Journal of Information Systems Education

Volume 37
Issue 4
Fall 2026

Unrecommended Use of Generative AI: Decision to Use and Effects on Performance and Learning

Stéphane E. Collignon, A. Graham Peace, Lauren Peace, and Nanda C. Surendra

Recommended Citation: Collignon, S. E., Peace, A. G., Peace, L., & Surendra, N. C. (2026). Unrecommended Use of Generative AI: Decision to Use and Effects on Performance and Learning. *Journal of Information Systems Education*, 37(4), 526-545. https://doi.org/10.62273/SGYZ6748

Article Link: https://jise.org/Volume37/n4/JISE2026v37n4pp526-545.html

Received:	October 23, 2025
First Decision:	December 1, 2025
Accepted:	February 16, 2026
Published:	December 15, 2026

Find archived papers, submission instructions, terms of use, and much more at the JISE website:
https://jise.org

ISSN: 2574-3872 (Online) 1055-3096 (Print)

Journal of Information Systems Education, 37(4), 526-545, Fall 2026
https://doi.org/10.62273/SGYZ6748

Unrecommended Use of Generative AI: Decision to Use and Effects on Performance and Learning

Stéphane E. Collignon
A. Graham Peace
John Chambers College of Business & Economics
West Virginia University
Morgantown, WV 26506, USA
stephane.collignon@mail.wvu.edu, graham.peace@mail.wvu.edu

Lauren Peace
College of Arts and Letters
University of Tampa
Tampa, FL 33606, USA
laur.peace@gmail.com

Nanda C. Surendra
John Chambers College of Business & Economics
West Virginia University
Morgantown, WV 26506, USA
NaSurendra@mail.wvu.edu

ABSTRACT

Student use of generative AI tools, such as ChatGPT, is a growing concern in all aspects of higher education. However, research on the impact of these tools on student learning is currently limited, especially in more technical programs such as information systems. The research that does exist yields mixed results and indicates that the impact on performance may be dependent on the course type. Universities and faculty have been left to struggle with the development of policies and procedures in the face of a technology that is rapidly becoming ubiquitous. This paper documents a study of students in a technical database management systems course, where students were allowed but discouraged to use AI tools to aid in the completion of assignments. Student use of AI was tracked throughout the semester. The results indicate that the use of AI is predicted by performance earlier in the semester. AI users did not perform at lower levels than non-users on technical assignments but did perform at a lower level than their non-AI using colleagues on the final exam, where the tools were unavailable. AI users were also found to perform consistently over time. These results shed light on the effects of unrecommended AI use and indicate that faculty need to carefully incorporate AI tools and policies into their more technical courses.

Keywords: Generative AI, Generative AI in teaching, Structured query language (SQL), Learning improvement, Student learning

1. INTRODUCTION

Generative AI technologies have exploded onto the academic scene over the past few semesters, giving students a new tool that can do everything from creating presentations to writing code for computer classes. Students can't be blamed for using a new application that appears to make their lives significantly easier while also possibly improving their grades, but educators in the information systems (IS) field have to ask: *Is generative AI helping students learn, or just allowing them to more easily move through classes with less retained knowledge?* A search will find many articles on the topic, but the majority of these are simply opinions or editorials in nature (Jiang & Nakatani, 2025).

While the appeal to students is obvious, early academic adopters of ChatGPT, perhaps the best known example of these technologies, have expressed concerns that generative AI may promote superficial learning habits and erode critical thinking skills (Mogavi et al., 2024). A Pew Research Center survey found that four times as many teachers believed that AI tools do more harm than good, although the majority had no opinion (Lin, 2024). Students are generally satisfied with ChatGPT as a tool for aiding in the learning process, but concerns have been raised regarding higher-level learning (Sun & Deng, 2025). As with any new technology, many express opinions, both positive and negative, but research is needed to identify specific benefits and costs of student use, so that informed decisions can be made.

As IS academics, we want our students to leave our institutions with the knowledge and tools necessary to succeed in their chosen careers, and determining how to incorporate new technologies into the curriculum can be difficult. This is not a new occurrence. When graphing calculators were first introduced, there were similar concerns regarding student use and learning, and research yielded mixed results (Penglase & Arnold, 1996). However, the rapid assimilation of new technology into the educational setting has increased the time pressure on the academic community when it comes to assessing the impact. As academic institutions struggle with creating appropriate policies and procedures, faculty are left with the decision to recommend or unrecommend generative AI technology usage to their students, but research is scant on which is the better decision. This study aims to provide some help in this process.

In our institution's typical introduction to database management systems course, we strive to teach our students the fundamentals of relational databases, including entity-relationship diagrams (ERDs), the use of SQL to create and populate databases, and retrieve data from the databases. We have seen a large rise in the use of generative AI, particularly ChatGPT, by students in our courses. While use of these technologies is not penalized in our program, students are warned that they will need to manually draw ERDs and write SQL code in their exams, without the availability of generative AI. The use of these tools is unrecommended, as we strongly feel that the practice and experiential learning gained working unassisted through the assignments is a valuable learning process. Practice makes perfect, as the saying goes. Despite these warnings, many students still choose to use ChatGPT to aid in completing their assignments. This non-generalized use was initially obvious to us because the code produced is different from what we deal with among non-AI users, and generative AI leaves clues in its outputs. However, as AI advances, its usage is becoming more ubiquitous and more difficult to detect. It is clear that AI use by students is here to stay, and we will need to incorporate AI into our educational experience. In order to do so, we need to understand the impact generative AI is making on student learning.

The students' use of AI in our classes provided us with a context of unrecommended AI use, in which we could compare students who made the choice to use AI against those who did not. In this context, our research questions are:

What drives the use of generative AI?

How does generative AI use impact performance and learning?

In the following sections, we present a review of the current literature on the topic, the development of four hypotheses, and the results of data collected in our course on database management systems offered in the IS department of a large state land-grant university. Finally, we present our findings and implications for both future research and current pedagogy. We hope that our findings are useful to those currently dealing with student use of AI in their classes.

2. LITERATURE REVIEW

Concerns are often raised when new technology is introduced in the learning environment, but in many cases the new technology is found to be useful after the initial period of skepticism. Kastberg and Leatham (2005) found that student achievement with graphing calculators was positively impacted when student use of the tool was specifically considered and incorporated in the curriculum. The calculator use allowed for a wider array of problem-solving approaches to be implemented (Kastberg & Leatham, 2005). Similarly, Grammarly has been found to help writers identify passages potentially requiring review (Abu Qub'a et al., 2024). Technology can be used to automate menial tasks, thereby allowing more time for higher-order learning.

Generative AI refers to technologies that can take raw data as input and "learn" to generate probable outputs when prompted, based on the data on which they have been trained. In effect, the AI generates a statistically probable response to the input. Initially, these models focused on numerical data, but newer technologies have allowed the processing of complex data types such as speech and images (IBM, 2025). Generative AI is a tool that can help people to perform tasks, including schoolwork. In the specific world of classes on database management systems, tools such as MySQL Workbench have been developed to automate some tasks and provide graphical interfaces, but these tools still require the student to understand database concepts and the structure of SQL queries. Generative AI, such as ChatGPT, is a completely different level of tool, where a student, for example, can enter a question into the AI and be presented with a fully constructed query. Second, while tools such as Workbench are specific to the class at hand and require some specialized learning, ChatGPT and other generative AI tools are rapidly becoming a ubiquitous part of everyday life. Students interact with these tools for classwork, but beyond the classroom, students are using ChatGPT for everything from breaking up with significant others to the provision of therapy (Anderson, 2025).

Given the relatively short history of the use of generative AI by students, there is limited academic research available regarding its impacts on learning and student performance, although anecdotes and sensational headlines can be easily found in the popular press. One thing is clear: the use of these technologies by students is increasing rapidly. A recent Pew Research Center survey found that 26% of students aged 13 to 17 admitted to using ChatGPT for schoolwork, double the number from two years prior. The usage increased significantly as students moved up through the grades (Sidoti et al., 2025). In response, OpenAI has created an online course for K-12 educators to aid them in determining how best to use ChatGPT in the curriculum.

At the college level, OpenAI's own research finds that more than a third of college-aged young adults in the USA use ChatGPT, with a quarter using the technology for schoolwork (OpenAI, 2025). The report concluded that "many college and university students are teaching themselves and their friends about AI without waiting for their institutions to provide formal AI education or clear policies about the technology's use" (OpenAI, 2025, p.2).

A meta-analysis of research studying ChatGPT specifically found the technology to have a positive impact on improving learning performance, although the impact was moderated by course type, as was the influence of ChatGPT use on the development of higher-order thinking (Wang & Fan, 2025). Abbas et al. (2024) found that ChatGPT usage by college students led to a greater tendency for procrastination and lower academic performance. Also, students were more likely to use ChatGPT when facing higher academic workloads and time pressures. Similarly, Dhulipala and Bhujade (2025) found that AI-enabled exam preparation had a positive impact on both student performance and time management, although the authors caution that excessive dependence on AI tools requires more investigation. In a study using electroencephalography to assess cognitive load during essay writing, Kosmyna et al. (2025) found that students using ChatGPT had less brain activity than those students who did not and were also less able to recall accurate quotes from the essays they had written.

Comments from students who actually use AI provide anecdotal evidence of concern. One student pondered whether or not he would have learned more if he hadn't used AI, comparing it to seeing a tutor, "but what happens when a tutor starts writing lines of your paper for you"? Another likened the use of

ChatGPT to using power tools when building a house, but stated, "It's really hard to judge. Is this my work?" (Walsh, 2025).

The use of ChatGPT is causing issues for students who do their own work. "I can sometimes get frustrated…when I know for a fact I've written my thing from scratch and I've worked really hard to do it and then hear of other people who've got away with just using ChatGPT to write it and at the end of the day, they end up with a way better mark," commented one student (Eccles & Hayton, 2025). Students who observe other students cheating perceive a greater incentive to cheat (Chacko et al., 2024), so if students view the use of AI tools as a form of cheating, it is likely that this behavior will spread.

When students feel overwhelmed, generative AI could be providing a way to reduce the possibility of overload. Cognitive Load Theory (CLT) proposes that effective learning depends on managing the limited capacity of working memory while facilitating the construction of long-term memory schemas (Sweller, 1988; Sweller et al., 1998). In effect, our memories are limited, and learning can only occur when the total cognitive load does not exceed the individual's working memory. CLT distinguishes among intrinsic load, stemming from the inherent complexity of the material; extraneous load, generated by suboptimal instructional design; and germane load, the cognitive effort devoted to schema formation and automation (Paas et al., 2003). It argues that instructional methods should minimize extraneous load and calibrate intrinsic load to the learner's prior knowledge, thereby maximizing germane processing (van Merriënboer & Sweller, 2005). Empirical research further shows that techniques such as worked examples, goal-free problems, and modality integration enhance learning efficiency by aligning instructional design with cognitive processing constraints (Sweller et al., 2011). College students frequently report experiencing cognitive and emotional overload as they navigate increasing academic, social, and financial demands that exceed their perceived coping resources. Research shows that academic pressures, such as dense coursework, high-stakes assessments, and time-management challenges, are primary contributors to heightened stress and feelings of overload (Misra & McKean, 2000). These demands often interact with developmental transitions, including emerging independence and identity formation, compounding students' sense of being overwhelmed (Conley et al., 2013). Additionally, contemporary factors such as continuous digital connectivity and the expectation of constant multitasking further tax students' attentional and cognitive resources, exacerbating overload and reducing well-being. The cumulative effect of these stressors has been associated with declines in mental health, academic performance, and retention, underscoring the need for institutional interventions that promote resilience and reduce unnecessary cognitive burdens (Pascoe et al., 2020). Consequently, when presented with a new technology, such as generative AI, that appears to instantaneously reduce the academic load placed upon the student, it is not at all surprising that students have enthusiastically used the tool in the completion of assignments. Students have been found to be more likely to use ChatGPT when facing higher academic workloads and time pressures (Abbas et al., 2024). It would not be surprising if the use of generative AI increases in future semesters.

However, there is evidence that the use of tools such as generative AI has a negative impact on student learning (Kosmyna et al., 2025). One potential reason is that the use of a tool reduces the time the student spends experientially learning how to complete the task at hand. Experiential Learning Theory (ELT) posits that learning is an iterative process in which knowledge emerges through the transformation of experience rather than the passive acquisition of information (Kolb, 1984). According to ELT, effective learning involves a four-stage cycle comprised of concrete experience, reflective observation, abstract conceptualization, and active experimentation, with each stage supporting distinct cognitive processes that reinforce one another. ELT emphasizes that learners construct meaning by integrating direct experiences with reflective and conceptual analysis, enabling them to apply insights to new contexts (Kolb & Kolb, 2005). The theory has been widely applied in higher education, where experiential approaches, such as service learning, internships, simulations, and problem-based learning, have been shown to deepen engagement, enhance skill development, and foster the transfer of learning to authentic situations (Beard & Wilson, 2018). Overall, ELT underscores the central role of practice and reflection in fostering adaptive, self-directed learners capable of navigating complex environments. The use of generative AI tools to perform tasks such as completing homework clearly reduces the amount of practice and reflection present

in the learning process, and therefore violates the tenets of ELT. ChatGPT use by college students has been linked to lower academic performance (Abbas et al., 2024).

In the case of AI use in MIS courses, a survey of students who utilized AI tools in a database class found that students viewed the tool as helpful in many aspects of the course, although the author cautioned that educators must consider potential student over-reliance on AI (Zhang, 2025). Jiang and Nakatani (2025) experimented with redesigning learning activities in two IS courses and found that students using AI performed at or above the level of a control group. They also found that students responded favorably to assignments incorporating the use of AI, and that overreliance on the tools was not an issue in their study.

It is clear that more research into this important and evolving topic is warranted, as students increasingly see the use of AI as a potential tool to help reduce overload in all aspects of their lives. Academia must understand the issue in order to develop proper policies and procedures that balance the viability of the learning process with the use of new technologies that can potentially reduce student load.

3. HYPOTHESES DEVELOPMENT

Based on the literature review above and the structure of the course in which we embed our study, we developed four hypotheses.

In our specific course, students were informed that using generative AI was not recommended because exams are hand-written with no possible use of AI tools. However, it was explained that there was no policy against using AI in the course when completing assignments, as long as they indicated on the submission that AI had been used. It was made clear that there was no penalty for using AI, although there was a significant penalty for using AI and not reporting the usage. Most students did not use AI early in the course. However, it was observed that AI usage increased as the course went on. We believe that this was a student response to struggling in the course, and potentially a result of cognitive overload. As students move through a semester, workload increases and stress rises, possibly prompting students to turn to AI tools to help (Abbas et al., 2024). Using CLT as a foundational theory, we hypothesize that those students who reach a critical cognitive load and start to struggle with specific tasks are more likely to look for tools that reduce workload tied to these specific tasks, such as generative AI products. For example, a student who fails an assignment involving the use of SQL will be more likely to utilize AI to complete the next assignment involving SQL. Therefore, we propose:

H1: Negative performance on specific tasks will drive the future use of AI to complete related tasks.

AI is not specifically designed to perform explicit tasks, such as writing SQL code to solve a single problem. Our experience is that it often makes mistakes, although subtle (for example, ChatGPT was found to produce poorly constructed "Group By" statements specific to the queries required in the course). While some research has shown that students using AI tools may perform at a higher level than those who do not, this varies by course type (Wang & Fan, 2025). In the specific case of a technical course such as database management systems, we posit that students relying on AI will miss mistakes made by the tool, and that the practice and reflection called for by ELT will be lacking, leading to poor results. Thus, we hypothesize:

H2: Students who use AI will perform more poorly on assignments than those who do not use AI.

Similarly, using the foundational theory of ELT, we hypothesize that those students using AI tools will be less likely to develop the higher-order learning skills necessary to perform tasks such as designing queries from scratch without the help of AI. Students using AI to simply create answers for an assignment are skipping the most important aspects of ELT (practice, reflection, experimentation, etc.), and will therefore have reduced learning when compared to those students who do not use AI. This reduced learning will be demonstrated on future assessments where AI is not allowed. Furthermore, those students who use AI to simply produce answers for the assignments may not even realize that the code is incorrect, and may repeat those incorrect code structures on the exams. Therefore, we hypothesize:

H3: Students who use AI on assignments will perform more poorly on comprehensive evaluations than those who do not use AI.

However, we believe that students using AI will continue to learn the course material, as knowledge will be required to use the AI to perform specific tasks. Therefore, while we hypothesize above that those using AI will perform more poorly compared to those that do not use AI, we also hypothesize that these students will maintain a grade performance over time that is consistent with those students that do not use AI, albeit at a lower level of performance. Grade performance refers to the student's change in grade over time from assignment 5 to the final exam. In essence, we posit that students using AI will see a change in grade over time consistent with non-AI-using students. The use of AI does not imply that a student will fail to learn a basic level of the material. This leads to:

H4: Students who use AI will maintain a grade performance over time consistent with students who do not use AI.

We expect no meaningful differences for the control variables of gender and major.

4. METHODOLOGY AND DATA COLLECTION

We set our study in an undergraduate database management systems course offered by the Management Information Systems (MIS) department at a large state land-grant university in the mid-Atlantic region of the United States. Course topics included the development of ERDs, transformation of ERDs into relations, and the use of SQL to create databases, populate the databases, and query the data to solve business problems.

4.1. Data Collection and Sample

The sample for the study consisted of the students in the database course, which was offered via three sections in a face-to-face format, meeting three times per week. One section was offered in the Spring (37 clean data points out of 44 students) and two in the Fall (67 total clean data points out of 87 students). All three sections utilized the same syllabus, course policies, etc. In our study, we did not include students who did not receive grades on all assignments and the final exam.

The course has been taught by a group of experienced faculty in the same manner for several years, including the time prior to the advent of generative AI. Student performance was evaluated via a mix of assignments and in-class exams. Table 1 shows the chronological order in which these assessments occurred.

Evaluation	Content	AI Tracking
Assignment 1	Create a basic ERD from description – No code	
Assignment 2	Create a complex ERD from document + description – No code	AI use tracked
Exam 1 - ERDs	ERDs, short answers, No code	No AI use possible
Assignment 3	Transformation & Normalization Agt – No code	AI use tracked
Assignment 4	Project Description & ERD Agt – no code	AI use tracked
Exam 2 – SQL exam	SQL Code	No AI use possible
Assignment 5	SQL Agt – SQL Code	AI use tracked
Final Exam	Comprehensive	No AI use possible

Table 1. Course Evaluation Structure

Assignments 1, 2, and 3 were identical for all students in each section but changed each semester to reduce cheating (while requiring the same knowledge and skills). Assignments 4 and 5 were comprised of an individual project where students had to create their own scenario and develop the appropriate ERD in assignment 4, and then create the corresponding database and construct queries using SQL in assignment 5.

The assessments were all graded in the same manner and by the same instructor in each section, whether students used AI or not. Students were clearly informed that the use of AI was unrecommended but was allowed with no potential penalties when completing the assignments. Students were only required to indicate whether they utilized AI tools to complete the assignment. They were reminded of this several times throughout the semester and on the assignments themselves, and there was a significant penalty for using AI and not self-reporting the use. There were no indications that any student utilized an AI tool without making it clear that they had done so.

All assignments were completed on the students' own time, and they were given a minimum of two weeks to complete each assignment (and often longer). This provided the students with ample opportunity to use AI, if they wished to do so.

We chose to focus on assignments 4 and 5 for multiple reasons. Most importantly, these were unique assignments for each student. Each student was required to develop their own scenario, completely unique from any other student's scenario. As assignments 1-3 were the same for each student, there was the possibility that students were collaborating (and potentially even cheating), thus possibly confounding the results. This was much less of a possibility with assignments 4 and 5. Anecdotally, it was clear that more students collaborated on the first three assignments than on assignments 4 and 5. Secondly, assignments 4 and 5 were the most complex and comprehensive in the course, and assignment 5 required the generation of complex SQL code. Finally, due to the nature of AI generated SQL code, it was easier to tell if students were being honest about their usage of AI tools. ChatGPT, in particular, adds artifacts such as automatic generation of primary keys, trickle-down modification mechanisms to foreign keys, and multi-criteria "Group Bys," all of which were not significantly used in the course. This provided a level of assurance that students were being honest regarding their AI use. As stated, no instances of students failing to admit to AI use when it appeared that they had done so were discovered. As a consequence, in line with our hypothesis-testing needs, students subsequently identified as "AI users" are those who used AI on assignment 5.

The control variables of gender and major were also included in the data.

4.2. Statistical Tests

From a statistical perspective, for each assignment and exam, we conducted normality tests on the grades within each group (see Appendix A for details), and f-tests were used to see if the samples of AI users and non-users had the same variances in performance. Based on these tests, we conducted Mann-Whitney U tests when the grades were not normally distributed, and the appropriate t-tests (based on variance f-tests) when they were normally distributed. This allowed us to see whether the means were the same between AI and non-AI users on assignments and exams (Ott & Longnecker, 2010).

The evaluation structure of our course has been consistent for several years, with the same set of experienced instructors teaching the classes, and no noticeable fluctuations in performance observed during those years. Thus, we assume that the results of our statistical procedures are not impacted by measurement reliability issues. To quantitatively support our reliability assumption, we collected a secondary sample of student performance in sections offered one year prior to our study when AI use was still marginal (one section in the Spring, n=46; two sections in the Fall, n=54). We conducted the same statistical procedures used on the main sample on this secondary sample to observe whether means differ from one semester to the other. No statistically significant differences were found in performances on assignments and exams from one semester to the other. Therefore, we are confident that our course evaluation structure provides reliable performance measurements.

5. ANALYSIS AND RESULTS

A total of 104 students were included in our study, consisting of 78 (75%) males and 26 (25%) females. Sixty-seven (64%) of the students were majoring in MIS, with the rest distributed among other majors in the college and a pair of Computer Science majors in the College of Engineering. As per the professor's recommendation, very few students used AI early in the semester (only 5 students for assignments 2 and 3, and 14 for assignment 4). However, 52 (50%) of the students indicated that they had used generative AI on assignment 5, which was due late in the semester and required the generation of the course's most complex SQL code. We, therefore, focus our study on their decision to adopt AI for assignment 5. Table 2 provides the descriptive statistics for student grades for each assignment and exam of interest to our study, beginning with assignment 4. All grades are presented as percentages, and a straight scale was used in the class for assigning letter grades (90-100 = A, 80-89.9 = B, 70-79.9 = C, etc.).

Measure	Group	N	Mean	SD	Var.	Min	Max*	Skew. Stat.	Skew. Std Err.	Kurtosis Stat.	Kurtosis Std. Err.
AGT4	No AI	52	86.60	16.51	272.64	25	105	-1.922	.330	4.814	.650
	AI users	52	83.71	15.25	232.48	40	107	-1.019	.330	.643	.650
SQL Ex.	No AI	52	76.84	14.88	221.28	40	103	-.040	.330	-.758	.650
	AI users	52	65.66	19.47	379.00	12.5	96.43	-.613	.330	.051	.650
AGT5	No AI	52	88.83	9.00	80.97	66	100	-.476	.330	-.490	.650
	AI users	52	84.56	14.83	220.05	20	100	-1.980	.330	6.035	.650
Final	No AI	52	75.81	11.96	142.93	51.25	100.55	-.093	.330	-.244	.650
	AI users	52	69.16	13.38	179.10	33.75	91.25	-1.298	.330	1.929	.650
**Due to bonus points being available, maximum grades could exceed 100%.*											

Table 2. Descriptive Statistics of Course Assessments

Tables 3 and 4 are between-group comparisons of mean grades. For transparency, the results of all tests for each hypothesis are displayed, allowing the reader to compare results, even when test assumption violations forced us to conduct different tests despite similar settings. In Tables 3 and 4, results in italics violate the statistical test assumptions. Table 3 is relevant for Exam 2 (SQL exam), which had normally distributed grades within both groups. Table 4 shows the results of Mann-Whitney U-tests when the grades were found to be non-normally distributed. This test is relevant for assignments 4 and 5, and the final exam. These tests allow us to investigate hypotheses 1 to 3.

Hypo	Measure	Mean Diff.	df	t	p (one sided)	95% CI Lower	Upper	Cohen's d
H1	*AGT4*	*2.885*	*102*	*.926*	*.178*	*-3.297*	*9.067*	*.182*
H1	SQL Ex.	11.174	102	3.289	<.001	4.435	17.913	.645
H2	*AGT5*	*4.269*	*84.056*	*1.774*	*.040*	*-.515*	*9.054*	*.348*
H3	*Final*	*6.643*	*102*	*2.670*	*.004*	*1.707*	*11.580*	*.524*

Table 3. Between Group Comparison of Means of Course Assessments (independent samples t-test)

Hypo	Measure	Group	N	Mean Rank	M-W U	p (2-tailed)
H1	AGT4	No AI	52	56.19	1160.00	.211
		AI user	52	48.81		
H1	*SQL Ex.*	*No AI*	*52*	*61.05*	*907.50*	*.004*
		AI user	*52*	*43.95*		
H2	AGT5	No AI	52	55.97	1171.50	.239
		AI user	52	49.03		
H3	Final	No AI	52	59.76	974.50	.014
		AI user	52	45.24		

Table 4. Between Group Comparison, Course Assessment Means (Mann-Whitney U-test)

Table 5 shows the results of a paired t-test that compares within group grades and allows us to investigate hypothesis 4. Both Tables 4 and 5 also have the advantage of displaying Cohen's d, a measure of effect size that does not require normality if we restrain from making percentile inferences (0.2 = small; 0.5 = medium; 0.8 = large; Cohen, 1988). Below, we include this measure in the interpretation of our results.

Hypo	Measure*	Mean	Std Dev.	df	t	p (two-sided)	95% CI Lower	Upper	Cohen's d
H4	Adj.AGT5 – Final Ex.	2.373	19.454	51	.880	.383	-3.042	7.790	.122
** Adj. AGT5 is the grade for AGT 5 – 13.02 (the mean difference observed between AGT5 and the Final for AI non-users.)*									

Table 5. Within Group Comparison Between AGT5 and the Final (paired samples t-test)

H1: Negative performance on specific tasks will drive the future use of AI to complete related tasks. SUPPORTED

We tested two scenarios: did a negative result on assignment 4 drive the use of AI on assignment 5, which was an unrelated task (assignment 4 involving case descriptions and ERDs, while assignment 5 involved SQL coding), and did a negative result on the SQL exam drive the use of AI on assignment 5, a related assignment also focusing on SQL coding. Thus, two tests were performed.

In the first two rows of Table 2, we see that those who used AI for assignment 5 had an average grade of 83.71 on assignment 4. Those who did not use AI had an average of 86.6, barely a 3-point difference between groups. The first two rows of Table 3 show that we did not find a statistical difference in means on assignment 4 between those who ended up using AI on assignment 5 and those who did not ($p = .211$). It seems there was no significant difference in performance on assignment 4, which did not contain SQL coding, between those who later used AI on a SQL coding assignment and those who did not.

On the contrary, rows 5 and 6 of Table 2 show a 4.3 difference in means on the SQL exam between those who did not use AI and those who did on assignment 5 (88.83 vs. 84.56, respectively). In the second row of Table 4, we see that the difference in means between groups on the SQL exam is related to the use of AI on assignment 5 ($p < 0.001$). Also, a Cohen's d of .645 shows that the effect size is between medium and large.

We thus find support that performance on specific related tasks (in this case, SQL coding) drives the use of AI on later related tasks. The combination of these two tests supports hypothesis 1.

H2: Students who use AI will perform more poorly on assignments than those who do not use AI. NOT SUPPORTED

Based on Table 2, we can see that non-AI users received an average grade of 88.8 on assignment 5, whereas AI users received an average grade of 84.56. Using a t-test (Table 4), this difference of more than 4 points was sufficient to find a statistically significant difference ($p = .04$). Also, in Table 4, Cohen's d is between small and medium ($d = .348$), which would indicate a non-negligible effect size. However, the grades for assignment 5 are not normally distributed in each group, thus the Mann-Whitney U-test was the relevant test. The results shown in Table 3 (rows 5 and 6) do not show a statistically significant difference between the groups ($p = .239$). Thus, we do not find support for hypothesis 2.

H3: Students who use AI on assignments will perform more poorly on comprehensive evaluations than those who do not use AI. SUPPORTED

In this course, the final exam provided a comprehensive evaluation of student knowledge of the course material. The mean grade for non-users of AI in the course had a final grade of 75.81, while AI users had a mean grade of 69.16. However, grades on the final exam were not normally distributed, so we performed a U-test, with results shown in the last two rows of Table 3. There is a clear significant difference in mean results between AI users and non-users ($p = .014$); those who used AI for assignment 5 performed more poorly on the comprehensive final exam. Table 4 shows a medium Cohen's d ($d = .524$), thus providing support for hypothesis 3.

H4: Students who use AI will maintain a grade performance over time consistent with students who do not use AI. SUPPORTED

We performed a paired sample t-test to investigate if the performance on the comprehensive final exam was similar to the performance on assignment 5 for those using AI. The data show that those who did not use AI had a difference in grade of 13 points between assignment 5 and the final. Students who used AI on assignment 5 performed more poorly than non-users. However, they also performed more poorly on the final exam than non-users, and we did not find statistical evidence for an increase or decrease in difference when comparing the two groups. AI users seem to keep performing at the level they were performing with the help of AI, even when they do not use AI anymore. The statistical test in Table 5 did not find a significant difference between this difference and that of those who used AI in the course. Cohen's d is small ($d = .122$), thus supporting the hypothesis.

5.1 Control Variables Tests

We did not make any assumptions regarding differences in grades or AI use between genders or majors. Dividing students into two groups representing MIS majors and non-MIS majors (mostly from supply chain and marketing), we did not find any differences between groups in performance nor regarding AI use. The fact that a third of our 104 students were from other majors and followed the same trends as the MIS majors supports the conclusion that the results of our study apply to the whole population of business students taking similar MIS classes and not just to the subset of MIS students.

On the contrary, gender was found to be linked to AI use for assignment 5. Since both gender and AI use were dichotomous variables in our data (0 or 1), we performed a Pearson Chi-Square test (shown in Appendix B). No assumption was violated, and the test is significant ($p = .024$), which shows that men tended to adopt AI more than women. To further analyze this finding, we ran two logistic regressions to see how gender and results on the SQL exam predicted AI use for assignment 5 (see Appendix C for details). A model with gender and grade on the SQL exam predicts AI use at 18.6%, with both independent variables being significant. A model with an interaction term does not add percentage explained and the interaction term is insignificant. This means that gender and performance on the SQL exam predict some of AI use independently of one another. Males and those with poor grades on the assignment adopt AI, but poor coding men do not adopt AI more than others.

Table 6 summarizes our findings in terms of support for our hypotheses.

	Hypothesis	Result
H1	Negative performance on specific tasks will drive the future use of AI to complete related tasks.	Supported
H2	Students who use AI will perform more poorly on assignments than those who do not use AI.	Not Supported
H3	Students who use AI on assignments will perform more poorly on comprehensive evaluations than those who do not use AI.	Supported
H4	Students who use AI will maintain a grade performance over time consistent with students who do not use AI.	Supported

Table 6. Summary of Support for the Hypotheses

6. DISCUSSION

We want to emphasize that our results need to be interpreted in the context in which they are embedded. As presented in our literature review, current research in educational IS (e.g., Jiang & Nakani, 2025; Sun & Deng, 2025; Zhang, 2025) shows that AI can be beneficial when used properly in MIS classes. We adopted a different but complementary perspective. We looked at AI use when AI is unrecommended and its use is not structured by professors. The two angles we adopted were to investigate what drives AI use, and what effects exist on performance and learning.

Regarding our first research question, our study shows that the students' choice to use AI is not guided only by poor performance on previous assignments. It is guided by poor performance on the specific task of coding, knowing that more coding testing was coming in the latter stages of the semester. As control investigations, we also tried models (not displayed in this paper) where we attempted to link performance on all preceding assignments to AI use on assignment 5, but none of these models showed statistical significance. This reinforces our impression that AI is not used as a panacea to fix general performance but rather as a potential remedy when weakness on a specific task is identified. Students do not resort to using AI because they see their overall grades as being too low and AI as a way to fix the situation. They use AI because they identify a difficulty with a specific task, in this case coding in SQL, and the potential of AI to help in this domain. Additional qualitative data gathered in the course of our study supports the claim that students overwhelmingly resorted to ChatGPT for help when faced with specific challenges, not as an overall course strategy from the beginning of the semester. We argue that this was not the best choice of technology usage. In the instructional section of this discussion, we provide more details.

Regarding the second research question on performance and learning, our results may suffer from a lack of statistical power. Hypothesis 2 is not statistically supported, but hypotheses 3 and 4 are. Although it appears that students who used AI performed more poorly on assignment 5, the lack of statistical support might be interpreted as AI acting as a mediocre learning tool in this situation. There is some support for this finding. An analysis of more than one million written discussion board submissions at a major university between 2021 and 2024 found that writing quality improved in the years after ChatGPT was introduced. The biggest improvement was found with students who were "linguistically disadvantaged," such as those who spoke English as a second language (Barshay, 2025). Our hypothesis was based on the idea that students would not be knowledgeable enough to identify the errors in the ChatGPT-produced code, but it could be that the code produced is still of high enough quality to receive grades comparable to the code submitted by non-AI-using students. However, this use of AI is of short benefit, since students who used AI then perform worse on evaluations where they do not have access to AI. The effect is not entirely disastrous because students seem to learn some content thanks to their use of AI. It may be that they are placing too much trust in the AI tool, and learning code that is incorrectly produced by the AI.

Some elements that are dependent on the design of the syllabus were not included in this study. In our specific setting, we found a negative effect of AI use on the semester grade and the letter grade (e.g., AI users performed at a C level, as opposed to the B- level of non-users). Of course, this is dependent on the weight given to SQL coding assignments and exams in the semester grade. However, assuming that we

distributed the weight of our assignments as per academic standard practices, it is interesting to see that the difference in performance observed at the assignment and exam levels has repercussions on semester performance. This shows that students should not count on their work on early assignments to cancel a potential deficit when they use AI for the SQL coding part of the semester. Zhang (2025) suggests developing assignments that require reflections and detailed documentation of the problem-solving process, as a potential means to reducing over-reliance on AI tools leading to less rigorous learning. This is a potential course of action to counter the negative impact of AI use.

6.1 Research Contribution

Our research contributes to cognitive overload theory by showing how students react to being overloaded. The data show that students did not suddenly substitute every piece of their work with AI outputs, but rather just the pieces that were creating their cognitive overload. In our case, it means using AI only for coding tasks even when comprehensive exams are given.

The finding showing that the performance of AI users did not decrease more than that of non-users on the final exam, as compared to assignment 5, shows that using AI can be a way to learn. As per the literature that advocates for formative assessment, assignments including the recommended use of AI could be a way for the professor to get students actively involved with generating code and then identifying where this code is incorrect. In this way, students would learn through the performance of the AI-assisted assignments and through the in-class feedback on that assignment, thus utilizing the tenets of ELT. Our study shows that students start using unrecommended AI when they start coding in SQL. Drawing from CLT, this behavior is a sign of saturation. However, ELT then tells us that the use of AI may lead to a lack of higher order learning, as the students, now under stress, utilize AI instead of performing the necessary iterative steps to properly learn the material. In cases such as these in other academic fields, there are actions that are identified as helping. Taking a break might not be a possibility, as we are constrained by the structure of academic semesters, but changing the IS environment is a possibility. For example, switching from typing commands to generating code with AI and seeing what it does in a playful and iterative way may help desaturate cognitively loaded students.

Several research studies on students' use of AI have been surveys, which depend on students self-reporting their usage of AI and their perceptions of AI benefits or drawbacks for learning (e.g., Zhang, 2025). A few research studies, such as that by Walsh (2025), have used students' comments and anecdotal data. One study (Jiang & Nakatani, 2025) used control groups (separating students who used AI and those who did not) set up by the authors. Our research makes unique contributions by being different from all the above in that (a) we used a naturally occurring control group (the students in the course who chose not to use AI for their assignments) and (b) we used an objective criterion in the form of their performance on a comprehensive examination (not self-reported, perceived, or anecdotal data) to evaluate how the use or non-use of AI affected their learning and performance.

It is to be noted that, in our study, instructors provided feedback on the mistakes made when coding, whether AI was used or not. However, feedback was provided not emphasizing the mistakes AI makes. Such emphasis could be what reinforces learning when AI is included in the curriculum.

6.2 Instructional Contribution

In a context where the course is designed for students to grasp the essentials of relational databases and SQL, this study shows that AI use does not help students develop a good understanding of more complex concepts (such as "Group By" queries in SQL), nor does it aid students in understanding how to apply these concepts to solve new problems. Instructors could identify such concepts in their respective courses to emphasize to students the need to avoid over-relying on AI use for learning. They could also demonstrate in class how an AI tool provides misleading and ambiguous responses to cases where those concepts need to be applied. Instructors could use these examples to motivate students to understand when AI can and cannot be used effectively. To do so, instructors can use reverse engineering techniques and generate code with AI tools. Later in the course, instructors can ask the students to find out what the code is supposed to do and what seems to be off.

Similarly, in the beginning of the semester, instructors should encourage students to work without AI. The results of our study contribute to giving incentives to do so. Students do not perform better with AI than without, and the evidence clearly shows that it can affect their course grade. Whether the instructor recommends the use of AI or not, a warning regarding what AI to use can be helpful. Popular AI tools such as ChatGPT or Claude can help convert text into relational models. On assignment 2, the 5 people who used AI seemed to perform well, although the sample size is too small to make any claims. One of our co-authors is an expert in journalism and has seen students doing well on their assignments when they use AI to convert a manuscript into a short summary listing the characteristics of main topics. This structure is akin to transforming a natural language text into relations with a list of attributes. Thus, natural language based generative AI, such as Claude and ChatGPT, can be relevant for building relations. Between ChatGPT and Claude, anecdotal evidence suggests that Claude could be a better "all-purpose" tool when it comes to translating natural language into SQL DDL code. Other more dedicated tools are clearly better for queries, and some are free; for example, EverSQL and Text2SQL. When it comes to converting ERDs to relations and SQL code, the MySQL Workbench still appears to be a better option than readily available AI, in our experience. If instructors want to encourage the use of AI, they can guide which one to use. If they do not want to encourage the use of AI, they can combine their recommendation not to use AI with a warning that the use of familiar general-purpose AI tools tends to be counterproductive in the later stages of the semester, notably when it comes to building queries.

6.3 Limitations and Future Research

It should be noted that this study did not consider the extent of the use of AI by each student, which is a limitation of the study. It may be that some students simply used AI to provide some feedback on their own ideas, while others used AI to complete the entire assignment. We recommend that future studies consider the extent of the use of AI as a factor in their analysis because its impact is not obvious. Two students might be using AI on the same question, but the extent to which a student iterates with the machine can substantially impact the student learning if outputs are questioned.

Beyond the impact of generative AI on education, future research is clearly needed into how academia better prepares students for the use of AI in the workplace. A Washington State University survey found that 83% of professionals believe that students should be prepared to use AI upon their entry into the workforce, and 74% believed that incoming employees should already have AI experience prior to employment (Hollenbeck, 2024). An even larger 88% placed the responsibility on colleges and universities to provide these educational opportunities. However, a report from Cengage Group (2024) found that over half of college graduates did not feel that their institution had prepared them to use generative AI tools, and 69% required further training in their current jobs. Cost has been reported as an issue for institutions when it comes to providing students with access to AI tools (Flaherty, 2025). Research is needed on the best way to prepare students for AI's role in the workplace. Also, if educators are going to integrate AI tools into the curriculum, another area of future research and consideration is the need to establish equitable access to AI tools through allocation of resources and enablement (Wang, 2024). Similarly, Al-Jahwari and Yousif (2025) argue that while the use of ChatGPT in education holds great promise, more consideration needs to be given to ethical considerations, educator training, and potential regulation.

While this paper has focused on the use of AI by students, a growing area of interest is the use of AI by the educators themselves. Students are now voicing complaints about faculty using AI for instruction or grading (Hill, 2025). There have been several reported issues with the use of AI-detection tools incorrectly penalizing students (e.g., Holtermann, 2025). We most likely wish to avoid a situation where faculty use AI to generate content and assignments, students use AI to complete the assignments, and faculty use AI to grade the assignments. It is not difficult to envision such a possibility in the not-so-distant future.

7. CONCLUSION

It is clear that AI tools are no passing fad, nor are they a panacea for students looking to improve their grades and manage their time load. Our study demonstrates that there are significant learning impacts when

students utilize AI as an aid in completing assignments, particularly when it is unrecommended. While students continue to learn at a rate similar to their non-AI-using colleagues, they learn at a significantly lower grade level and show a greater lack of understanding of the material on comprehensive evaluations where AI is not allowed, especially when it comes to specific tasks such as coding. The use of AI is understandable. The tools are ubiquitous, easy to use, and perform well on many tasks. Students faced with cognitive overload and the stress of deadlines cannot be faulted for resorting to such a tool. In fact, anecdotal evidence suggests that faculty are acting in a similar manner. However, it is our duty as IS educators to ensure that our students have the knowledge necessary to succeed. This research will hopefully help guide us toward that goal.

8. REFERENCES

Abbas, M., Jam, F. F., & Khan, T. I. (2024). Is It Harmful or Helpful? Examining the Causes and Consequences of Generative AI Usage Among University Students. *International Journal of Educational Technology in Higher Education*, 21(10). https://doi.org/10.1186/s41239-024-00444-7

Abu Qub'a, A., Abu Guba, M. N., & Fareh, S. (2024). Exploring the Use of Grammarly in Assessing English Academic Writing. *Heliyon*, 10(15). https://doi.org/10.1016/j.heliyon.2024.e34893

Al-Jahwari, M., & Yousif, M. J. (2025). The Impact of AI Tools on Education: ChatGPT in Focus. *Artificial Intelligence & Robotics Development Journal*, 4(4), 314-336, https://doi.org/10.52098/airdj.20244430

Anderson, L. (2025, May). How Are Students Using ChatGPT? For Therapy, Breakups and Even Texting Friends. *Teen Vogue*. http://www.teenvogue.com/story/how-students-using-chatgpt-therapy-breakups

Barshay, J. (2025, April 14). International Students May Be Among the Biggest Early Beneficiaries of ChatGPT. *The Hechinger Report*. http://hechingerreport.org/proof-points-international-students=chatgpt/

Beard, C., & Wilson, J. P. (2018). *Experiential Learning: A Practical Guide for Training, Coaching and Education* (3rd ed.). New York, NY, USA: Kogan Page.

Cengage Group. (2024). *2024 Graduate Employability Report: Preparing Students for the GenAI-Driven Workplace*. https://cengage.widen.net/s/bmjxxjx9mm/cg-2024-employability-survey-report

Chacko, J., Collignon, S., Green, C. L., & Peace, A. G. (2024). Understanding the Impact of Perceptions of Student Assessment Congruence, Satisfaction, and Peer Behavior on Cheating Incentive in Higher Education. *Journal of Information Systems Education*, 35(4), 481-494. https://doi.org/10.62273/CSGL6637

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Conley, C. S., Travers, L. V., & Bryant, F. B. (2013). Promoting Psychosocial Adjustment and Stress Management in First-Year College Students: The Benefits of Engagement in a Psychosocial Wellness Course. *Journal of American College Health*, 61(2), 75-86. https://doi.org/10.1080/07448481.2012.754757

Dhulipala, M., & Bhujade, A. (2025). Analyzing the Impact of AI on Exam Preparation and Student Performance in Higher Education. *Journal of Informatics Education and Research*, 5(2). https://doi.org/10.52783/jier.v5i2.2924

Eccles, L., & Hayton, L. (2025, May 24). "An Existential Crisis": Can Universities Survive ChatGPT? *The Sunday Times*. http://www.thetimes/uk/technology-uk/article/an-existential-crisis-can-universities-survive-chatgpt-7r38r72sv

Flaherty, C. (2025). The Digital Divide: Student Generative AI Access. *Inside Higher Ed*. https://www.insidehighered.com/news/tech-innovation.artificial-intelligence/2025/04/21/half-colleges-dont-grant-students-access/

Hill, K. (2025, May 14). College Professors Are Using ChatGPT. Some Students Aren't Happy. *The New York Times*. https://www.nytimes.com/2025/05/14/technology/chatgpt-college-professors.html

Hollenbeck, E. (2024). Workers Sound the Alarm on AI Workplace Readiness. *WSU Insider*. https://news.wsu.edu/news/2024/01/25/american-workers-sound-the-alarm-on-ai-workplace-readiness/

Holtermann, C. (2025, May 17). How Students Are Fending off Accusations That They Used AI To Cheat. *The New York Times*. https://www.nytimes.com/2025/05/17/style/ai-chatgpt-turnitin-students-cheating.html

IBM. (2025). *What Is Generative AI?* https://research.ibm.com/blog/what-is-generative-AI

Jiang, Y., & Nakatani, K. (2025). Exploring Implementations of GenAI in Teaching IS Subjects and Student Perceptions. *Journal of Information Systems Education*, 36(2), 180-194. https://doi.org/10.62273/WFHO1011

Kastberg, S., & Leatham, K. (2005). Research on Graphing Calculators at the Secondary Level: Implications for Mathematics Teacher Education. *Contemporary Issues in Technology and Teacher Education*, 5(1), 25-37.

Kolb, D. A. (1984). *Experiential Learning: Experience as the Source of Learning and Development*. Wilmington, DE, USA: Prentice-Hall.

Kolb, A. Y., & Kolb, D. A. (2005). Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. *Academy of Management Learning & Education*, 4(2), 193-212. https://doi.org/10.5465/amle.2005.17268566

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your Brain on ChatGPT: Accumulation of Cognitive Debt When Using an AI Assistant for Essay Writing Task. Ithaca, NY, USA: Cornell University. https://doi.org/10.48550/arXiv.2506.08872

Lin, L. (2024). A Quarter of U.S. Teachers Say AI Tools Do More Harm Than Good in K-12 Education. *Pew Research Center*. https://www.pewresearch.org/short-reads/2024/05/15/a-quarter-of-u-s-teachers-say-ai-tools-do-more-harm-than-good-in-k-12-education/

Misra, R., & McKean, M. (2000). College Students' Academic Stress and Its Relation to Their Anxiety, Time Management, and Leisure Satisfaction. *American Journal of Health Studies*, 16(1), 41-51.

Mogavi, R. H., Deng, C., Kim, J. J., Zhou, P., Kwon, Y. D., Metwally, A. H. S, Tlili, A., Bassanelli, S., Bucchiarone, A., Gujar, S., Nackle, L. E., & Hui, P. (2024). ChatGPT in Education: A Blessing or a Curse? A Qualitative Study Exploring Early Adopters' Utilization and Perceptions. *Computers in Human Behavior: Artificial Humans*, 2(1), 100027. https://doi.org/10.1016/j.chbah.2023.100027

OpenAI. (2025). *Building an AI-Ready Workforce: A Look at College Student ChatGPT Adoption in the US*. https://cdn.openai.com/global-affairs/openai-edu-ai-ready-workforce.pdf

Ott, R. L., & Longnecker, M. (2010). *An Introduction to Statistical Methods and Data Analysis*. Boston, MA, USA: Cengage Learning Inc.

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, 38(1), 1-4. https://doi.org/10.1207/S15326985EP3801_1

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The Impact of Stress on Students in Secondary School and Higher Education. *International Journal of Adolescence and Youth*, 25(1), 104-112. https://doi.org/10.1080/02673843.2019.1596823

Penglase, M., & Arnold, S. (1996). The Graphics Calculator in Mathematics Education: A Critical Review of Recent Research. *Mathematics Education Research Journal*, 8(1), 58-90. https://doi.org/10.1007/BF03355481

Sidoti, O., Park, E., & Gottfried, J. (2025). *About a Quarter of U.S. Teens Have Used ChatGPT for Schoolwork – Double the Share in 2023*. Washington, DC, USA: Pew Research Center. https://www.pewresearch.org/short-reads/2025/01/15/about-a-quarter-of-us-teens-have-used-chatgpt-for-schoolwork-double-the-share-in-2023/

Sun, R., & Deng, X. (2025). Using Generative AI to Enhance Experiential Learning: An Exploratory Study of ChatGPT Use by University Students. *Journal of Information Systems Education*, 36(1), 53-64. https://doi.org/10.62273/ZLUM4022

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., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. New York, NY, USA: Springer. https://doi.org/10.1007/978-1-4419-8126-4

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

van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, 17(2), 147-177. https://doi.org/10.1007/s10648-005-3951-0

Walsh, J. D. (2025). Everyone Is Cheating Their Way Through College. *New York*. https://nymag.com/intelligencer/article/openai-chatgpt-ai-cheating-education-college-students-school.html

Wang, J., & Fan, W. (2025). The Effect of ChatGPT on Students' Learning Performance, Learning Perception, and Higher-Order Thinking: Insights From a Meta-Analysis. *Humanities and Social Sciences Communications*, 12(621). https://doi.org/10.1057/s45199-025-04787-y

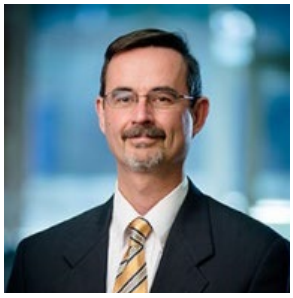
Wang, Y. (2024). The Impact of AI Applications in Education on Students. *Journal of Education Humanities and Social Sciences*, 45, 116-122. https://doi.org/10.54097/ggld3q19

Zhang, X. (2025). Incorporating AI Tools Into Database Classes. *Journal of Information Systems Education*, 36(1), 37-52. https://doi.org/10.62273/GKZI2477

AUTHOR BIOGRAPHIES



Stéphane E. Collignon is an Associate Professor with the Department of MIS and Supply Chain Management at West Virginia University. He received his Ph.D. from Virginia Tech. Collignon's research interests are information systems issues in logistics, privacy management, and education. Collignon works with quantitative methodologies but is also an expert in qualitative research (Grounded Theory). He teaches information systems, data management, decisions sciences, and business intelligence. Collignon worked for 7 years in R&D for two French distribution companies. His research notably appeared in *Information & Management*, *Journal of Business Logistics*, *European Journal of Operational Research*, *Decision Sciences Journal of Innovative Education*, and *Journal of Information Systems Education*.



A. Graham Peace is an Associate Professor of Management Information Systems in the John Chambers College of Business & Economics at West Virginia University. He formerly worked at IBM and obtained his doctorate from the University of Pittsburgh. Dr. Peace has served as Associate Dean for Undergraduate Programs and Chair of multiple departments while at WVU. His research and teaching interests include information ethics, database management systems, and pedagogical issues. His research has appeared in the *Journal of Management Information Systems*, *Communications of the ACM*, *Journal of Information Systems Education*, and *Journal of Global Information Technology Management*.



Lauren Peace is an Adjunct Professor of Journalism in the College of Arts and Letters at the University of Tampa in Tampa, FL. She received her Master's degree in journalism from Columbia University, where she studied as a Stabile Investigative Fellow. She was previously a Fulbright Student Research Scholar in Kosovo where she completed an oral history project and public art exhibition featuring women who had acted as catalysts for social and political change**.**



Nanda C. Surendra is an Associate Professor of Management Information Systems in the John Chambers College of Business & Economics at West Virginia University. He received his Ph.D. from the Department of Information Systems at the University of Cincinnati. His main research interest is studying the design and development of user-centered, effective, and ethical software systems. He has published in the *European Journal of Information Systems*, *Information & Management*, and *Information Technology and Management*.

APPENDICES

Appendix A. Tests of Normality

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Measure		Statistic	df	Sig.	Statistic	df	Sig.
AGT2	No AI	.170	52	<.001	.859	52	<.001
	AI Users	.131	52	.026	.893	52	<.001
AGT3	No AI	.188	52	<.001	.834	52	<.001
	AI Users	.139	52	.014	.907	52	<.001
AGT4	No AI	.167	52	<.001	.820	52	<.001
	AI Users	.149	52	.006	.914	52	.001
AGT5	No AI	.138	52	.015	.934	52	.007
	AI Users	.175	52	<.001	.830	52	<.001
SQL Exam	No AI	.109	52	.171	.966	52	.136
	AI Users	.083	52	.200*	.962	52	.094
Final Exam	No AI	.077	52	.200*	.976	52	.359
	AI Users	.205	52	<.001	.866	52	<.001

*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Interpretations: The test helps to figure if the distribution of grades is significantly different from a normal distribution. Non-significant results mean that we fail to find a significant difference and thus we need to assume it is normal. Significant results mean it is found to be different and thus should not be considered normally distributed.

Most of the grades have non-normal distribution except for the SQL exam (for both groups) and the final exam (for the AI non-users).

Appendix B. Effect of Gender on AI Adoption - Chi-Square Test

Use AI AGT5 * Gender (0 = Male) Crosstabulation

			Gender (0 = Male)		
			0	1	Total
use AI AGT5	0	Count	34	18	52
		Expected Count	39.0	13.0	52.0
	1	Count	44	8	52
		Expected Count	39.0	13.0	52.0
Total		Count	78	26	104
		Expected Count	78.0	26.0	104.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.128[a]	1	.024		
Continuity Correction[b]	4.154	1	.042		
Likelihood Ratio	5.233	1	.022		
Fisher's Exact Test				.040	.020
Linear-by-Linear Association	5.079	1	.024		
N of Valid Cases	104				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 13.00.
b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approximate Significance
Nominal by Nominal	Phi	-.222	.024
	Cramer's V	.222	.024
N of Valid Cases		104	

Appendix C. Logistic Regression Analysis

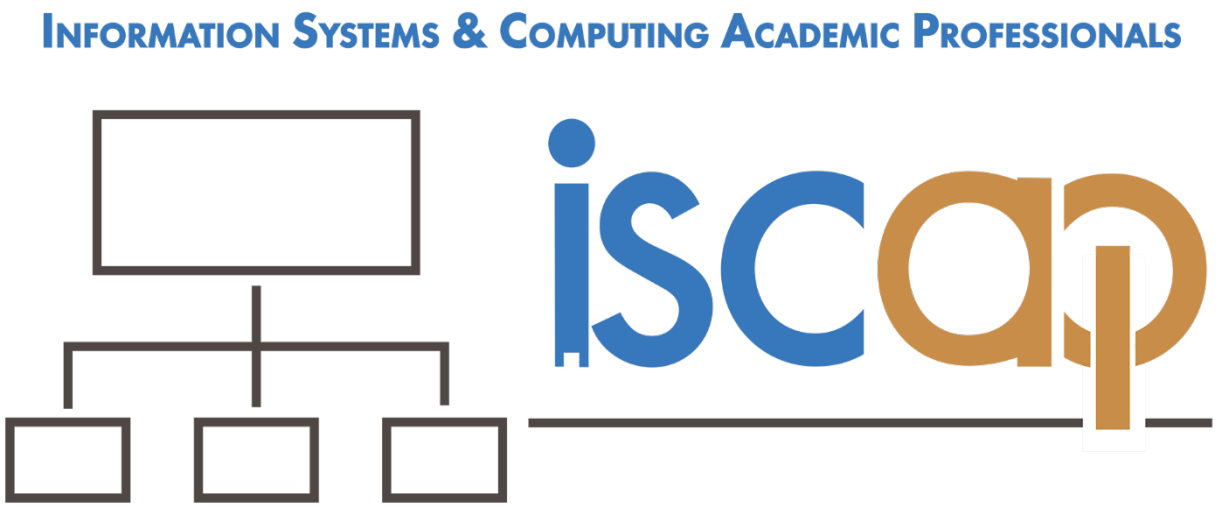
DV – AI adoption for AGT5	**Variables**	**Model 1**	**Model 2**
Key Independent Variable	SQL Exam Grade	-.039** (.013)	-.039** (.013)
Interaction	SQL Ex. X Gender	--------	.000 (.037)
Control variable	Gender (0= Male)	-1.119* (.512)	-1.097 (3.278)
Constant	Constant	3.088** (.991)	3.087** (1.004)
Nagelkerke R^2		.186	.186

Standard errors are indicated in parentheses
*p<0.05; **p<0.01; ***p<0.001.

Classification Table for Model 1[a]

			Predicted Use AI AGT5		
	Observed		0	1	Percentage Correct
Step 1	Use AI AGT5	No AI	31	21	59.6
		AI Users	18	34	65.4
	Overall Percentage				62.5

a. The cut value is .500



STATEMENT OF PEER REVIEW INTEGRITY

All papers published in the ***Journal of Information Systems Education*** have undergone rigorous peer review. This includes an initial editor screening and double-blind refereeing by three or more expert referees.

Copyright ©2026 by the Information Systems & Computing Academic Professionals, Inc. (ISCAP). Permission to make digital or hard copies of all or part of this journal for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial use. All copies must bear this notice and full citation. Permission from the Editor is required to post to servers, redistribute to lists, or utilize in a for-profit or commercial use. Permission requests should be sent to the Editor-in-Chief, ***Journal of Information Systems Education***, editor@jise.org.

ISSN: 2574-3872 (Online) 1055-3096 (Print)