OPEN ACCESS
Jurnal Ilmiah Ilmu Terapan Universitas Jambi
Vol. 10 No. 4. August 2026, pp. 1878-1889 DOI. 10.22437/jiituj.v10i4.58071
Research Article

GENERATIVE AI REDUCES PERIPHERAL BUT NOT MENTAL COGNITIVE LOAD IN PROGRAMMING EDUCATION

Nurkasym Arkabaev[1,*], Roza Oichueva[1], Aizada Ajibekova[1]

[1] Osh State University, Osh, Kyrgyz Republic
Corresponding author email: nurkasym@gmail.com

Article Info
Received: Jul 01, 2026
Revised: Jul 24, 2026
Accepted: Aug 19, 2026
OnlineVersion: Aug 30, 2026

Abstract
Generative AI tools have entered programming classrooms faster than universities have learned to manage them, and evidence from resource-constrained education systems remains scarce. This study examined how engineering students use ChatGPT when learning programming and whether this use is associated with changes in perceived cognitive load. A quantitative cross-sectional survey design was applied: 140 undergraduate engineering students completed a three-block questionnaire combining a demographic and usage profile, a modified NASA-TLX cognitive load scale, and a perceived usefulness scale adapted from the Technology Acceptance Model. The results show near-universal adoption: 94.3% of respondents use ChatGPT for learning, mainly for code debugging and explanations of theoretical concepts. Compared with conventional study, working with ChatGPT was rated significantly lower on five of six NASA-TLX subscales (physical demand, temporal demand, frustration, perceived difficulty, and effort), whereas mental demand did not differ. Usefulness ratings were moderate, risk awareness was low, and usage intensity declined from the second to the fourth year of study (H=7.78; p=0.021). The novelty of this study lies in providing the first field evidence from Central Asia that self-directed generative AI use selectively removes the peripheral load of programming work while leaving its intellectual core intact. The main limitations are the single-institution sample and the use of self-reported measures. The findings support scaffolded integration of AI tutors and compulsory modules on verifying AI-generated code within AI literacy provision.

Keywords: ChatGPT, Cognitive Load, Generative AI, Higher Education, Programming Education.

© 2026 by the author(s)
This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

INTRODUCTION

The release of ChatGPT in November 2022 compressed what would normally be a decade of technology adoption into two years. A global survey covering more than 23,000 students in 109 countries found that ChatGPT had already become a routine study aid, used mainly for brainstorming, summarising texts, and searching for academic sources (Ravšelj et al., 2025). National surveys point in the same direction: roughly two thirds of students in Germany work with AI-based tools (von Garrel & Mayer,

2023), and students in Hong Kong combine broadly positive attitudes with concerns about accuracy and their own development (Chan & Hu, 2023). Acceptance studies consistently identify habit, performance expectancy, and perceived usefulness as the strongest predictors of continued use (Zou & Huang, 2023; Shahzad et al., 2024; Strzelecki, 2024). Programming is a special case within this picture. Because large language models can generate, explain, and repair code in most mainstream languages, they intersect with almost every activity in a programming course (Ouh et al., 2023; Stoyanova et al., 2025).

Evidence from computing education research shows how deep this intersection runs. OpenAI Codex, the model behind early versions of GitHub Copilot, outperformed the average student on typical first-year exam problems and remained competitive on more advanced material (Finnie-Ansley et al., 2022; Finnie-Ansley et al., 2023). Reviews of the field describe a discipline rethinking its assessments, assignments, and learning objectives in response (Becker et al., 2023; Rahman & Watanobe, 2023; Denny et al., 2024). The benefits are real: generative tools adapt explanations to a student's level (Lombarts et al., 2025; Sevilla et al., 2025), give immediate feedback on errors (Lepp & Kaimre, 2025), and a meta-analysis of experimental studies reports an overall positive effect on academic achievement (Sun & Zhou, 2024). Broader reviews reach similarly encouraging, if cautious, conclusions about learning support and personalisation (Grassini, 2023; Lo, 2023).

The experimental record, however, is far from uniformly positive. In a large field experiment, school students who practised with an unrestricted ChatGPT tutor solved more practice problems but then performed worse on an unassisted exam than peers who never had access (Bastani et al., 2024). High school students in a quasi-experimental study rated the actual help received from ChatGPT below their expectations (Yang et al., 2025). A controlled study with novice programmers found that access to a code generator improved code-authoring performance without harming later code-modification tasks, although outcomes depended strongly on how learners used the tool (Kazemitabaar et al., 2023). Self-regulation appears to be the decisive moderator of such effects (Jin et al., 2025), and model accuracy itself degrades as algorithmic complexity grows (Elhambakhsh, 2025). Jose et al., (2025) capture this tension as a cognitive paradox: the same tool can amplify or erode cognition depending on whether learners delegate peripheral work or core thinking. The concern echoes an older line of research on cognitive offloading, which shows that people readily externalise memory and reasoning whenever a tool is at hand (Risko & Gilbert, 2016).

Cognitive Load Theory (CLT) offers the most direct framework for separating these two possibilities. Since Sweller's (1988) original formulation and its systematisation in instructional design (Sweller et al., 1998; Sweller, 2011), the theory has distinguished intrinsic load, driven by the complexity of the material; extraneous load, produced by the way the material is presented; and germane load, the productive effort of building cognitive schemas. Later work reframes the practical task of instruction as load management, that is, replacing unproductive load with productive load (Paas & Sweller, 2012; Sweller et al., 2019; Paas & van Merriënboer, 2020). Digital environments complicate this management because interface features can both add and disguise extraneous load (Skulmowski & Xu, 2022), and recent syntheses argue that AI systems could, in principle, manage load adaptively and in real time (Gkintoni et al., 2025). The Load Reduction Instruction model takes the idea furthest, proposing that AI-supported scaffolding should fade as competence grows (Martin et al., 2025). Whether everyday, self-directed ChatGPT use actually behaves this way, trimming extraneous load while sparing germane processing, is the kind of question that requires field data rather than laboratory tasks.

Central Asia gives that question particular urgency. Kazakhstan plans to train one million citizens in AI-related skills by 2030 and has released the Kazakh-language model KazLLM; research from Kazakhstani universities already examines AI-driven personalised learning pathways (Tapalova & Zhiyenbayeva, 2022). Uzbekistan has adopted a national AI strategy and opened dedicated research laboratories. The Kyrgyz Republic has moved along the same path with fewer resources: a National Council for AI Development has been established (Resolution of the Cabinet of Ministers of the Kyrgyz Republic, 2025), an association of AI developers operates, and leading universities offer specialised courses. A UNESCO regional symposium held in Bishkek in October 2025 placed ethical and human-centred AI in education at the top of the regional agenda (UNESCO, 2025). At the same time, the conditions under which students meet these technologies differ sharply from those in high-income systems: quality learning materials in Kyrgyz and Russian are scarce, internet access is uneven outside major cities, and AI literacy provision is only beginning to appear. Prior work at Osh State University illustrates both sides of this situation. Emerging technologies such as virtual reality can strengthen inclusive instruction, yet infrastructural and methodological barriers remain substantial (Arkabaev et al.,

2025), while a piloted conceptual model for preparing engineering students to apply machine learning produced large gains in knowledge, practical skills, and motivation (Arkabaev, 2026). Adaptive personalisation has also been tested there directly in Python instruction: a gamified system that used reinforcement learning and collaborative filtering to adjust task difficulty and reward mechanics to individual learner profiles raised intrinsic motivation and code quality and shortened the time needed to master basic concepts (Arkabaev et al., 2026). In such an environment, ChatGPT is not one convenience among many; for many students it substitutes for missing textbooks, missing tutors, and missing documentation in their working languages.

Against this background, three gaps motivate the present study. First, the empirical record on generative AI in programming education is dominated by controlled experiments and by samples from North America, Europe, and East Asia, while studies of routine, self-initiated use under authentic study conditions remain rare. Second, cognitive load in AI-assisted programming is usually inferred from performance or discussed theoretically; direct measurement with a standardised instrument such as NASA-TLX (Hart & Staveland, 1988) in a field setting is almost absent. Third, no published survey data describe how students in Central Asian universities actually use these tools, even though universities in the region are drafting AI policies right now and international AI literacy frameworks assume baseline evidence that does not yet exist locally (Long & Magerko, 2020; Ng et al., 2021; Laupichler et al., 2022). The urgency is practical: without local data, institutional rules risk either banning a tool that students treat as infrastructure or endorsing it without safeguards.

The purpose of this study is therefore to describe and analyse how engineering students use ChatGPT when learning programming and to test whether this use is associated with changes in perceived cognitive load. Four research questions guide the analysis: (RQ1) How prevalent is ChatGPT use among engineering students, and which learning tasks does it serve? (RQ2) How does perceived cognitive load when working with ChatGPT compare with load under conventional study across the NASA-TLX subscales? (RQ3) How do students evaluate the benefits and risks of ChatGPT for learning programming? (RQ4) Does the intensity of use differ across years of study? Answering these questions provides the first survey-based evidence from the Kyrgyz Republic and a baseline for designing pedagogical support for AI tutors in the region.

RESEARCH METHOD

This study employed a quantitative, non-experimental, cross-sectional descriptive survey design. The design fits the research questions: RQ1, RQ3, and RQ4 require a structured description of behaviour and attitudes in a defined population, while RQ2 requires a comparison of two sets of standardised self-ratings collected under identical conditions. No experimental manipulation was involved, and all variables were measured as they naturally occur.

The target population comprised undergraduate students of engineering specialities at the Faculty of Mathematics, Physics, and Information Technologies of Osh State University (Kyrgyz Republic). A purposive sampling technique was applied: eligible participants were Year 2 to Year 4 students who had completed at least one programming course, because the questionnaire asked them to compare AI-assisted and conventional programming work, which requires experience of both. The final sample included 140 students from the following programmes: Information Systems and Technologies (46.4%), Computer Science and Computational Engineering (23.6%), Software Engineering (17.1%), Information Security (10.7%), and related programmes (2.1%). The sample contained 66 male (47.1%) and 74 female (52.9%) students; 55 were in Year 2 (39.3%), 31 in Year 3 (22.1%), and 54 in Year 4 (38.6%). The mean age was 20.1 years (SD=1.8). Python was the primary programming language for 66 respondents (47.1%), followed by JavaScript and web technologies (35; 25.0%), C/C++ (20; 14.3%), and other languages (19; 13.6%).

The questionnaire consisted of 32 items organised in three blocks; its structure is summarised in Table 1. Block I recorded demographic characteristics and patterns of ChatGPT use and was developed by the authors. Block II measured subjective cognitive load with the six subscales of the NASA-TLX instrument (Hart & Staveland, 1988) adapted for educational use: each subscale was rated twice, once for programming work with ChatGPT and once for conventional work without it, on a 7-point scale, and one summary item asked about the overall change in academic workload. Block III measured perceived usefulness and perceived risks with items adapted from the Perceived Usefulness scale of the Technology Acceptance Model (Davis, 1989) on a 5-point Likert scale, together with open-ended questions on motives of use. The questionnaire was administered in Russian with Kyrgyz-language support for key

terms. Internal consistency was high for all multi-item constructs: Cronbach's α=0.856 for the cognitive load block in the ChatGPT condition, α=0.922 in the conventional condition, and α=0.914 for the perceived usefulness block.

Table 1. Instrument grid of the survey questionnaire

Block	Constructs (indicators)	Items	Scale	Source	n
I	Demographic profile; prevalence, frequency, and scenarios of ChatGPT use (usage status, frequency, experience, version, task types, alternative tools)	12	Nominal / ordinal	Authors' design	140
II	Subjective cognitive load in two conditions: mental demand, physical demand, temporal demand, frustration, perceived difficulty, effort; overall workload change	7	7-point Likert	Adapted from NASA-TLX (Hart & Staveland, 1988)	110
III	Perceived usefulness; perceived risks; motives of use; open-ended items	13	5-point Likert; open-ended	Adapted from TAM (Davis, 1989)	96

Data were collected in February and March 2026 through an online questionnaire (Google Forms) distributed via the educational WhatsApp groups of the participating programmes. The three blocks were administered sequentially, and students could stop after completing any block. This explains the declining number of respondents across blocks (140, 110, and 96) and was accepted as a deliberate trade-off that reduced fatigue-related careless responding at the cost of unequal subsample sizes. Sample sizes are reported separately for every analysis.

Quantitative data were analysed with descriptive statistics (means, standard deviations, frequencies) for RQ1 and RQ3. The Shapiro-Wilk test indicated non-normal distributions of the Likert-type variables, so non-parametric methods were used throughout. To answer RQ2, the distributions of subscale ratings in the two conditions were compared with the Mann-Whitney U test; to answer RQ4, differences in usage intensity across years of study were tested with the Kruskal-Wallis H test. Reliability of multi-item constructs was estimated with Cronbach's alpha. All computations were performed in Python 3.11 using the pandas, SciPy, and Matplotlib libraries. The significance threshold was set at $p \leq .05$.

Participation was voluntary and anonymous. Before completing the survey, all participants received information about the purpose of the research, data processing procedures, and confidentiality assurances. No identifying data were collected.

RESULTS AND DISCUSSION

Prevalence and Dominant Use Scenarios

Of the 140 respondents, 132 (94.3%) reported using ChatGPT for educational purposes; only eight students (5.7%) did not use generative AI at all. Among active users, 38 (28.8%) worked with the tool daily, 67 (50.8%) several times a week, 13 (9.8%) about once a week, and 14 (10.6%) less often; thus 79.6% of users interacted with ChatGPT at least several times a week. Most users (79.3%) relied on the free version, and 65% had been using ChatGPT for more than a year. Alongside ChatGPT, 54.3% of respondents also used Google Gemini, 14.3% GitHub Copilot, and 8.6% Claude.

Table 2 and Figure 1 present the distribution of use across learning tasks for the 132 active users. Code debugging led the list: 29.5% of users turned to ChatGPT frequently for this purpose and 76.5% at least sometimes. Explaining theoretical concepts came second (28.0% frequently), followed by exam preparation (25.8%), code generation (25.0%), and code refactoring (21.2%). For every task type the modal answer was “sometimes”, which indicates situational rather than systematic integration: students call on the tool when a specific obstacle appears, not as a constant working environment.

Table 2. Distribution of ChatGPT usage frequency by type of learning task (n=132)

Type of Learning Task	Frequently (%)	Sometimes (%)	Rarely (%)	n
Code debugging	29.5	47.0	18.2	132
Explaining concepts	28.0	40.9	22.7	132

Type of Learning Task	Frequently (%)	Sometimes (%)	Rarely (%)	n
Exam preparation	25.8	39.4	25.8	132
Code generation	25.0	37.1	25.8	132
Code refactoring	21.2	40.2	24.2	132

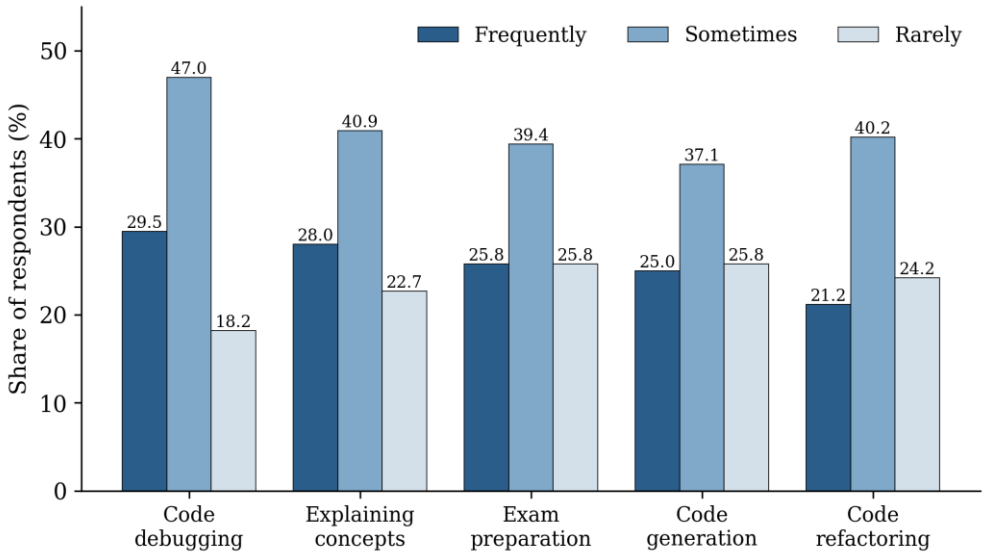

Figure 1. ChatGPT usage frequency by type of learning task (n=132)

Subjective Cognitive Load

To answer RQ2, the ratings of the six NASA-TLX subscales in the two conditions were compared with the Mann-Whitney U test for the 110 respondents who completed Block II (Table 3). Working with ChatGPT received lower mean ratings on all six subscales, and the difference was statistically significant for five of them: physical demand (M=2.63 vs 3.35; p=.004), temporal demand (M=2.48 vs 3.35; p<.001), frustration (M=2.54 vs 3.00; p=.032), perceived difficulty (M=2.59 vs 3.17; p=.010), and effort (M=2.72 vs 3.34; p=.019). Mental demand was the single exception: the two ratings did not differ significantly (M=3.00 vs 3.44; p=.127).

Table 3. Subjective cognitive load ratings on the modified NASA-TLX scale (1-7 scale, n=110)

Parameter	M	SD	Min	Max	U / p
Mental demand (with ChatGPT)	3.00	1.67	1	7	U=5340
Mental demand (without ChatGPT)	3.44	1.92	1	7	p=0.127
Physical demand (with ChatGPT)	2.63	1.61	1	7	U=4731
Physical demand (without ChatGPT)	3.35	1.94	1	7	p=0.004
Temporal demand (with ChatGPT)	2.48	1.57	1	7	U=4475
Temporal demand (without ChatGPT)	3.35	1.97	1	7	p < 0.001
Frustration (with ChatGPT)	2.54	1.57	1	7	U=5063
Frustration (without ChatGPT)	3.00	1.71	1	7	p=0.032
Perceived difficulty (with ChatGPT)	2.59	1.63	1	7	U=4852
Perceived difficulty (without ChatGPT)	3.17	1.78	1	7	p=0.010
Effort level (with ChatGPT)	2.72	1.69	1	7	U=4966
Effort level (without ChatGPT)	3.34	1.99	1	7	p=0.019

The summary item confirms this picture (Figure 2): 70.0% of the 110 respondents reported a decrease in academic workload when using ChatGPT (29.1% a significant decrease and 40.9% a moderate one), 17.3% noticed no change, and 12.7% reported an increase. Read through the lens of CLT, the pattern is selective. The five subscales that dropped describe the supporting envelope of programming work: routine typing and searching, time pressure, emotional strain, and overall exertion. The one subscale that

did not drop, mental demand, describes the intellectual work itself. Effective interaction with ChatGPT still requires formulating a question, judging the relevance of the answer, and integrating it into one's own solution, so the core cognitive engagement is preserved even as the periphery gets lighter.

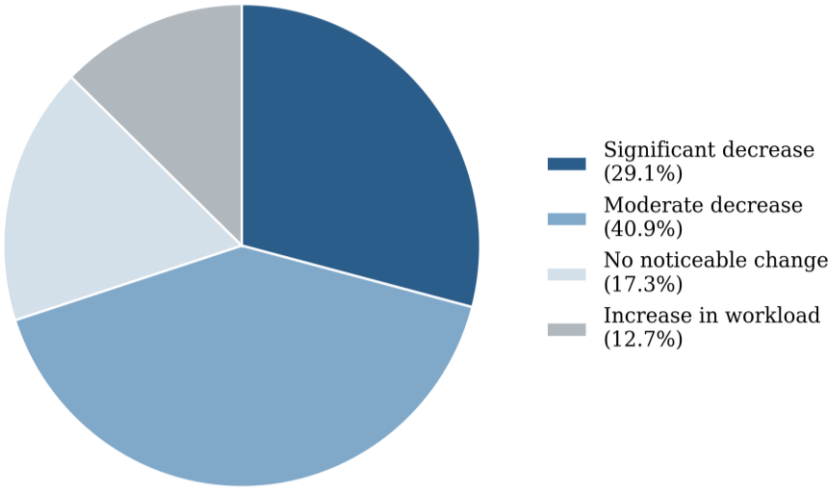

Figure 2. Overall assessment of ChatGPT's impact on academic workload (n=110)

Perceived Benefits and Risks

For the 96 respondents who completed Block III, mean usefulness ratings fell within a narrow moderate band from 2.95 to 3.25 on the 5-point scale (α=0.914). The highest-rated statement concerned the clarity of ChatGPT's explanations relative to official technical documentation (M=3.25; 50.0% agreement), followed by the possibility of learning at one's own pace (M=3.21; 43.8% agreement). The lowest rating went to confidence in code written with ChatGPT's help (M=2.95; 32.3% agreement). Among the risks (Table 4), respondents most often acknowledged that generated answers may be outdated or incorrect (42.7% agreement). At the same time, only 20.8% agreed that verifying the correctness of generated code is difficult, while 37.5% disagreed, and the risk of developing dependence on prompts was accepted by 33.3% of respondents but rejected by 39.6%. Students thus recognise the abstract possibility of errors yet see little difficulty in the concrete task of catching them, a gap that points to overestimated verification competence rather than informed confidence.

Table 4. Perceived benefits and risks of ChatGPT use (n=96)

Statement	Agree (%)	Disagree (%)
Benefits		
ChatGPT accelerates the code debugging process	43.8	28.1
ChatGPT explanations are clearer than documentation	50.0	31.2
24/7 availability of ChatGPT is a significant advantage	39.6	35.4
ChatGPT supports learning at one's own pace	43.8	32.3
Risks		
There is a risk of developing dependence on prompts	33.3	39.6
It is difficult to verify the correctness of generated code	20.8	37.5
ChatGPT may produce outdated or incorrect answers	42.7	32.3

Differentiation by Year of Study

A Kruskal-Wallis test revealed significant heterogeneity of usage intensity across years of study (H=7.78; p=0.021). Mean usage frequency declined monotonically from 3.27 in the second year to 3.00 in the third and 2.69 in the fourth (on a 4-point scale). Senior students appear to rely less on ChatGPT as their repertoire of information sources and independent strategies widens, whereas second-year students, who face the steepest part of the programming learning curve, use it most intensively.

The adoption rate of 94.3% observed here is well above the 66-71% range reported in the global student survey (Ravšelj et al., 2025) and fits the general position of computing students as early adopters

of code-capable AI (Rogers et al., 2024). Two local factors plausibly amplify the effect. Programming tasks give an immediate, testable payoff from a generated answer, and the shortage of quality learning materials in Russian and Kyrgyz turns a general-purpose chatbot into a substitute for documentation and tutoring. High prevalence combined with the dominance of the "sometimes" response across all task types suggests that students have settled into a pragmatic, on-demand mode of use rather than wholesale delegation of their coursework.

The cognitive load results add substance to that interpretation. The profile obtained, in which five peripheral subscales decrease while mental demand stays flat, is consistent with the reduction of extraneous rather than germane load and speaks against the strongest versions of the deskilling argument, at least for routine self-directed use. The same asymmetry has now been observed outside education: knowledge workers using generative AI report less cognitive effort while their role shifts towards verifying and integrating AI output (Lee et al., 2025). This convergence matters because verification is precisely where our respondents look weakest. Only one in five sees any difficulty in checking generated code, although hallucinated and subtly wrong output is a well-documented property of large language models (Kasneci et al., 2023), and dedicated tool support for validating AI-generated code is only beginning to emerge (Ferdowsi et al., 2024). The combination of reduced frustration, unchanged mental demand, and unexamined trust also helps reconcile the conflicting experimental findings discussed in the introduction: the harm documented by Bastani et al. (2024) arises when learners offload the mental core, while the neutral-to-positive outcomes in Kazemitabaar et al. (2023) arise when they offload the periphery. Our respondents describe themselves on the safe side of that line, but self-report cannot fully guarantee it.

Moderate usefulness ratings, clustered around three on a five-point scale, deserve a separate explanation, since attitude studies in other regions often report distinctly warmer evaluations (Chan & Hu, 2023). Part of the answer is linguistic: most respondents interact with ChatGPT in Russian, and the quality of generated explanations and code is known to be higher in English. Part of the answer lies in competence. A growing body of work shows that perceived and actual benefit from AI tools tracks the user's AI and digital literacy: AI literacy predicts educational attainment in online settings (Jingyu et al., 2024) and practice achievement in vocational education (Lin et al., 2025); digital literacy shapes achievement through learning adaptation and self-regulation (Chen, 2025) and buffers the stress of technology-mediated study (Feng & Liu, 2024); and attitudes mediate the path from literacy to learning outcomes (Fan & Zhang, 2024). Conceptual models likewise tie AI-enhanced learning to self-efficacy and critical thinking awareness (Jia & Tu, 2024). Where AI literacy provision is minimal, as in the present context, moderate usefulness scores are exactly what these models predict, and validated instruments now exist for measuring both digital literacy in the AI era (Hwang et al., 2023) and AI literacy proper (Ng et al., 2024), which makes the deficit measurable and addressable.

The monotonic decline of usage intensity across years of study admits two readings that future work should separate. Within CLT it resembles an expertise effect: as schemas consolidate, the marginal value of external support falls, so seniors need the tool less. Alternatively, senior students formed their study habits before ChatGPT existed and simply integrated it less deeply, while current second-year students have never studied programming without it. Motivational evidence showing that AI-based tools raise computational thinking and programming self-efficacy mainly in earlier phases of learning is compatible with both readings (Yilmaz & Karaoglan Yilmaz, 2023).

Generalising from these findings, we propose describing the observed mode of generative AI use as peripheral optimisation: students use ChatGPT to strip away the supporting load of programming work, and by their own account its intellectual core stays with them. The theoretical contribution of the study is to ground this description in standardised load measurement collected in the field rather than in the laboratory, in a resource-constrained, multilingual context that the literature has so far ignored. The practical implications follow directly. First, universities in Kyrgyzstan and the wider region need explicit institutional guidelines for generative AI, and ready policy frameworks exist to adapt (Chan, 2023); their success will depend on instructors, whose trust in educational AI varies widely across countries (Viberg et al., 2024) and who currently navigate competing pressures largely without institutional support (Zuo et al., 2026). Second, a scaffolded model of AI use, in which prompt support fades as competence grows, matches both the Load Reduction Instruction principles (Martin et al., 2025) and the empirical usage decline across years observed here; machine-learning-driven adaptive gamification offers one locally validated mechanism for delivering such individualised support in programming courses (Arkabaev et al., 2026). Third, programming courses should include a compulsory module on evaluating AI output,

covering verification, testing, and code revision, because this is the single competence our data show to be both critical and absent; assessment practice will have to change together with it (Xia et al., 2024), and reviews of AI-enhanced education consistently favour such engagement-centred design over prohibition (Almusaed et al., 2023).

Several limitations bound these conclusions. The sample comes from one institution, so the estimates describe a single, although large, regional university. All measures are self-reported and vulnerable to social desirability bias; in particular, the reassuring stability of mental demand is a perception, not a performance measure. The cross-sectional design supports no causal claims about the relationship between ChatGPT use and cognitive outcomes. The sequential administration of blocks produced unequal subsamples, and the with/without load ratings were collected within a single questionnaire and compared as independent distributions; paired designs would give stronger control. Future research should extend the geography to universities in Kazakhstan, Uzbekistan, and Tajikistan, combine surveys with interviews and interaction log analysis, adopt longitudinal designs with objective learning outcomes, and localise AI literacy instruments for Kyrgyz-speaking students.

CONCLUSION

This study set out to establish how engineering students at a Kyrgyz university use ChatGPT in learning programming and how this use relates to their perceived cognitive load, and the answers to the four research questions form a consistent picture: adoption is near-universal (94.3%), concentrated in code debugging and concept explanation and situational in character (RQ1); working with ChatGPT is rated significantly lower on physical demand, temporal demand, frustration, perceived difficulty, and effort, while mental demand remains unchanged (RQ2); students judge the tool's usefulness as moderate and substantially underestimate its risks, above all the difficulty of verifying generated code (RQ3); and usage intensity declines from the second to the fourth year of study (RQ4). Taken together, the findings support the concept of peripheral optimisation, whereby routine self-directed use of generative AI removes the supporting load of programming work without, by students' own accounts, displacing its intellectual core. The impact of the research is a local evidence base on which universities in Kyrgyzstan and neighbouring countries can build institutional guidelines for generative AI, scaffolded models of AI-assisted learning with gradually fading support, and compulsory modules on verifying AI-generated code within AI literacy provision; as the first survey-based account of generative AI in programming education in Central Asia, the study also sets a baseline against which the effects of such measures can later be assessed.

ACKNOWLEDGMENTS

The authors would like to acknowledge assistance of the Editor and Reviewers in the preparation of the article for publication.

AUTHOR CONTRIBUTIONS

Conceptualization, Nurkasym Arkabaev. Methodology, Nurkasym Arkabaev; Software, Roza Oichueva; Validation, Nurkasym Arkabaev and Roza Oichueva; Formal Analysis, Aizada Ajibekova; Investigation, Nurkasym Arkabaev; Resources, Nurkasym Arkabaev; Data Curation, Nurkasym Arkabaev; Writing – Nurkasym Arkabaev, X.X.; Writing – Review & Editing, Nurkasym Arkabaev; Visualization, Aizada Ajibekova; Supervision, Aizada Ajibekova; Project Administration, Nurkasym Arkabaev.

CONFLICTS OF INTEREST

The author(s) declare no conflict of interest.

USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors declare that no artificial intelligence (AI) tools were used in the generation, analysis, or writing of this manuscript. All aspects of the research, including data collection, interpretation, and manuscript preparation, were carried out entirely by the authors without the assistance of AI-based technologies.

REFERENCES

Almusaed, A., Almssad, A., Yitmen, I., & Homod, R. Z. (2023). Enhancing Student Engagement:

Harnessing "AIED"'s Power in Hybrid Education – A Review Analysis. *Education Sciences*, 13(7), 632. https://doi.org/10.3390/educsci13070632.

Arkabaev, N. K. (2026). A conceptual model for integrating machine learning competencies into technical higher education: Development and pilot validation in a Central Asian context. *Journal of Applied Artificial Intelligence in Education*, 2(2), 191-205. https://doi.org/10.66053/jaaie.v2i2.429..

Arkabaev, N., Isakov, T., Yagoub, I. H., Elziber, H. E., Elsir, S. M., Sulima, M. Y., & Abdelgalal, O. I. A. (2026). Adaptive gamification in Python learning: Integrating machine learning algorithms for educational process personalization. *Journal of Statistics Applications and Probability*, 15(3), 519-542. https://doi.org/10.18576/jsap/150311.

Arkabaev, N., Tran, DV., Chu, VT., Hashimova, S. & Nasirova S. (2025). Use of Virtual Reality for Inclusive Education: Assessing the Availability and Adaptation of Educational Materials. *Premier Journal of Science*, 14, 100148. https://doi.org/10.70389/PJS.100148.

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: evidence from high school mathematics. *The Wharton School Research Paper, Proceedings of the National Academy of Sciences*, 122(26), 10.1073 https://doi.org/10.2139/ssrn.4895486.

Becker, B.A., Denny, P., Finnie-Ansley, J., Luxton-Reilly, A., Prather, J., Santos, E.A. (2023). Programming Is Hard - Or at Least It Used to Be: Educational Opportunities and Challenges of AI Code Generation. SIGCSE 2023: *Proceedings of the 54th ACM Technical Symposium on Computer Science Education*, 1, 500-506. https://doi.org/10.1145/3545945.3569759.

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20:38. https://doi.org/10.1186/s41239-023-00408-3.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20:43, 43. https://doi.org/10.1186/s41239-023-00411-8.

Chen, F. (2025). The relationship between digital literacy and college students' academic achievement: The chain mediating role of learning adaptation and online self-regulated learning. *Frontiers in Psychology*, 16, 1590649. https://doi.org/10.3389/fpsyg.2025.1590649.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, 13(3), 319-340. https://doi.org/10.2307/249008.

Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM*, 67(2), 56-67. https://doi.org/10.1145/3624720.

Elhambakhsh, S. E. (2025). Evaluating ChatGPT-3's efficacy in solving coding tasks: Implications for academic integrity in English language assessments. *Language Testing in Asia*, 15:37. https://doi.org/10.1186/s40468-024-00333-w.

Fan, J., & Zhang, Q. (2024). From literacy to learning: The sequential mediation of attitudes and enjoyment in AI-assisted EFL education. *Heliyon*, 10(17), e37158. https://doi.org/10.1016/j.heliyon.2024.e37158.

Feng, X., & Liu, H. (2024). I feel blue – teacher, can you help me? A study on the effect of digital literacies on language learners' technostress, on-line engagement, autonomy, and academic success. *BMC Psychology*, 12:143. https://doi.org/10.1186/s40359-024-01637-5.

Ferdowsi, K., Huang, R., James, M. B., Polikarpova, N., & Lerner, S. (2024). Validating AI-generated code with live programming. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1-8. https://doi.org/10.1145/3613904.3642495.

Finnie-Ansley, J., Denny, P., Becker, B. A., Luxton-Reilly, A., & Prather, J. (2022). The robots are coming: Exploring the implications of OpenAI Codex on introductory programming. *Proceedings of the 24th Australasian Computing Education Conference*, 10-19. https://doi.org/10.1145/3511861.3511863.

Finnie-Ansley, J., Denny, P., Luxton-Reilly, A., Santos, E. A., Prather, J., & Becker, B. A. (2023). My AI wants to know if this will be on the exam: Testing OpenAI's Codex on CS2 programming exercises. *Proceedings of the 25th Australasian Computing Education Conference*, 97-104. https://doi.org/10.1145/3576123.3576134.

Gkintoni, E., Antonopoulou, H., Sortwell, A., & Halkiopoulos, C. (2025). Challenging cognitive load

theory: The role of educational neuroscience and artificial intelligence in redefining learning efficacy. *Brain Sciences*, 15(2), 203. https://doi.org/10.3390/brainsci15020203.

Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. https://doi.org/10.3390/educsci13070692.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in Psychology*, 52, 139-183. https://doi.org/10.1016/S0166-4115(08)62386-9.

Hwang, H. J., Liu, C., & Qin, C. (2023). Development and validation of a digital literacy scale in the artificial intelligence era for college students. *KSII Transactions on Internet and Information Systems*, 17(8), 2241- 2258. https://doi.org/10.3837/tiis.2023.08.016.

Jia, X.-H., & Tu, J.-C. (2024). Towards a new conceptual model of AI-enhanced learning for college students: The roles of artificial intelligence capabilities, general self-efficacy, learning motivation, and critical thinking awareness. *Systems*, 12(3), 74. https://doi.org/10.3390/systems12030074.

Jin, F., Lin, C.-H., & Lai, C. (2025). Modeling AI-assisted writing: How self-regulated learning influences writing outcomes. *Computers in Human Behavior*, 165, 108538. https://doi.org/10.1016/j.chb.2024.108538.

Jingyu, X., Alibakhshi, G., Zamanpour, A., Zarei, M. A., Sherafat, S., & Behzadpoor, S.-F. (2024). How AI literacy affects students' educational attainment in online learning: Testing a structural equation model in higher education context. *The International Review of Research in Open and Distributed Learning*, 25(3), 179-198. https://doi.org/10.19173/irrodl.v25i3.7720.

Jose B, Cherian J, Verghis AM, Varghise SM, S M and Joseph S (2025) The cognitive paradox of AI in education: between enhancement and erosion. *Frontiers in Psychology*, 16, 1550621. https://doi.org/10.3389/fpsyg.2025.1550621.

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., … Groh, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274.

Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). Studying the effect of AI code generators on supporting novice learners in introductory programming. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1-23. https://doi.org/10.1145/3544548.3580919.

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. *Computers and Education: Artificial Intelligence*, 3, 100101. https://doi.org/10.1016/j.caeai.2022.100101.

Lee, H.-P. H., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1-22. https://doi.org/10.1145/3706598.3713778.

Lepp, M. & Kaimre, J. (2025). Does generative AI help in learning programming: Students' perceptions, reported use and relation to performance. *Computers in Human Behavior Reports*, 18, 100642. https://doi.org/10.1016/j.chbr.2025.100642.

Lin, X., Xu, G., & Xiong, B. (2025). Artificial intelligence literacy, sustainability of digital learning and practice achievement: A study of vocational college students. *PLOS ONE*, 20(10), e0332175. https://doi.org/10.1371/journal.pone.0332175.

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. https://doi.org/10.3390/educsci13040410.

Lombarts, T., Walravens, G., Seraj, M., Ochoa, L. and Brand, M. (2025). Personalization of Programming Education: An NLP-Based Bi-dimensional Classification of Programming Exercises. *Proceedings of the 2025 ACM SIGPLAN International Symposium on SPLASH-E (SPLASH-E '25). Association for Computing Machinery, New York, NY, USA*, 90-101. https://doi.org/10.1145/3758317.3759684.

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-16. https://doi.org/10.1145/3313831.3376727.

Martin, A. J., Collie, R. J., Kennett, R., Liu, D., Ginns, P., Sudimantara, L. B., Dewi, E. W., &

Rüschenpöhler, L. G. (2025). Integrating generative AI and load reduction instruction to individualize and optimize students' learning. *Learning and Individual Differences*, 121, 102723. https://doi.org/10.1016/j.lindif.2025.102723.

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, 100041. https://doi.org/10.1016/j.caeai.2021.100041.

Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & 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.

Ouh, E.L., Gan, B., Shim, K.J., Wlodkowski, S. (2023). ChatGPT, Can You Generate Solutions for My Coding Exercises? An Evaluation on Its Effectiveness in an Undergraduate Java Programming Course. ITiCSE 2023: *Proceedings of the 2023 Conference on Innovation and Technology in Computer Science Education*, 1, 54-60. https://doi.org/10.1145/3587102.3588794.

Paas, F., & Sweller, J. (2012). An evolutionary upgrade of cognitive load theory: Using the human motor system and collaboration to support the learning of complex cognitive tasks. *Educational Psychology Review*, 24(1), 27-45. https://doi.org/10.1007/s10648-011-9179-2.

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, 29(4), 394-398. https://doi.org/10.1177/0963721420922183.

Rahman, M.M. & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13, 9, 5783. https://doi.org/10.3390/app13095783.

Ravšelj, D., Keržič, D., Tomaževič, N., Umek, L., Brezovar, N., et al. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE*, 20(2), e0315011. https://doi.org/10.1371/journal.pone.0315011.

Resolution of the Cabinet of Ministers of the Kyrgyz Republic No. 29 of January 22, 2025, "On Amending the Resolution of the Cabinet of Ministers of the Kyrgyz Republic 'On the National Council for the Development of Artificial Intelligence'." URL: https://cbd.minjust.gov.kg/7-51185/edition/47014/ru.

Rogers, M. P., Hillberg, H. M., & Groves, C. L. (2024). Attitudes towards the use (and misuse) of ChatGPT: A preliminary study. *Proceedings of the 55th ACM Technical Symposium on Computer Science Education*, 1, 1147-1153. https://doi.org/10.1145/3626252.3630784.

Sevilla, A., Cuevas-Ruiz, P., Rello, L. et al. (2026). Artificial Intelligence in Education: Computer-Assisted Learning and AI-guided Tutors. *Italian Economic Journal, 12, 1-38* https://doi.org/10.1007/s40797-025-00354-1.

Shahzad, M. F., Xu, S., & Asif, M. (2024). Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model. *British Educational Research Journal*, 51(2), 489-513. https://doi.org/10.1002/berj.4084.

Skulmowski, A., & Xu, K. M. (2022). Understanding cognitive load in digital and online learning: A new perspective on extraneous cognitive load. *Educational Psychology Review*, 34, 171-196. https://doi.org/10.1007/s10648-021-09624-7.

Stoyanova, D., Stoyanova-Petrova, S., Shotarova, S., Lyubomirov, S., & Mileva, N. (2026). ChatGPT in Programming Education: An Empirical Study on Its Impact on Student Performance, Creativity, and Teamwork. *Education Sciences*, 16(1):19. https://doi.org/10.3390/educsci16010019.

Strzelecki, A. (2024). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 32(9), 5142-5155. https://doi.org/10.1080/10494820.2023.2209881.

Sun, L., & Zhou, L. (2024). Does Generative Artificial Intelligence Improve the Academic Achievement of College Students? A Meta-Analysis. *Journal of Educational Computing Research*, 62(7), 1676-1713. https://doi.org/10.1177/07356331241277937.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257-285. https://doi.org/10.1207/s15516709cog1202_4.

Sweller, J. (2011). CHAPTER TWO - Cognitive Load Theory. *Psychology of Learning and Motivation*, 55, 37-76. https://doi.org/10.1016/B978-0-12-387691-1.00002-8.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review*, 10(3), 251-296. https://doi.org/10.1023/A:1022193728205.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261-292. https://doi.org/10.1007/s10648-019-09465-5.

Tapalova, O., & Zhiyenbayeva, N. (2022). Artificial intelligence in education: AIED for personalised learning pathways. *The Electronic Journal of e-Learning*, 20(5), 639-653. https://doi.org/10.34190/ejel.20.5.2597.

UNESCO. (2025). Empowering the future of digital learning: UNESCO initiated regional dialogue for equitable, ethical and human-centred AI in education. URL: https://www.unesco.org/en/articles/empowering-future-digital-learning-unesco-initiated-regional-dialogue-equitable-ethical-and-human.

Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., Wasson, B., Tømte, C., Spikol, D., Milrad, M., Coelho, R., & Kizilcec, R. F. (2025). What explains teachers' trust in AI in education across six countries? *International Journal of Artificial Intelligence in Education*, 35(3), 1288-1316. https://doi.org/10.1007/s40593-024-00433-x.

von Garrel, J., & Mayer, J. (2023). Artificial intelligence in studies — use of ChatGPT and AI-based tools among students in Germany. *Humanities and Social Sciences Communications*, 10:799. https://doi.org/10.1057/s41599-023-02304-7.

Xia, Q., Weng, X., Ouyang, F., Lin, T., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21:40. https://doi.org/10.1186/s41239-024-00468-z.

Yang, T. C., Hsu, Y. C., & Wu, J. Y. (2025). The effectiveness of ChatGPT in assisting high school students in programming learning: evidence from a quasi-experimental research. *Interactive Learning Environments*, 33(6), 3726-3743. https://doi.org/10.1080/10494820.2025.2450659.

Yilmaz, R., Karaoglan Yilmaz, F.G. (2023). The effect of generative artificial intelligence (AI)-based tool usage on students' computational thinking skills, programming self-efficacy and motivation. *Computers & Education: Artificial Intelligence*, 4, 100147. https://doi.org/10.1016/j.caeai.2023.100147.

Zou, M., & Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology*, 14, 1259531. https://doi.org/10.3389/fpsyg.2023.1259531.

Zuo, Y., Xu, X., & Dunning, L. A. (2026). "Will AI steal my glory?": Power relations perceived by college instructors when grappling with generative AI. *Journal of Applied Learning & Teaching*, 9(1). https://doi.org/10.37074/jalt.2026.9.1.8.