Akçayır & Akçayır, 2018). López-Villanueva et al. (2024) discussed how AI tools can support the pre-class phase, for example by generating materials and offering support adapted to learners. These studies indicate that the feasibility of AI-integrated flipped instruction depends on students' access to devices, their readiness to prepare before class and the rules that govern their use of AI, which are the conditions examined under RQ3.

3. METHOD

The study employed a cross-sectional survey design. Data were collected at Bao Yen No.2 High School, Lao Cai Province, Vietnam. Participants were students in Grades 10, 11 and 12 who took part voluntarily. The questionnaire was distributed online through Google Forms. Its opening section described the purpose of the study and stated that no identifying information would be collected, that participation was voluntary and could be discontinued at any time, and that responses would not be used to grade students or reported to teachers at the individual level. Students gave consent by selecting an agreement option before accessing the items. The questionnaire was administered in Vietnamese and took approximately 10 minutes to complete. Of the 492 responses received, 12 contained no answers to the main items and were excluded, leaving 480 valid responses. Grade 10 students, the target group of the larger project, accounted for 38.5% of the sample, and female students for 53.3% (Table 1).

Table 1: Distribution of participants by grade and gender (N = 480)

Grade	Male	Female	Not specified	Total	%
Grade 10	95	89	1	185	38.5
Grade 11	66	70	3	139	29.0
Grade 12	59	97	0	156	32.5
Total (%)	220 (45.8)	256 (53.3)	4 (0.8)	480	100.0

The questionnaire was developed from the four-dimension structure presented in Section 2.1. It contained 18 five-point Likert items in three groups and one optional open-ended question, which is not analysed in this report. Group A (items A1–A6) measured the frequency of AI-related practices in physics learning on a scale from 1 (never) to 5 (very often). Four items addressed learning purposes: previewing lesson content (A1), obtaining solutions to exercises (A2), asking about experimental set-up or procedure (A3), and processing data or plotting graphs from experiments (A4). Item A5 asked how often students submitted AI-

generated output to the teacher without editing it, and item A6 how often students received teacher guidance on using AI in physics lessons. Group B (B1–B6) measured self-assessed competence in working with AI tools on a scale from 1 (strongly disagree) to 5 (strongly agree). Items B1 and B2 corresponded to dimension D1, B3 to D2, B4 and B5 to D3, and B6 to D4. Group C (C1–C6) used the same agreement scale to measure conditions and preferences related to AI-integrated instruction: access to devices and connectivity for home study (C1), readiness to complete pre-class tasks (C2), preference for using class time for discussion and experiments (C3), the wish for guidance on checking AI answers (C4), the need for clear rules on AI use (C5), and willingness to submit prompts and AI responses together with learning products (C6).

Internal consistency was estimated with Cronbach's α and McDonald's ω . Values were .849 and .851 for Group A, .914 and .915 for Group B, and .913 and .913 for Group C, all above the conventional threshold of .70 (Nunnally & Bernstein, 1994). ω was reported alongside α because it does not assume equal factor loadings across items (Hayes & Coutts, 2020). For Group A, ω increased to .861 when A5 was removed, consistent with the fact that A5 describes a practice with a direction opposite to the other items. A5 was therefore analysed separately and excluded from the Group A composite score.

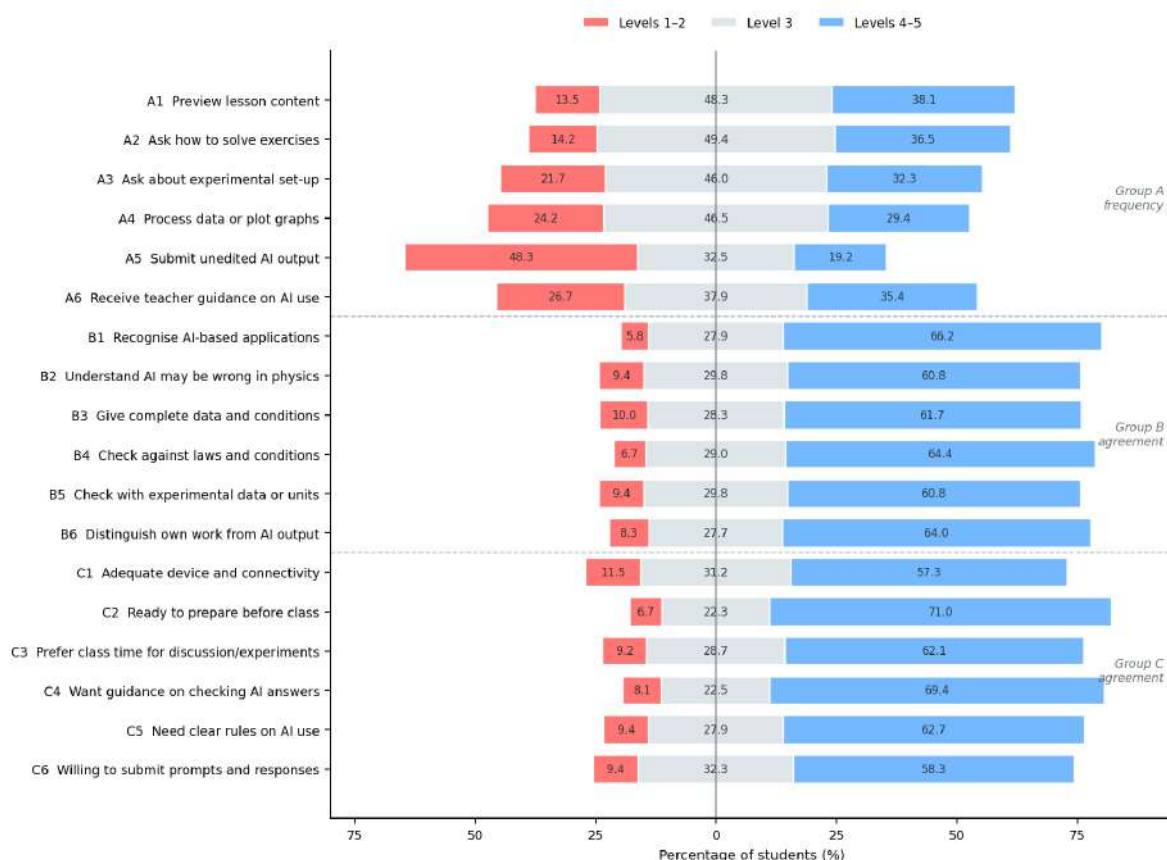
Data were coded and analysed in JASP Ver 0.96. For each item, the percentages of responses at Levels 1–2, Level 3 and Levels 4–5 were computed, together with the mean (M) and standard deviation (SD). The five response levels were grouped into three categories for presentation, since the two lowest and the two highest levels express the same direction of response. Means and standard deviations were reported alongside the percentages to allow comparison with other studies, with the recognition that Likert responses are ordinal and that the percentage distributions provide the primary description. Mean scores were interpreted using five intervals of width 0.8 corresponding to the five scale levels: 1.00–1.80, 1.81–2.60, 2.61–3.40, 3.41–4.20 and 4.21–5.00. To address the second part of RQ2, the distributions of self-assessment items in Group B were compared with the distributions of conceptually related practice items in Group A, specifically B6 with A5 and B5 with A4. This comparison was made at the level of item distributions rather than individual students, and its interpretation is limited accordingly.

4. RESULTS

Figure 2 presents the distribution of responses for all 18 items. The results are reported below by research question.

Figure 2

Distribution of responses to the 18 questionnaire items (N = 480)



Note. For Group A, Levels 1–2 = never or rarely, Level 3 = sometimes, Levels 4–5 = often or very often. For Groups B and C, Levels 1–2 = strongly disagree or disagree, Level 3 = neutral, Levels 4–5 = agree or strongly agree. Values are percentages of students.

4.1. Extent and purposes of AI use in physics learning (RQ1)

Table 2 reports the frequency of AI-related practices. The mean scores of A1–A4 and A6 ranged from 3.02 to 3.24, within the interval corresponding to “sometimes”, and the composite score of Group A excluding A5 was 3.13 ($SD = 0.74$). The two most frequent purposes were previewing lesson content (A1) and obtaining solutions to exercises (A2), both with $M = 3.24$; 38.1% and 36.5% of students, respectively, reported these practices often or very often. Purposes related to experimental work were less frequent: asking about experimental set-up (A3; $M = 3.10$) and processing data or plotting graphs (A4; $M = 3.02$), with 21.7% and 24.2% of students, respectively, reporting that they never or rarely did so. Students therefore used AI tools mainly for acquiring knowledge and solving exercises and less for handling experimental data.

Table 2: Frequency of AI-related practices in physics learning (N = 480)

Item	Content	Levels 1–2 (%)	Level 3 (%)	Levels 4–5 (%)	M	SD
A1	Using AI to preview lesson content	13.5	48.3	38.1	3.24	0.85
A2	Using AI to obtain solutions to exercises	14.2	49.4	36.5	3.24	0.83
A3	Using AI to ask about experimental set-up or procedure	21.7	46.0	32.3	3.10	0.93
A4	Using AI to process data or plot graphs from experiments	24.2	46.5	29.4	3.02	0.94
A5	Submitting AI-generated output without editing	48.3	32.5	19.2	2.48	1.12
A6	Receiving teacher guidance on AI use in physics lessons	26.7	37.9	35.4	3.03	1.06
Group A composite (excluding A5)					3.13	0.74

Note. Levels 1–2 = never or rarely; Level 3 = sometimes; Levels 4–5 = often or very often. Reliability: $\alpha = .849$ and $\omega = .851$ for all six items; $\omega = .861$ without A5.

Two further practices are relevant to the instructional context. Item A6 showed the second-largest variation in Group A after A5 ($SD = 1.06$): 26.7% of students reported that they never or rarely received teacher guidance on using AI, whereas 35.4% reported receiving it often or very often. Item A5 had a mean of 2.48, in the interval corresponding to “rarely”; however, 51.7% of students reported submitting unedited AI output at least sometimes, including 19.2% who did so often or very often.

4.2. Self-assessed competence and its correspondence with reported practice (RQ2)

Table 3 reports students’ self-assessments. The mean scores of B1–B6 ranged from 3.58 to 3.70, within the interval corresponding to agreement, and the composite score was 3.64 ($SD = 0.73$). The proportion of students who agreed or strongly agreed ranged from 60.8% to 66.2%, and the proportion who selected the neutral option ranged from 27.7% to 29.8%. The highest means were observed for recognising AI-based applications (B1; $M = 3.70$) and checking AI answers against laws, formulas and conditions of application (B4; $M = 3.69$). The lowest means were observed for formulating requests with complete data and conditions (B3) and checking AI results with experimental data or units (B5), both with $M = 3.58$. Differences between the dimensions were small: items associated with D1 had means of 3.70 and 3.63, the item associated with D2 had a mean of 3.58, items associated with D3 had means of 3.69 and 3.58, and the item associated with D4 had a mean of 3.65.

Table 3: Self-assessed competence in working with AI tools (N = 480)

Item	Content	Levels 1–2 (%)	Level 3 (%)	Levels 4–5 (%)	M	SD
B1	Recognising which applications in use are based on AI (D1)	5.8	27.9	66.2	3.70	0.81
B2	Understanding that AI answers may be wrong in physics content despite coherent presentation (D1)	9.4	29.8	60.8	3.63	0.92
B3	Providing complete data, conditions and the desired form of the result when asking AI (D2)	10.0	28.3	61.7	3.58	0.91

Item	Content	Levels 1–2 (%)	Level 3 (%)	Levels 4–5 (%)	M	SD
B4	Checking AI answers against laws, formulas and conditions of application (D3)	6.7	29.0	64.4	3.69	0.86
B5	Using experimental data or unit analysis to check AI results (D3)	9.4	29.8	60.8	3.58	0.88
B6	Distinguishing one's own work from AI-generated content when submitting (D4)	8.3	27.7	64.0	3.65	0.88
Group B composite					3.64	0.73

Note. Levels 1–2 = strongly disagree or disagree; Level 3 = neutral; Levels 4–5 = agree or strongly agree. Reliability: $\alpha = .914$, $\omega = .915$.

Comparing the distributions of self-assessment items with those of related practice items revealed two discrepancies. First, 64.0% of students agreed that they distinguished their own work from AI-generated content when submitting assignments (B6), while 51.7% reported submitting unedited AI output at least sometimes (A5). Second, 60.8% of students agreed that they used experimental data or unit analysis to check AI results (B5), while processing experimental data was the least frequent AI-related purpose (A4; $M = 3.02$), with 24.2% of students never or rarely using AI for it. Because the comparison was made between item distributions, these figures do not indicate how many individual students held both positions; they show that the level of self-assessed competence was higher than the level suggested by the reported practices.

4.3. Conditions and preferences regarding AI-integrated instruction (RQ3)

Table 4 reports conditions and preferences. The mean scores of C1–C6 ranged from 3.52 to 3.77, within the interval corresponding to agreement, and the composite score was 3.65 ($SD = 0.75$). The highest means were observed for readiness to complete pre-class tasks (C2; $M = 3.77$; 71.0% agreement) and the wish for guidance on checking AI answers (C4; $M = 3.73$; 69.4% agreement). More than 62% of students agreed that class time should be used for discussion and experiments (C3; 62.1%) and that clear rules on AI use are needed (C5; 62.7%). Regarding the submission of prompts and AI responses together with learning products (C6), 58.3% agreed and 32.3% were neutral. The lowest mean concerned access to devices and connectivity for home study (C1; $M = 3.52$): 11.5% of students disagreed that they had adequate conditions, and 31.2% were neutral.

Table 4: Conditions and preferences regarding AI-integrated instruction (N = 480)

Item	Content	Levels 1–2 (%)	Level 3 (%)	Levels 4–5 (%)	M	SD
C1	Having adequate devices and connectivity for home study with digital materials	11.5	31.2	57.3	3.52	0.94
C2	Being ready to study materials and complete tasks before class	6.7	22.3	71.0	3.77	0.85
C3	Preferring class time for discussion and experiments	9.2	28.7	62.1	3.66	0.92
C4	Wishing to be guided on checking the correctness of AI answers	8.1	22.5	69.4	3.73	0.88
C5	Needing clear rules on the extent of AI use in assignments	9.4	27.9	62.7	3.63	0.90
C6	Being willing to submit prompts and AI responses with learning products	9.4	32.3	58.3	3.57	0.86
Group C composite					3.65	0.75

Note. Levels 1–2 = strongly disagree or disagree; Level 3 = neutral; Levels 4–5 = agree or strongly agree. Reliability: $\alpha = .913$, $\omega = .913$.

5. DISCUSSION

This study examined the extent and purposes of AI use in physics learning, students' self-assessed competence in working with AI tools, and the conditions and preferences relevant to AI-integrated instruction among 480 upper-secondary students in Vietnam.

5.1. AI use concentrated on knowledge acquisition and exercises

In response to RQ1, students used AI tools occasionally, and mainly to preview lessons and obtain solutions to exercises. Uses connected with experimental work, including asking about experimental set-up and processing data, were less frequent. This pattern matters for physics learning because the situations in which students most often consult AI are those in which a fluent but incorrect answer is least likely to be challenged. When a student asks for a solution to an exercise, there is often no independent reference against which to compare the answer except the student's own understanding, which is precisely what the exercise is meant to develop. The errors documented in chatbot solutions to mechanics problems (Gregorcic & Pendrill, 2023) and the dependence of output quality on prompt formulation (Polverini & Gregorcic, 2024) suggest that this form of use carries a risk of accepting incorrect reasoning. Conversely, experimental data offer an external reference for checking AI output, but students rarely linked AI use to such data. The variation in teacher guidance (A6) indicates that the extent to which students are supported in using AI productively currently depends on individual teachers or classes.

5.2. Self-assessment higher than reported practice

In response to RQ2, students assessed their competence at the agreement level, and differences between the four dimensions were small. About three in ten students chose the neutral option on each item, which may reflect uncertainty about their own practices or limited experience with the tasks described. The comparison with reported practice indicated that self-assessed competence was higher than practice would suggest, particularly for distinguishing one's own work from AI output (D4) and for verifying AI results with data (D3). This finding is consistent with research showing that students tend to overestimate their digital competence relative to their actual performance (Porat et al., 2018). Two interpretations are possible. Students may regard the ability to verify or to distinguish their own work as something they could do, while not doing it in practice; alternatively, they may interpret "distinguishing" broadly, for example as knowing privately which parts came from AI without marking them in the submitted work. In either case, the result supports the view that self-report questionnaires such as that of Ng et al. (2024), while useful for large-scale description, are insufficient for assessing AI competence in a subject-specific sense. Evidence from learning products, such as records of prompts and AI responses and the verification that follows, would provide a more direct basis for judging D3 and D4.

5.3. Conditions for AI-integrated flipped instruction

In response to RQ3, most students reported readiness to prepare before class and a preference for using class time for discussion and experiments. These responses are favourable to flipped learning, given that non-completion of pre-class tasks is among the most frequently reported difficulties of the model (Akçayır & Akçayır, 2018). Students also expressed a wish to be taught how to check AI answers and a need for clear rules on AI use, indicating that guidance and rules would be accepted rather than resisted. Willingness to submit prompts and AI responses was somewhat lower, with about one third of students neutral, which suggests that disclosure requirements should be introduced with explanations of their purpose. The 11.5% of students reporting inadequate devices or connectivity at home, and the 31.2% who were neutral on this item, indicate that a pre-class phase relying entirely on home access would exclude part of the class. Designs that combine AI tools with flipped learning (López-Villanueva et al., 2024) therefore need alternatives such as printable materials or access to school devices, in line with the observation that the effects of flipped instruction vary with how it is implemented (Strelan et al., 2020).

5.4. Implications for physics teaching

Three implications follow. First, physics lessons that allow AI use should include explicit verification tasks in which AI output is treated as an object to be checked against laws, conditions of application, unit analysis and experimental data, as in the activities proposed by Bitzenbauer (2023). Linking AI use to experimental work would shift it from the purposes where verification is least likely to those where an external reference is available. Second, disclosure of prompts and AI responses should be built into assignments, both to support responsible use (Akgun & Greenhow, 2021) and to create evidence for assessment. Third, the assessment of students' AI competence should draw on learning products and on students' explanations of their decisions, rather than on self-assessment alone.

5.5. Limitations

The sample was drawn from one school in one province through voluntary participation, so the findings cannot be generalised to all Vietnamese upper-secondary students. All data were self-reported; item A5 in particular may be affected by social desirability, which would lead to underestimation of unedited submission. The comparison between self-assessment and reported practice was made at the level of item distributions and does not show the extent of divergence for individual students. Group B contained one or two items per dimension, which does not allow the dimensional structure to be tested and limits conclusions about differences between dimensions. Finally, the cross-sectional design describes the situation at one point in time and does not permit causal interpretation.

6. CONCLUSION

Upper-secondary students in this study used AI tools in physics learning occasionally, mainly to preview lessons and obtain solutions to exercises, and seldom in connection with experimental data. They assessed their competence in working with AI at the agreement level, but the distributions of their reported practices, particularly the submission of unedited AI output and the limited use of AI with experimental data, indicate that self-assessed competence exceeded what their practices suggested. Most students were ready to prepare before class and wished to be taught how to verify AI answers, while a minority lacked adequate conditions for home study. For physics teaching, these results support instructional designs that make verification of AI output with physical principles and data an explicit learning task, that require disclosure of AI use, and that provide alternatives for students without home access. Future research should assess AI competence through learning products and interviews, and examine individual-level relationships between self-assessment and practice; the larger project of which this survey is a part pursues this direction in the context of Grade 10 Dynamics taught through flipped learning.

REFERENCES

1. Akçayır, G., & Akçayır, M. (2018). The flipped classroom: A review of its advantages and challenges. *Computers & Education*, 126, 334–345. <https://doi.org/10.1016/j.compedu.2018.07.021>
2. Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2(4), 431–440. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8455229/>
3. Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, 15(3), Article ep430. <https://doi.org/10.30935/cedtech/13176>
4. Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, 10, Article 29. <https://doi.org/10.1186/s40594-023-00418-7>
5. Gregorcic, B., & Pendrill, A.-M. (2023). ChatGPT and the frustrated Socrates. *Physics Education*, 58(3), Article 035021. <https://doi.org/10.1088/1361-6552/acc299>
6. Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
7. Lê, A. V., Đỗ, Đ. L., Nguyễn, H. N., Lê, Q. Q., Bùi, T. H., Nguyễn, M. H., Vũ, V. L., Lê, A. L., Vũ, K. C., & Lương, Đ. H. (2025). Khung năng lực trí tuệ nhân tạo dành cho học sinh phổ thông Việt Nam [An artificial intelligence competency framework for Vietnamese general education students]. *Tạp chí Khoa học Giáo dục Việt Nam*, 21(7), 1–8. <http://vjes.vnies.edu.vn/vi/khung-nang-luc-tri-tue-nhan-tao-danh-cho-hoc-sinh-pho-thong-viet-nam>
8. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. <https://doi.org/10.1145/3313831.3376727>
9. López-Villanueva, D., Santiago, R., & Palau, R. (2024). Flipped learning and artificial intelligence. *Electronics*, 13(17), Article 3424. <https://doi.org/10.3390/electronics13173424>
10. Miao, F., & Shiohira, K. (2024). *AI competency framework for students*. UNESCO. <https://doi.org/10.54675/JKJB9835>
11. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, Article 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
12. Ng, D. T. K., Luo, W., Chan, H. M. Y., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, 55(3), 1082–1104. <https://doi.org/10.1111/bjet.13411>
13. Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
14. Polverini, G., & Gregorcic, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics*, 45(2), Article 025701. <https://doi.org/10.1088/1361-6404/ad1420>
15. Porat, E., Blau, I., & Barak, A. (2018). Measuring digital literacies: Junior high-school students' perceived competencies versus actual performance. *Computers & Education*, 126, 23–36. <https://doi.org/10.1016/j.compedu.2018.06.030>
16. Strelan, P., Osborn, A., & Palmer, E. (2020). The flipped classroom: A meta-analysis of effects on student performance across disciplines and education levels. *Educational Research Review*, 30, Article 100314. <https://doi.org/10.1016/j.edurev.2020.100314>