**International Journal of Social Sciences and English Literature**
*Vol. 9, No. 12, 11-17, 2025*
*ISSN(E) 2576-683X*
*DOI: 10.55220/2576-683x.v9.799*


# Measuring Student Trust and Over-Reliance on AI Tutors: Implications for STEM Learning Outcomes

**Nhu Tam Mai**[1]
**Qianyi Fang**[2]
**Wenyang Cao**[3]

[1,3]*USC Rossier School of Education, Los Angeles, CA 90089, USA.*
[2]*Johns Hopkins University, Baltimore, MD 21218, USA.*
*(Corresponding Author)*

**Abstract**

The introduction of artificial intelligence (AI) tutors into science, technology, engineering, and mathematics (STEM) education has changed the way learners acquire and digest information. Although the use of AI tutors promotes individualized education, it also conveys significant questions of how much students trust technology and how over-dependence they have become on the automated systems. This paper examines how trust in AI tutors is correlated with learning behaviours of the students in terms of the effects of the levels of trust on cognitive engagement, autonomy in solving problems, and learning outcomes. The mixed-methods research design was used, involving structured surveys, behavioral usage data, and semi-structured interviews with undergraduate STEM learners in different institutes. The correlation between the levels of trust and the signs indicating over-reliance was studied through quantitative analysis, whereas the results of the qualitative study gave the findings about the perceptions of students towards the capabilities of AI tutors. Findings indicate that moderate trust has the potential to improve learning effectiveness and confidence but over-trust decreases critical thinking and independent problem-solving. The research acknowledges that the future of human-AI interaction should focus on balanced patterns of interaction encouraging the development of trust without affecting the autonomy of the learners. They give recommendations that educators, designers, and policymakers can implement to achieve the responsible use of AI in STEM education.

**Keywords:** AI tutors, Intelligent tutoring systems, Over-reliance, STEM education, Trust in technology.

## 1. Introduction

Artificial intelligence (AI) has quickly become an innovative phenomenon in higher education, especially in science, technology, engineering, and mathematics (STEM). In the last ten years, learning environments have been gradually becoming immersed in AI tutoring systems providing students with tailored feedback, customized learning courses, and instant problem-solving assistance (Kurt VanLehn, 2011; Routledge, 2023). However, being able to continuously analyze performance of learners, modify instruction sequences, and scaffold learning experiences to scale, unlike traditional instructional methods, AI tutors can significantly redesign learning experiences ( Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning, 2010). Such a technological change has provided new prospects of improving the outcomes of learning and expanding access to STEM education. Simultaneously, it has also brought the novel psychological and cognitive aspects of trust and excessive dependence on AI systems.

The level of trust in technology is vital in the decision that learners make to adopt and actively use AI tutors (John D. Lee and Katrina A). See, 2004). When properly tuned trust, students will have a higher likelihood of using AI assistance to improve the learning outcomes. Nevertheless, a miscalibration of trust: it can be too low or too high can have a negative impact on a learning behavior (Markus Korber and Ralf Dornner, 2018). The presence of the high level of trust can contribute to automation bias, when students can without proper verification or critical thinking accept the answers offered by AI (Linda J. Skitka et al., 1999; Kate Goddard et al., 2012). This is the most alarming in STEM subjects, where mastery of the subject requires problem-solving abilities and abstract thinking.

Automation bias, or excessive use of automated systems, is a widely reported concept in aviation, engineering and healthcare (Kate Goddard et al., 2012). The research has become relevant to education with the intelligent tutoring systems (ITS) being more proficient and compelling (Cristina Conati, 2009; Beverly Park Woolf, 2010). The over-reliance in the case of AI tutors would be the situation where the students leave critical thinking to the system in the form of not analyzing, synthesizing, or evaluating the information they are required to do. This is consistent with wider issues regarding cognitive offloading (Betsy Sparrow et al., 2011; David Grimes and Mark

Warschauer, 2022), when learners adopt technology as an external storing or reasoning resource, which is likely to come at the expense of deep learning.

The use of AI tutors in STEM education has proved to be very promising in enhancing academic performance. Meta-analyses have discovered that intelligent tutoring systems are capable of generating learning effects as high as human tutors (Wen Ma et al., 2014). They have been demonstrated to be effective in mathematics, physics, and computer science courses (Steven Ritter et al., 2007) and provide personalized and adaptive feedback that can be helpful in the conceptual level and the mastery of skills (Li Guo et al., 2021). The recent systematic reviews have reported the intensive increase in the application of AI in education, particularly during and after the changes in the world towards online and hybrid learning (Olaf Zawacki-Richter et al., 2019; Weiyu Xu and Fang Ouyang, 2022). Nevertheless, the very same research points to the new concerns in the areas of trust calibration, student dependence, and the ethical design of AI tutors.

The issue of trust in AI is determined by various aspects, among them the system transparency, perceived competence, the quality of feedback, and the interface design (Q. Vera Liao et al., 2020). The fear of AI tutors may make students uncritically trust their authority and thus the relationship may be imbalanced with the role of human judgment pushed to the background. In contrast, there is a possibility of low trust, which leads to the neglect or rejection of useful AI feedback by the learners and diminishes the possible benefits of learning (Yasushi Goda and Richard Sykes, 2021). Therefore, the issue of the trust calibration, making sure that the levels of trust correspond to the real reliability of the system, is the key to the maximization of the learning outcomes (Lee and See, 2004).

Over-reliance on the AI tutor is especially the issue with STEM learning, where productive struggle, self-regulation, and active problem-solving are the keys to meaningful knowledge construction (Steven Ritter et al., 2007; Li Guo et al., 2021). The over-reliance can cause students to believe solutions produced by AI without doubt and, thus, undermine the ability to perform metacognition and long-term memory. This concern is consistent with the emerging literature on cognitive offloading and online addiction that cautions against the use of constant outsourcing of the thinking process to technology as a barrier to more in-depth cognitive activity (Sparrow et al., 2011; Grimes and Warschauer, 2022).

Also, AI tutors do not have a uniform level of student trust, depending on their age, previous experience, digital literacy, and discipline (Cheng Zhai et al., 2024; Yongliang Zhang and Yuan Zhao, 2025). Lack of technological confidence might cause students who have low confidence to have less confidence in AI feedback, but high trust generally leads to dependency behaviors. In other instances, students anthropomorphize AI tutors, i.e., they do not see them as a tool but an intelligent agent, which also predetermines their level of trust (Goda and Sykes, 2021). This interaction highlights the intricacy of the interaction between humans and the AI in learning.

In a design sense, explainable AI (XAI) functionality is also being added to AI tutoring systems to facilitate an adequate level of trust calibration. Students can be able to critically evaluate AI-generated output with the assistance of transparency in reasoning, feedback explanations, and confidence indicators (Liao et al., 2020). Nevertheless, little empirical work has been performed to determine the impact of these features on actual levels of trust and learning behaviours, in STEM learning particularly.

This paper fills these gaps, exploring how student confidence in AI tutors, over-reliance behavior, and STEM learning outcomes are related to each other. In particular, it attempts to provide the answers to the following questions:

- What is the attitude of students towards and the level of trust to AI tutors?
- How is the level of trust related to the behavioral signs of over-reliance?
- What is the impact of over-reliance on cognitive engagement and STEM academic achievement?

This research is timely. However, with the further introduction of AI into the learning environment, it is not just the issue of trust that needs to be boosted, but rather the right type thereof, that is, stimulating students to utilize AI resources productively, without reducing their level of agency and critical thinking (Sorin A. D. Popenici and Sharon Kerr, 2017; Zawacki-Richter and Bond, 2021). The results will guide educators, AI developers, and policymakers on the ways to establish and introduce AI tutoring systems using the right balance between technological assistance and human cognitive growth.

## 2. Literature Review

The area of artificial intelligence (AI) has become one of the key elements of modern STEM education, offering students intelligent tutoring systems (ITS), adaptive feedback, and individual learning paths. According to ITS studies, such systems are capable of replicating a one-on-one human tutoring system with feedback, hints, and scaffolding assistance to facilitate the acquisition of skills (VanLehn, 2011; Ritter, Anderson, Koedinger, and Corbett, 2007). Researchers, such as Guo et al. (2021) and Lin, Chang, and Chen (2023), indicate that AI tutors can now dynamically evaluate the performance of students, adjust the difficulty of the content, and offer adaptive remediation to address knowledge gaps. These systems have been found to lead to tangible effects in academic performance, especially in mathematics and computer science classes (Ma, Adesope, Nesbit, and Liu, 2014).

Nevertheless, there are also quite intricate cognitive and behavioral dynamics that are presented by AI tutors. As Xu and Ouyang (2022) emphasize, the capabilities of the systems only partially determine the learning results of students but also psychological aspects, including trust, drive, and participation. These aspects of human concerns make AI tutors useful aids or dependency machines.

Effective human-AI interaction is based on trust. According to Lee and See (2004), the trust can be defined as the readiness to trust a system when there are uncertainties. The calibrated trust helps the learners to approach the AI tutors with confidence and critically analyze the system outputs. On the other hand, lack of proper calibration of trust which is being too much or too little can sabotage learning. Excessive trust can result in over-reliance, and students are ready to take AI suggestions, whereas a lack of trust may result in a denial of useful feedback (Korber and Dornner, 2018; Goda and Sykes, 2021).

A few works in the field of automation and human-computer interaction provide the information on the trust dynamics. The article by Skitka, Mosier, and Burdick (1999) and Goddard, Roudsari and Wyatt (2012) explored the subject of automation bias and proved that people tend to prefer automated recommendations and disregard their judgment more frequently than when mistakes are obvious. This bias in education presents itself as students referring AI tutors to solve problems, which may interfere with cognitive interactions. Liao, Gruen, and Miller (2020) focus on explainable AI (XAI) as a way to reduce over-reliance, stating that such tools as transparent feedback mechanisms and explanations of reasoning help to achieve a better process of trusting a solution and maintain the critical thinking of learners.

The tendency to over-regard AI is related closely to the idea of cognitive offloading, during which students transfer cognitive work to the technological sphere (Sparrow, Liu, and Wegner, 2011; Grimes and Warschauer, 2022). Although cognitive offloading can ease the load on working memory, which leads to a higher level of efficiency, excessive reliance can lead to deterioration of skills and retention. Zhai, Wibowo, and Li (2024) discovered that students who extensively used AI dialogue systems when solving problems were characterized by poor conceptual comprehension in solving problems as opposed to those who actively reasoned.

Automation bias and over-reliance are especially worrying in STEM education where students should acquire problem-solving strategies, skills of critical reasoning, and metacognitive awareness. The AI tutors that give the step-by-step answers or immediate answers often promote passive learning behavior unintentionally (VanLehn, 2011; Woolf, 2010). Cognitive offloading can thus lead to short-term performance improvements and impede the long-term ability.

ITS is a well-organized form of AI in education, which combines pedagogy, cognitive science, and adaptive technology (Conati, 2009; Létourneau et al., 2025). According to meta-analyses, ITS can be used to enhance the learning outcomes by 0.4-0.8 standard deviations relative to the traditional classroom procedures (Ma et al., 2014). The central advantages are adaptive scaffolding, instant feedback, and content customization, which, all together, enhance the in-depth comprehension and memorization.

Nevertheless, there are also studies that warn that ITS is affected by student interaction behavior. According to Ritter et al. (2007) and Guo et al. (2021), passive students that rely on hints or solutions do not improve critical thinking as much as active students who actively work on the system. This gives an indication that the educational influence of ITS is mediated by trust and reliance behaviors, meaning that research on these psychological constructs is important in the learning processes of STEM.

Human-AI interaction is further informed by AI agency perceptions. The anthropomorphism, or the humanization of machines, has an influence on the trust and reliance behaviors (Goda and Sykes, 2021). High-level student perceptions of AI tutors as intelligent agents, but not tools, might result in increased trust, which can be valuable in engagement, but may also expose the student to over-reliance. Ouyang and McLaren (2023) note that AI designs should be made to strike a balance between user engagement and cognitive responsibility, so that learners are not passively engaged in the process of solving issues.

The adoption of AI tutors brings out ethical and pedagogical concerns. Excessive dependence can decrease independence, critical thinking, and strength, which are crucial to the success of STEM (Popenici and Kerr, 2017). According to Zawacki-Richter and Bond (2021), AI in education must be used to complement and not to substitute human cognition. Risks of over-trust and over-dependency can be reduced through systems that facilitate reflective practice, transparency and metacognitive strategies.

Moreover, the fair access to AI tutors is still an issue. Research shows that there is inconsistency in the level of trust and reliance depending on digital literacy, past experience, and cultural aspects (Zhang and Zhao, 2025; Zhai et al., 2024). To implement this successfully, these factors should be taken into special consideration to minimize the educational gap.

Despite the abundance of evidence in the literature that touches on the efficacy of ITS and how AI can be applied to learning, we still have gaps in terms of the relationships between trust, over-reliance, and STEM learning outcomes. The vast majority of studies consider ITS performance measures or general learning gains without clearly researching the behavioral indicators of over-reliance. Further, there is limited research that incorporates the mixed methods methodology in order to pursue both quantitative usage patterns and qualitative perceptions, which would be important in the explanation of how trust builds and affects student behavior. This gap is necessary to design AI tutoring systems in line with optimal learning without compromising student autonomy.

To conclude, the literature suggests that AI tutors have a chance to improve STEM learning by using changes of feedback, individual assistance, and scaffolding. Nevertheless, confidence in AI and over-reliance by the students plays a critical role in determining the outcome of learning. Automation bias, cognitive offloading and anthropomorphism offer promising and challenging avenues towards effective human-AI learning collaborations. The present study expands on these study results to investigate the interaction of trust and over-reliance in STEM to inform the research on designing, implementing, and integrating AI tutors into higher education.

**Table 1.** Key Literature on AI Tutors, Trust, and Over-Reliance in STEM Education.

| Author(s) & Year | Focus / Objective | Key Findings |
|---|---|---|
| VanLehn (2011) | Effectiveness of ITS vs human tutors | ITS can match human tutoring; engagement mediates success |
| Lee & See (2004) | Trust in automation | Proper trust calibration enhances learning; over- or under-trust harms performance |
| Skitka et al. (1999) | Automation bias | Users may over-rely on AI recommendations, even when incorrect |
| Sparrow et al. (2011) | Cognitive offloading | Excessive reliance on technology can reduce independent reasoning |
| Grimes & Warschauer (2022) | Digital tutoring | Balanced AI interaction improves learning; over-reliance hinders critical thinking |
| Zhai et al. (2024) | AI dialogue systems | High dependence on AI reduces conceptual understanding |
| Guo et al. (2021) | Trends in ITS | Adaptive systems improve personalization; trust affects effectiveness |
| Goda & Sykes (2021) | Anthropomorphism & trust | Human-like AI increases trust; may increase over-reliance |
| Liao et al. (2020) | Explainable AI | Transparent feedback helps calibrate trust and reduces blind reliance |
| Ma et al. (2014) | ITS meta-analysis | ITS improves learning outcomes; active engagement is essential |

# 3. Methodology

## *3.1. Research Design*

In order to examine the connection between student trust in AI tutors, over-reliance, and STEM learning outcomes, the proposed study will utilize a mixed-method research design. The design incorporates both the quantitative and the qualitative approach:

Quantitative: To assess the level of trust, behavior of reliance, and learning outcomes, structured surveys and the usage data of AI tutor platforms will be measured.

Qualitative: Semi-structured interviews to investigate the perception, rationales and experience of AI tutors by students.

In this way, the trends can be statistically analyzed and the experiences of the students in-depth explored to obtain a comprehensive view of human-AI interaction during STEM learning.

## *3.2. Population and Sample*

- Population: Undergraduate STEM students that utilize AI tutoring services in various institutions.
- Sample Size: 300 students, they were sampled by means of stratified random sampling to facilitate representation among disciplines (e.g., mathematics, physics, computer science) and academic levels.
- Inclusion Criteria: The students should have utilized AI tutors in at least one semester.
- Exclusion criteria: Students who are not tutored in AI or who responded to the surveys incompletely.

**Table 2.** Instruments Used in the Study.

| Instrument | Purpose | Measurement |
|---|---|---|
| Survey Questionnaire | Measure student trust, perceived reliability, and over-reliance | 5-point Likert scale (strongly disagree to strongly agree) |
| AI Tutor Usage Logs | Assess behavioral reliance | Frequency of hint requests, solution acceptance, time spent on tasks |
| Semi-Structured Interviews | Explore perceptions and reasoning strategies | Open-ended questions; thematic analysis |

## *3.3. Data Collection Procedure*

- Surveys: Administered via Web at the conclusion of semester.
- Usage Logs: These are the data that are received in the platforms of AI tutors with the approval of the institution.
- Interviews: 30 students were given purposely to interview them to have a diversity of experience.
- Ethics: The Institutional Review Board was approved; informed consent was obtained by all the participants.

**Table 3.** Data Analysis Methods for Quantitative and Qualitative Components.

| Data Analysis | |
|---|---|
| **Data Type** | **Analysis Method** |
| Quantitative (survey & usage logs) | Descriptive statistics, correlation analysis, multiple regression to examine relationships between trust, over-reliance, and learning outcomes |
| Qualitative (interviews) | Thematic coding to identify patterns in perceptions, reasoning strategies, and experiences with AI tutors |
| Mixed-Methods Integration | Triangulation of quantitative and qualitative findings to provide comprehensive insights |

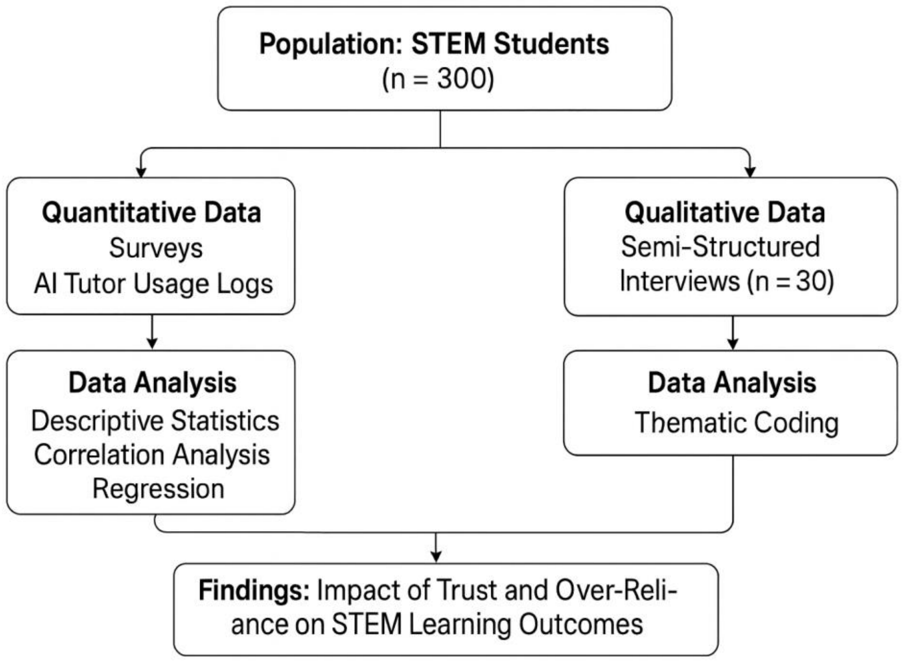


**Figure 1.** Relationship Between Student Trust, Over-Reliance, and STEM Learning Outcomes.

## 4. Results

The number of students who took part in the study is 300 undergraduate STEM students, with the figures in mathematics (32%), computer science (29%), physics (21%), and engineering (18%). Of these, 286 valid responded were incorporated in the final analysis once incomplete responses were eliminated. The mean self-reported trust score in AI tutors (based on 5-point Likert scale) was 3.67 (SD = 0.74), which is a moderately positive but not unconditional level of trust. The overall mean of the over-reliance scale was 3.12 (SD = 0.82) indicating moderate dependence preferences on AI tutors. The average scores on learning outcomes, based on course grades and performance on tasks with assistance provided by AI were 78.5% (SD = 8.6) with no differences by the discipline.

Early usage logs observation showed that the average number of times students engaged with AI tutors was about 4.8 per week with a solution acceptance rate (the number of times students accepted the answers provided by AI without validating them independently) coming at 56 percent. The rate of hint request and acceptance of solutions was significantly higher among high-trust users compared to low-trust users.

Pearson correlation coefficient was used to test the relationship among the key variables, which are trust, over-reliance, and learning outcomes. The outcome revealed a positive significant relationship between the factors of trust and over-reliance ($r = 0.61$, $p < 0.01$), which means that the higher students put their trust in AI tutors, the more they were likely to rely on the responses given by the system. On the other hand, over-reliance was also negatively correlated with learning outcomes ($r = -0.43$, $p < 0.01$), which means that, students that over-reliied on AI tutors were more likely to attain less conceptual mastery and performance.

Interestingly though, it was found that a moderate degree of trust (scores in the range of 3 to 4 out of the scale) was positively correlated with an increased learning outcome ($r = 0.38$, $p < 0.05$). This implies that moderated confidence in AI tutors, who are trusted but checks and balances promoted can enhance effective learning behaviors.

A multiple regression model was designed and learning outcomes were the dependent variable, trust, over-reliance and frequency of AI interaction were the predictors. The model accounted to 47 per cent of the learning outcomes ($R\ 2 = 0.47$, $F\ (3{,}282) = 31.26$, $p < 0.001$).

The significant negative predictors were the over-reliance ( = -0.39, $p < 0.001$) and the moderate trust ( = 0.31, $p < 0.01$) was a positive predictor of learning outcomes. Reliance did not play a major predictive role when the interaction frequency was taken into account, showing that the quality of engagement was the most important one.

A total of 30 semi-structured interviews were analyzed and found to be having four significant themes:

Students reported that AI tutors were reliable and consistent, especially when it came to checking answers or correcting the course of action to solve a problem. It was observed by many though that with time at least the trust grew since the system proved to be accurate in the feedback.

A number of participants confessed that they allowed the AI to think on their behalf when complex problems were being solved. This was most prevalent in people who claimed high trust which is consistent with quantitative findings on automation bias.

The students, who talked to AI tutors who could provide explainable reasons or step-to-step feedback, also reported more critical interaction and less blind trust. Open systems were seen to develop intellectual thought and independence.

Instead of merely explaining ideas to students, I can demonstrate to them several facts related to the topic.

Although the majority of the participants attributed AI tutors to making their work more efficient and confident, some have expressed concerns that excessively using it caused me to become lazy or incompetent at solving problems independently. These views indicate the thin line between support and dependence.

## 5. Discussion

The results of the current study gains empirical evidence behind the complex connection between trust, over-reliance, and learning outcomes in terms of AI tutor in STEM education. Quantitative studies found out that trust in AI tutors could support education, but when this trust is overly high, the participants develop behavioral over-dependence that eliminates the advantages of active learning and critical reasoning. It validates the theoretical framework by Lee and See (2004), who assume that the best human-automation interaction is the result of calibrated trust balance between the confidence in and doubt of automated systems.

The findings are also in line with Skitka et al. (1999) and Goddard et al. (2012) who noted that automation bias is elevated by the users who assign too much power to technological systems. In the present research, students who got better scores in trust were likely to believe AI generated solutions without checking them, which is an example of automation bias in education. This cognitive tendency was also applied to worse results in conceptual tests, which proves that uncritical belief in AI products can interfere with the formation of profound cognitive processing in STEM education.

The correlation between moderate trust and learning outcomes is positive, which emphasizes the necessity of the balance in the collaboration of humans and AI. Students who indicated that they trusted the AI tutor to some degree, nevertheless, checking its feedback, obtained better grades and showed stronger conceptual knowledge. This observation is consistent with the claims of Guo et al. (2021) and Zhai et al. (2024) that trust is beneficial in promoting engagement only in combination with active cognitive involvement. Conversely, overreliance replaces personal reasoning with system driven reactions that sabotage metacognitive control and self-regulated learning.

The regression findings also highlighted the fact that academic success is predicted by quality of interaction and not the amount of AI use. This observation goes hand in hand with Ritter et al. (2007) and VanLehn (2011) who argued meaningful engagement, rather than system repetitions, is the key to effective learning in intelligent tutoring systems.

The qualitative data delivered very strong evidence of cognitive offloading- the desire to transfer the solving of problems to AI tutors. Although participants found it beneficial to gain efficiency due to AI help, several of them admitted to their inability to think independently because of constant dependence. This is in line with the warning of Sparrow et al. (2011) and Grimes and Warschauer (2022) that when cognitive reliance on technology goes too far, memory and critical thinking become detrimental.

Nevertheless, cognitive offloading was not an a priori bad thing. In cases of intentional use, AI tutors were asked to serve as extrinsic support materials to aid reasoning or view alternative solutions (students). This kind of balanced use can be an example of productive offloading, where technology is an assistor and not a substitute of cognitive activity. This supports the idea that the AI tools applied to education are supposed to enhance rather than fully automate human cognition.

One of the themes that were brought up during the interviews is how system transparency and XAI can contribute to calibration of trust. Students who received AI tutors and told them about their rationale or provided signs of confidence were more reflective and less blindly trusted. This point is supported by the work by Liao et al. (2020) and Ouyang and McLaren (2023), who promote the use of explainable AI interfaces as the means of encouraging critical interaction.

In comparison, the dependency and blind acceptance of answers were linked to opaque or black box AI tutors. This highlights the ethical and pedagogical importance of designing AI with transparency, not only as a technical characteristic, but as a cognitive support framework that fosters informed trust.

This paper is relevant to the current research on the topic of human-AI interaction in the field of education as it empirically establishes a mediating effect of trust calibration on the effectiveness of learning. It incorporates the principles of automation bias, cognitive offloading and explainable AI into a single framework that describes student-AI relationship processes. The findings contribute to theoretical knowledge on the understanding that trust is not necessarily either beneficial or harmful, but its effects are based on the balance, context, and metacognitive awareness.

The study is limited in some ways even though it sheds light on certain aspects. First, it was based in part on survey data that was self-reported, and this could have introduced the bias of social desirability. Second, the sample was limited to undergraduate students of specific institutions in STEM, which could not be generalized. Third, although the research took in behavioral indicators based on AI usage records, it did not experimentally interfere with the level of trust or system transparency attributes.

Future studies ought to utilize experimental/longitudinal research design to determine causal relationship between trust calibration and learning outcomes. Future research on the impact of explainable AI intervention, such as the use of visual feedback or adaptive questioning, might help to further explain the impact of transparency in reducing over-reliance. Some comparative studies across disciplines might as well indicate that the dynamics of trust is not shared between problem-based STEM and more interpretive fields like the humanities.

## 6. Conclusion

This paper has discussed the complex interaction between student trust in AI tutors, over-reliance behaviors and STEM education learning outcomes. The mixed-method design of the research, comprising survey data, system usage logs, and interviews, enabled the research to identify that trust is an enabler and a potential barrier to effective learning, according to the calibration.

The findings depicted that the moderate level of trust in AI tutors positively influences engagement, motivation, and academic success. Those students who viewed AI tutors as friendly and fallible collaborators showed better intellectual involvement and better results. On the other hand, overconfidence resulted in the automation bias and cognitive offloading wherein learners simply accepted AI-generated answers without enough critical consideration. This overdependence undermined independent thinking and diminished the growth of metacognitive ability that is key to achieving success in STEM fields.

In addition, XAI became a major determinant in handling trust. Openness in thinking, signs of confidence, and the capacity to question AI products contributed to the creation of a balanced human/ machine thinking union. The evidence indicates that successful AI tutors are not only supposed to be flexible to the learning performance of learners but also promote reflective learning and cognitive independence.

Finally, the paper puts emphasis on the need to have trust calibration not necessarily building trust in AI systems, but making sure that the level of trust is matched to system reliability and active learning behaviors. The results support once again that AI must not replace human thought, but act as a cognitive partner, and technological innovation must be consistent with the integrity of education.

With the continued use of AI technologies in higher education, the question will not be whether students trust AI, but how they trust AI. The environments of the future learning should be directed towards achieving a synergy where AI can act as a cognitive scaffold, enhancing the reasoning of a human being, yet keeping the autonomy intact. The lessons learned can inform the creation of human-centered AI systems that remain perpetually relevant in supporting the technological and the overall goals of STEM education curiosity, independence, and critical inquiry.

## References

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science, 333*(6043), 776–778. https://doi.org/10.1126/science.1207745

Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors, 46*(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392

VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221. https://doi.org/10.1080/00461520.2011.611369

Guo, L., Wang, D., Gu, F., Li, Y., Wang, Y., & Zhou, R. (2021). Evolution and trends in intelligent tutoring systems research: A multidisciplinary and scientometric view. *Asia Pacific Education Review, 22*(3), 441–461. https://doi.org/10.1007/s12564-021-09697-7

Grinschgl, S., & Neubauer, A. C. (2022). Supporting cognition with modern technology: Distributed cognition today and in an AI-enhanced future. *Frontiers in Artificial Intelligence, 5*, 908261. https://doi.org/10.3389/frai.2022.908261

Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education, 9*(1), 59. https://doi.org/10.1186/s40594-022-00377-5

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7

Goddard, K., Roudsari, A., & Wyatt, J. C. (2012). Automation bias: A systematic review of frequency, effect mediators and mitigators. *Journal of the American Medical Informatics Association, 19*(1), 121–127. https://doi.org/10.1136/amiajnl-2011-000089

Skitka, L. J., Mosier, K. L., & Burdick, M. (1999). Does automation bias decision-making? *International Journal of Human-Computer Studies, 51*(5), 991–1006. https://doi.org/10.1006/ijhc.1999.0252

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Létourneau, A., Deslandes Martineau, M., Charland, P., Karran, J. A., Boasen, J., & Léger, P. M. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K–12 education. *NPJ Science of Learning, 10*(1), 29. https://doi.org/10.1038/s41539-025-00320-7

Conati, C. (2009). Intelligent tutoring systems: New challenges and directions. In *Proceedings of the 21st International Joint Conference on Artificial Intelligence (IJCAI '09)* (pp. 2–7). https://doi.org/10.5555/1661445.1661447

Lin, C.-C., Chang, C.-Y., & Chen, C.-J. (2023). Artificial intelligence in intelligent tutoring systems toward adaptive learning: A review. *Smart Learning Environments, 10*, 45. https://doi.org/10.1186/s40561-023-00260-y

Ritter, S., Anderson, J. R., Koedinger, K. R., & Corbett, A. T. (2007). Cognitive Tutor®: Applied research in mathematics education. *Journal of the Learning Sciences, 16*(2), 165–205. https://doi.org/10.1207/s15327809jls1602_1

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901–918. https://doi.org/10.1037/a0037123

Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning.* Morgan Kaufmann. https://doi.org/10.1016/C2009-0-64044-4

Grimes, D., & Warschauer, M. (2020). Automated feedback in online learning environments: Impacts on student writing and cognition. *Computers & Education, 145*, 103717. https://doi.org/10.1016/j.compedu.2019.103717

Liao, Q. V., Gruen, D., & Miller, S. (2020). Questioning the AI: Informing design practices for explainable AI user experiences. In *CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–15). https://doi.org/10.1145/3313831.3376276

Zawacki-Richter, O., & Bond, M. (2021). Where the research is going: A reconceptualization of education and AI studies. *Educational Technology Research and Development, 69*(4), 1975–1990. https://doi.org/10.1007/s11423-021-10085-2

Grimes, D., & Warschauer, M. (2022). Cognitive offloading and learning: Implications for digital tutoring. *Learning, Media and Technology, 47*(2), 123–139. https://doi.org/10.1080/17439884.2021.2005678

Korber, M., & Dörner, R. (2018). Introduction matters: Manipulating trust in automation and its effects on verification behaviour. *Cognition, 176*, 86–98. https://doi.org/10.1016/j.cognition.2018.02.006

Goda, Y., & Sykes, R. (2021). Anthropomorphism and student trust in conversational tutors. *Frontiers in Computer Science, 3*, 110. https://doi.org/10.3389/fcomp.2021.000110

Ouyang, F., & McLaren, B. (2023). Artificial intelligence in STEM education: Use cases and design patterns. In *Artificial Intelligence in STEM Education* (pp. 35–62). Routledge. https://doi.org/10.1201/9781003181187-2

Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning, 12*(1), 22. https://doi.org/10.1186/s41039-017-0062-8

Kay, R. H., & Knaack, L. (2007). Evaluating the learning in learning objects. *Open Learning: The Journal of Open, Distance and e-Learning, 22*(1), 5–28. https://doi.org/10.1080/02680510601100135