International Journal on Science and Technology (IJSAT)
E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

Enhancing Trust in AI-Powered Learning: The Role of Explainable AI in the Education Field

K. Swathi[1], Rampilla Yuvatez[2], Satyala Kovid[3]

[1,2,3] Computer Science and Engineering (CSE), Koneru Lakshmaiah Educational Foundation, Guntur, India.

Abstract

The integration of Artificial Intelligence (AI) in education has shown immense potential in personalizing learning, automating assessments, and supporting educators in decision-making. However, the opaque nature of many AI models raises concerns about trust, fairness, and accountability, particularly in high-stakes educational contexts. This paper explores the role of Explainable AI (XAI) in enhancing trust and transparency within AI-powered learning systems.

The purpose of this study is to examine how explainability can bridge the gap between AI-driven decision-making and user acceptance in education by enabling students, teachers, and administrators to better understand AI recommendations. The methods involve a systematic literature review of existing XAI techniques applied in educational settings, supported by case analyses of AI-powered learning tools that integrate interpretability features. Surveys and feedback from educators and learners are also examined to capture perceptions of trust and usability.

The results indicate that XAI not only improves transparency and accountability but also enhances user trust and adoption of AI tools. In particular, interpretable feedback on student performance predictions and adaptive learning paths fosters confidence among learners, while educators gain actionable insights for personalized teaching. Moreover, the presence of XAI reduces bias-related concerns and improves overall acceptance of AI in classrooms.

The conclusion underscores that explainability is not merely a technical feature but a pedagogical necessity in ensuring responsible and trustworthy AI integration in education. Future directions highlight the need for hybrid approaches combining technical explainability with user-centered design, fostering both ethical and effective AI-driven learning environments.

Keywords: Artificial Intelligence in Education; Trust and Transparency; Personalized Learning; Adaptive Learning; Ethical AI; User Acceptance; Accountability

1. Introduction

In recent years, Artificial Intelligence (AI) has increasingly played a significant role in education. It is widely used in learning applications and tools to provide students with personalized learning, timely feedback, and to assist teachers in understanding student progress. However, a major challenge is that AI

often functions as a "black box," producing results or recommendations without clearly showing how or why they were generated. This lack of transparency makes it difficult for students and educators to fully trust AI systems. Modern approaches in Explainable AI (XAI), such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations), address this issue by revealing how complex AI models arrive at their decisions, making predictions more interpretable and trustworthy.

Explainable AI (XAI) is an important solution to this problem. It helps AI systems to show the reasons behind their answers in simple and clear ways. For example, if an AI says a student needs to study more, it can also explain which topic the student is weak in. This builds more trust [1] between students, teachers, and AI systems.

Artificial Intelligence (AI) is becoming an important part of our daily life, and education is no exception. Many schools, colleges, and learning platforms [2], are now using AI tools to make learning faster and easier. These tools can give students personalized lessons, check their answers quickly, and even guide teachers by showing which students need more help. AI can also save time for teachers by doing tasks like grading or preparing study plans.

However, there is a big challenge with AI. Most AI systems do not explain how they make their decisions [6]. For example, if an AI says a student is weak in mathematics, it usually does not explain why it came to that conclusion. This makes students and teachers confused and sometimes worried. If they do not understand the reason behind AI's suggestions, they may not trust it completely. Trust is very important in education because wrong or unfair decisions can affect a student's learning and confidence.

This is where Explainable AI (XAI) becomes very useful. Explainable AI means AI systems that can clearly show the reasons for their results. For instance, if the AI says a student should practice more in science, it will also explain which topics in science are causing problems which topic the student is weak [3],[5]. This way, students can improve in the right areas, and teachers [4] can make better decisions. XAI makes AI more open, fair, and easier to accept.

But there is also a problem. Most AI systems only give answers or suggestions without showing why. They work like a "black box," where we can see the result but not the process [7]. For example, if an AI system says, "This student may fail in the exam," it does not tell the student or teacher the reasons. This makes it hard for people to fully trust the AI. In education, trust is very important because it directly affects the future of students. If students or teachers cannot understand why the AI made a decision, they may ignore it, misuse it, or even lose confidence in it.

This is where Explainable AI (XAI) becomes very important. Explainable AI is a type of AI that not only gives results but also explains them in a simple way. For example, instead of only saying "You are weak in math," XAI can say, "You made more mistakes in fractions and geometry, so you need extra practice in those topics." This explanation helps students know exactly where they need to improve, and it helps teachers make better choices for their teaching methods.

XAI makes AI systems more clear, fair, and easy to trust. It can also reduce bias, because if an AI is unfair, explainability helps to find and correct the problem. In this way, XAI can make education more honest and supportive [8].

This paper explores how Explainable AI (XAI) can build trust in education, how it can help students and teachers, and why explainability is an important step for using AI in a safe and responsible way in learning [9].

2. AI in Education: Current Uses and Benefits

Artificial Intelligence (AI) is slowly becoming an important part of education all over the world. It is not only used in advanced research but also in everyday classrooms and online learning platforms. One of the main uses of AI in education is personalized learning. Unlike traditional teaching where every student is taught in the same way, AI can study the progress of each student [10] and provide lessons according to their needs. For example, if a student is weak in fractions, the AI system will give more practice problems on that topic. This makes learning more effective because every student gets what they actually need instead of following a fixed path.

Another important use of AI is in grading and assessment. Checking tests, assignments, and projects is a time-consuming task for teachers. AI tools can quickly grade multiple-choice questions, short answers, and even some essays [11]. This helps teachers save time and focus more on teaching rather than spending long hours on correction. At the same time, students benefit because they receive instant feedback about their performance. This quick feedback helps them identify mistakes and improve faster.
AI also provides virtual tutors and chatbots [12] that can answer students' questions at any time. For example, if a student is studying late at night and gets stuck, they can use an AI-powered tutor to get explanations. This makes learning more flexible and independent, as students do not have to wait for the next class to clear their doubts.

For teachers, AI offers strong support by analyzing large amounts of student data. It can highlight which students are doing well and which students are struggling [13]. With this information, teachers can give extra attention to weaker students and adjust their teaching methods. This makes classrooms more inclusive and fair.

In today's classrooms and online platforms, technology is becoming an important part of the learning process, and Artificial Intelligence (AI) is leading this change. One of the biggest contributions of AI is the idea of personalized learning. Instead of following the same lesson plan for everyone, AI studies how each student learns and then suggests topics, exercises, and resources that fit their level. This makes learning more focused and ensures that students get the right kind of support where they need it most.

A second area where AI is making a difference is in grading and assessment. Checking tests and assignments is one of the most time-consuming tasks for teachers. With AI tools, much of this work can be done automatically. Students can also see their mistakes immediately through instant feedback, which helps them learn from errors without delay. This quick response is very valuable for keeping students engaged and motivated.

Another growing benefit of AI is the availability of virtual tutors and chatbots. These tools work like assistants that are always available, even outside classroom hours. A student studying late at night can ask the AI a question and get a clear explanation right away. This reduces dependency on teachers for every small doubt and gives students the freedom to learn at their own pace.

AI is not only useful for learners but also provides valuable insights for teachers. By analyzing student performance data, AI can show which topics are difficult for most students, which students need extra attention, and which teaching methods are most effective. This information helps teachers adjust their lessons and support each student more effectively.

Apart from these uses, AI also helps in creating digital study materials. Smart systems can generate quizzes, flashcards [12], and practice questions based on the current syllabus. Some tools can even turn a textbook into an interactive learning experience with videos, summaries, and examples. This makes education more engaging and easier to access for students everywhere.

When all these uses are combined, the result is a learning system that is more efficient, flexible, and student-friendly. Students receive learning experiences suited to their needs, while teachers save time and gain tools to improve their teaching. The overall benefit is that AI creates an environment where education becomes more inclusive, fair, and effective for everyone.

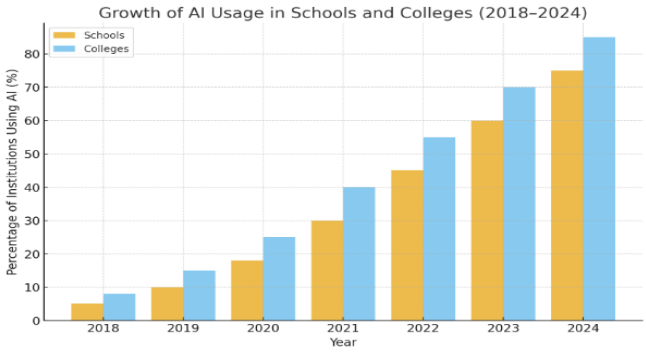

Fig. 2.1: The graph illustrates the steady rise in AI usage in schools and colleges from 2018 to 2024, with colleges adopting AI at a faster rate than schools.

3. Challenges of Using AI in Education

While Artificial Intelligence (AI) is bringing many benefits to education, it also comes with several challenges. One of the biggest problems is the lack of transparency. Many AI systems work like a "black box," which means they give results without explaining how those results were made. For example, if an AI system predicts that a student may fail, neither the student nor the teacher can easily understand the reason behind it. This lack of explanation creates confusion [14], and reduces trust in AI systems.

Another issue is the risk of bias. AI systems are trained on data, and if that data has errors or unfair patterns, the AI will also give unfair [15] results. For instance, if the training data is more focused on students from one background, the system might not give accurate results for students from other backgrounds. This can create inequality in education, which is very harmful because every student should be treated fairly.

High cost and resources are also a big challenge. Not all schools and colleges can afford AI tools, as they require advanced computers, internet connection, and proper training for teachers. In many rural or underdeveloped areas, it is very difficult to bring such technologies into classrooms [16]. This creates a digital divide where some students enjoy the benefits of AI while others are left behind.
Another challenge is the fear of replacing teachers. Some people think that if AI becomes too powerful, it might reduce the need for human teachers. While AI can support teaching, it cannot replace the personal care, motivation, and emotional support that a human teacher provides. Still, this fear sometimes makes teachers hesitant to accept AI in their classrooms.

There are also privacy and security concerns. AI systems often collect large amounts of data about students, such as their learning progress, exam results, and even personal information. If this data is not protected properly, it could be misused [17]. Parents and teachers worry about how safe the data really is and whether it might fall into the wrong hands.
Lastly, there is the challenge of training and awareness. Many teachers and students are not fully familiar with how AI works. Without proper training, they may find it difficult to use AI tools effectively [18]. This can lead to frustration and misuse of technology.

Even though AI is helpful in education, there are many challenges that make it difficult to use everywhere. One major challenge is the black-box problem, where AI gives answers but does not explain how it reached them. This makes it hard for students and teachers to trust the results.

Another problem is bias in data. If the data used to train AI is unfair, the results will also be unfair. This means some students may get wrong feedback or fewer chances, which can create inequality in learning.
AI also brings the issue of high cost. Setting up AI tools needs good internet, strong computers, and regular updates. Many schools and colleges, especially in rural or poor areas, cannot afford these resources. This creates a gap between students who have access to AI and those who do not.

There is also the fear of job loss. Some people worry that if AI can teach or grade, teachers may not be needed in the future. Even though AI cannot replace the care and support of a real teacher, this fear still makes many teachers doubtful about using AI.

Privacy concerns are another challenge. AI collects a lot of data about students, like test results, study habits, and personal details. If this data is not handled carefully, it can be misused. Parents and teachers often worry about how safe this information really is.

Another issue is lack of awareness and training. Many teachers and students are not trained to use AI tools. Without proper knowledge, they may not use AI effectively or may even avoid it completely.
Apart from these, there is also the problem of over-dependence on technology. If students rely too much on AI for answers, they may stop thinking critically or solving problems on their own. This can affect creativity [19] and independent learning.

Although AI has many advantages for learning, the challenges show that it is not always easy to use in education. Problems like lack of transparency, bias in results, high costs, privacy issues, and the fear of

replacing teachers make many people unsure about trusting AI. In addition, not every school has the resources or training to use these tools effectively, which creates inequality among students.

These challenges do not mean AI should be avoided, but they remind us that AI must be used carefully and with proper planning. Teachers, students, and policymakers need to work together to make AI fair, safe, and useful. Only then can AI reach its true potential in supporting education without creating new problems.

4. U Role of Explainable AI (XAI) in Building Trust in Education

Explainable AI (XAI) plays a crucial role in making AI-powered learning systems transparent, fair, and trustworthy. Traditional AI often works like a "black box," leaving educators unsure why a model made certain recommendations or assessments. XAI solves this by clearly explaining decisions in simple terms, empowering teachers to validate AI-driven insights and adapt their teaching strategies [20].

For students, XAI provides clarity about personalized feedback and recommendations, turning AI into a trusted learning partner rather than an intimidating system. It also helps identify and correct biases in grading and resource allocation, ensuring fairer [21] and more inclusive education.
Beyond classrooms, XAI supports administrators and parents by offering interpretable analytics for curriculum design and student progress tracking. It also aligns with growing regulations that demand ethical, transparent AI use in sensitive fields like education.

In essence, XAI doesn't just explain algorithms—it builds trust [22]. By making AI's decision-making understandable, it creates a stronger partnership between humans and machines, making education more transparent, equitable, and student-centered.

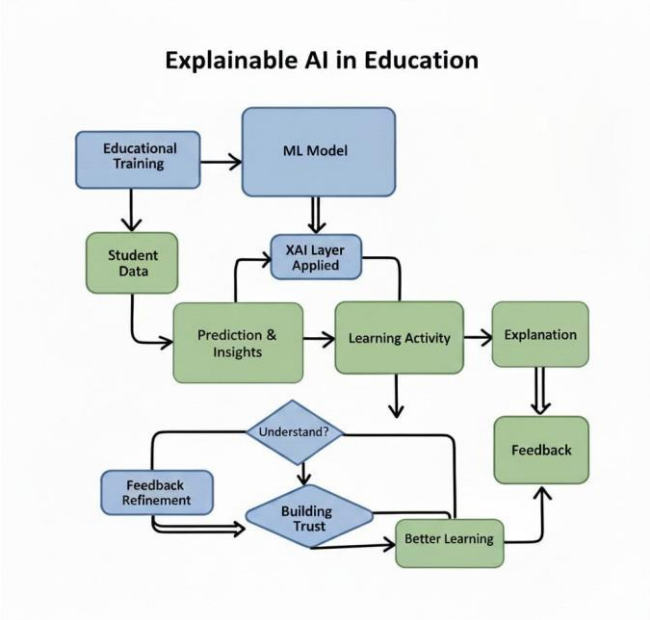

Fig. 4.1: This flowchart illustrates how Explainable AI builds student trust and improves learning by making AI-driven educational activities transparent and understandable.

5. Comparative Analysis of Traditional AI vs. Explainable AI in Education

In education, Artificial Intelligence tools have made learning more personalized and efficient. However, most of these systems are built on black-box models like deep neural networks, which are difficult to interpret [23]. This creates a gap in trust, as educators and students often cannot fully understand the reasoning behind predictions or recommendations.

Explainable AI (XAI) bridges this gap by offering transparency and interpretability. It allows stakeholders to see why a student was flagged for extra support, how grading algorithms work, and which data points influence recommendations [24]. XAI doesn't only make AI systems clearer but also empowers teachers to take informed decisions, ensures fairness by detecting biases, and improves adoption rates since users are more likely to trust transparent systems [25].

This comparison highlights that while Traditional AI focuses on automation and efficiency, XAI emphasizes collaboration between humans and AI, making education more ethical, inclusive, and accountable [26].

Feature/Aspect	Traditional AI in Education	Explainable AI (XAI) in Education
Model Transparency	Opaque models, difficult to interpret	Clear, interpretable explanations of decisions
Decision Justification	No clear reasoning provided	Provides reasoning for every prediction/recommendation
Human Trust	Low trust, "black-box" perception	High trust due to explain ability and transparency
Teacher Role	Minimal involvement in AI decision-making	Teachers can verify, adjust, and collaborate with AI recommendations
Bias Identification	Hard to detect algorithmic bias	Easier detection and correction of biases
Student Empowerment	Students feel judged by AI	Students understand and learn from AI feedback
Regulatory Compliance	Difficult to meet ethics regulations	Supports compliance with education AI policies and standards
Examples	Plagiarism detection, basic grading	Transparent learning analytics, adaptive learning platforms

Fig. 5.1: Comparative Analysis of Traditional AI vs. Explainable AI in Education

This figure highlights key differences between Traditional AI and XAI in education. While Traditional AI operates as a "black box" with limited transparency, XAI provides clear reasoning, supports human-in-the-loop decisions, and improves trust, fairness, and accountability in learning environments.

6. Limitations and Challenges of XAI

Even though Explainable AI (XAI) helps in building trust, it also has some limitations. One problem is that sometimes the explanations given by XAI are too technical and difficult for students or teachers to understand. If the explanation is not clear, it cannot really help.

1. **Complexity of Explanations**

   XAI is designed to make AI decisions clear, but often the explanations are still too hard for normal users [27]. Students and teachers may not have the technical background to understand them. If the explanation is confusing, it does not serve its purpose.

2. **Trade-off Between Accuracy and Explainability**
   Some AI models are very accurate but act like a "black box." [28] The process of making AI explainable usually starts with simplification. This simplification sometimes reduces accuracy, so there is always a trade-off between performance and clarity.
3. **High Cost and Resources**
   Creating XAI systems requires strong computers, skilled developers, and extra money. Many schools and colleges, especially in developing areas, cannot afford these systems. This limits the reach of XAI in education.
4. **Risk of Oversimplification**
   To make explanations easy, XAI sometimes gives very short or basic reasons. These reasons may not show the full truth behind the decision. As a result, students and teachers may develop a false sense of trust.
5. **Slower Performance**
   XAI takes extra time to generate explanations for its decisions. In fast situations like real-time learning or exams, this delay can be a problem. Students usually want quick results, and waiting may reduce interest.
6. **Lack of Standard Methods**
   There is no single method that all XAI systems follow. Different tools provide explanations in different ways. This lack of standards can confuse teachers and students when switching between systems.
7. **Limited Understanding for Non-Experts**
   Even if XAI provides clear explanations, non-experts may still find them hard to apply in real situations. Teachers without technical training may not know how to use the given information effectively in classrooms.
8. **Data Privacy Concerns**
   To give explanations, XAI often requires detailed access to student data. This raises privacy issues [29], as sensitive information about learners could be exposed or misused if not handled properly.
9. **Maintenance and Updates**
   XAI systems need constant updates and maintenance to stay effective. Schools and colleges may not have the technical staff to manage these updates, which makes long-term use difficult.
10. **Risk of Misinterpretation**
    Even when XAI explains its decisions, students or teachers might misinterpret the explanation. A wrong understanding can lead to poor trust or incorrect use of AI in learning.

Explainable AI (XAI) is an important step toward building trust in AI-powered learning, but it is not free from problems. The challenges show that even though XAI helps make AI decisions clearer, it still faces issues like complexity, high cost, slower performance, and privacy concerns. Many of these problems come from the fact that there is no single standard method and not everyone has the technical knowledge to fully understand the explanations.

These limitations remind us that XAI must continue to improve before it can be used widely in schools and colleges. Instead of seeing these challenges as barriers, they should be treated as opportunities for

research and better design. If future systems can reduce these problems, XAI can truly become a powerful and fair tool for education.

7. Problem-Solving Approaches in XAI

- One way to solve the problems of XAI is by making the explanations very simple and clear. Many people in schools or colleges may not have a technical background, so long and complex terms confuse them. If XAI tools use easy language, short sentences, and real-life examples, everyone can follow them. This also builds confidence and reduces fear of using new technology.
- Training programs can help teachers and students understand XAI better. When people are given small workshops [30] or tutorials, they learn how to read the explanations properly. Even basic training makes a big difference in how they trust and use AI. Without training, many users ignore the AI's feedback because they cannot understand it.
- Another solution is to balance between accuracy and simplicity. AI should give correct and useful results, but the explanation should not be too complicated. If the system is too technical, students may lose interest, even if the AI is right. On the other hand, too much simplification may reduce trust, so a middle path works best.
- Cost is a big challenge, but this can be solved by using open-source tools. Open-source AI systems are free or low-cost and allow schools with fewer resources to benefit from technology. They also give flexibility to customize the system. This ensures that not only rich institutions but also small schools can access XAI.
- Data privacy can be managed by designing systems that explain decisions without sharing personal information. In education, student records are sensitive, so privacy is very important. If XAI respects privacy while giving clear reasons for decisions, parents and teachers will trust it more. Safe systems will be accepted more quickly in schools.
- Another way to solve challenges is to create standard rules for explanations. If all XAI tools explain in the same style, it will be easier for students and teachers to move from one tool to another. This reduces confusion and makes the learning process smoother. Standards also save time for educators who use multiple systems.
- Visual aids can also be used to improve explanations. Many students find it easier to understand charts, diagrams, or small graphics than reading long text. Simple visuals can quickly show why AI gave a result. This method can especially help younger learners who are not comfortable with too many words.
- Improving the speed of XAI responses is another solution. If the system takes too long to explain, students may get bored or lose focus. Faster feedback keeps students engaged and helps them learn immediately. Developers can work on making the process smooth and quick without reducing the quality of explanations.
- Taking feedback from teachers and students is a good way to solve problems in XAI. Users can share what is easy to understand and what is confusing. This feedback helps developers make better updates. In this way, the system grows in a direction that matches real classroom needs.
- Collaboration between educators and AI experts can bring better solutions. Teachers know what students need, and AI experts know how to design the system. By working together, they can build tools that are both technically strong and user-friendly. This teamwork ensures the system is practical and useful in real classrooms.
- Another solution is to build trust slowly. Instead of using XAI in all areas at once, schools can start with small applications. As students and teachers get comfortable, they can expand its use. This step-by-step approach reduces fear and allows people to adapt better.

- Giving examples along with explanations can also make XAI clearer. For instance, if the AI says a student is weak in math, it should also show which question type caused the issue. Such examples make the explanation more relatable and easy to act on. Without examples, explanations may feel too general.
- Making XAI systems interactive can also solve challenges. If students or teachers can ask "why" and get a clear answer, it feels more natural. Interactive systems create a two-way communication, unlike static results. This makes users more satisfied and engaged.
- [31] Personalization is another solution to XAI challenges. Not all users need the same level of detail in explanations. Teachers may want deep reports, while students prefer simple notes. If XAI adjusts its explanations based on the user, it becomes more effective and useful for everyone.
- Explainable AI tools such as SHAP and LIME help make AI decisions more understandable. SHAP uses Shapley values from game theory to show how each feature contributes to a prediction, while LIME creates simple, interpretable models to explain individual decisions. These tools can help teachers and administrators make sense of AI recommendations and take informed actions in educational settings
- Finally, continuous improvement is key. Technology changes fast, so XAI systems must keep updating to solve new challenges. Developers should regularly upgrade explanations, visuals, and privacy features. With regular improvement, XAI can stay relevant, useful, and trusted in education for a long time.

8. FUTURE WORK

Future work on Explainable AI (XAI) in education should focus on improving scalability, personalization, and usability of AI-powered learning systems by designing hybrid models that balance high accuracy with clear, user-friendly explanations, exploring adaptive explanation systems tailored to users' roles, establishing global standards for consistency and interoperability, addressing data privacy concerns through secure and ethical frameworks, conducting real-world classroom trials to measure impact on engagement, teacher effectiveness, and equitable access, and developing a prototype model integrating SHAP and LIME to provide detailed explanations of AI predictions that enhance transparency, fairness, and trust in AI-driven learning.

9. CONCLUSION

The integration of Artificial Intelligence (AI) in education has brought significant benefits, including personalized learning, automated assessments, and actionable insights for teachers. However, the black-box nature of many AI systems creates challenges of trust, fairness, and usability, making Explainable AI (XAI) a critical solution for ethical and responsible adoption. XAI not only improves transparency and accountability but also empowers students and teachers by providing clear reasoning behind AI-driven decisions. Although challenges such as complexity, high costs, and data privacy concerns remain, targeted strategies like training programs, standardization, and user-centered design can address these limitations. Overall, XAI is not merely a technical improvement but a pedagogical necessity, enabling equitable, trustworthy, and effective AI-powered learning environments. Future research should focus on building hybrid approaches that combine interpretability with scalability, ensuring AI becomes a supportive partner in education rather than a source of doubt.

10. ACKNOWLEDGMENT

We want to give a huge thank you to Ms. K. Swathi, our Associate Professor in Artificial Intelligence. Her constant guidance, encouragement, and valuable insights were absolutely essential to this research. And to our amazing team, thank you for all your dedication, collaboration, and hard work. We couldn't have done this without you.

References

1. H. Khosravi, et al. (2022). Explainable Artificial Intelligence in education. (Cited by 744) — discusses the role and need for XAI to build trust and transparency in educational AI systems.
2. G. Türkmen (2025). The Review of Studies on Explainable Artificial Intelligence in ... — a systematic analysis (35 studies) of educational XAI from 2019–2024 via the PICOS framework.
3. M. Saqr et al. (2024). Why explainable AI may not be enough: predictions and ... — emphasizes the value and limitations of instance-level explainability in educational AI.
4. Y. Feldman-Maggor, M. Cukurova, C. Kent & G. Alexandron (2025). The Impact of Explainable AI on Teachers' Trust and ... — demonstrates how XAI enhances understandability, trust, and acceptance among teachers, with domain-driven explanations more effective.
5. J. Prentzas (2025). Explainable Artificial Intelligence Approaches in Primary Education — a review of XAI in primary education contexts using Google Scholar & Scopus; highlights explainability to increase trust and effectiveness.
6. B. Leichtmann (2023). Effects of Explainable Artificial Intelligence on trust and … (Sciencedirect) — empirically investigates how XAI methods affect user understanding and trust.
7. R. Farrow (2023). The possibilities and limits of explicable AI in education (Taylor & Francis) — explores XAIED as a means to improve trust and ethical usage in algorithmic education.
8. K Sai Geethanjali & Nidhi Umashankar (2025). Enhancing Educational Outcomes with Explainable AI: Bridging Transparency and Trust in Learning Systems. (Discusses how XAI promotes transparency, trust, and addresses the opaque nature of AI in education.)
9. Zaid M. Altukhi & Sojen Pradhan (2025). Systematic Literature Review: Explainable AI Definitions and Challenges in Education. (Provides a detailed overview of how XAI contributes to trust and highlights challenges in educational contexts.)
10. Tu, Y., Binhammad, M. H. Y., Othman, A., Abuljadayel, L., Al Mheiri, H., & Almarri, M. (2024). Investigating how generative AI can create personalized learning materials tailored to individual student needs. Creative Education, 15(7), 1499–1523.
11. Gobrecht, A., Tuma, F., Möller, M., Zöller, T., Zakhvatkin, M., Wuttig, A., Sommerfeldt, H., & Schütt, S. (2024). Beyond human subjectivity and error: a novel AI grading system. arXiv.
12. Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2023). Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education. arXiv.
13. Luo, J. (2025). Design and assessment of AI-based learning tools in higher education. Educational Technology Research and Development.
14. Lipton, Z. C. (2016). The Mythos of Model Interpretability. Proceedings of the 2016 ICML Workshop on Human Interpretability in Machine Learning, 96–100.
15. Holstein, K., Wortman Vaughan, J., Wallach, H., Daumé III, H., & Wallach, D. (2019). Improving Fairness in Machine Learning Systems: What Do Industry Practitioners Need to Know? Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, 1–16.

16. Williamson, B., & Piattoeva, N. (2019). Education Governance and Datafication: Critical Perspectives and Emerging Issues. Learning, Media and Technology, 44(1), 1–14.
17. Pardo, A., & Siemens, G. (2014). Ethical and Privacy Principles for Learning Analytics. British Journal of Educational Technology, 45(3), 438–450.
18. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. Center for Curriculum Redesign.
19. Selwyn, N. (2019). Should Robots Replace Teachers? AI and the Future of Education. PoliPointPress.
20. Swamy, V., Du, S., Marras, M., & Käser, T. (2023). Trusting the Explainers: Teacher Validation of Explainable Artificial Intelligence for Course Design. Proceedings of the 13th International Conference on Educational Data Mining.
21. Deck, L., Schoeffer, J., De-Arteaga, M., & Kühl, N. (2024). A critical survey on fairness benefits of explainable AI. Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency.
22. Weller, A. (2019). Transparency: motivations and challenges. In Explainable AI: Interpreting, Explaining and Visualizing Deep Learning (pp. 23-40). Springer.
23. Longo, L., Goebel, R., Lecue, F., Kieseberg, P., & Holzinger, A. (2020). Explainable artificial intelligence: Concepts, applications, research challenges, and visions. Proceedings of the 2020 IEEE European Symposium on Security and Privacy (EuroS&P).
24. Fiok, K., Farahani, F. V., Karwowski, W., & Ahram, T. (2021). Explainable artificial intelligence for education and training. Proceedings of the 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC).
25. Swamy, V., Du, S., Marras, M., & Käser, T. (2023). Trusting the explainers: Teacher validation of explainable artificial intelligence for course design. Proceedings of the 13th International Conference on Learning Analytics & Knowledge (LAK)
26. Hall, P., & Gill, N. (2019). An introduction to machine learning interpretability: An applied perspective on fairness, accountability, transparency, and explainable AI. Proceedings of the 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA).
27. Wenli Yang et al., “Survey on Explainable AI: From Approaches, Limitations and Applications Aspects,” Human-Centric Intelligent Systems, 2023.
28. A. Bell, “It’s Just Not That Simple: An Empirical Study of the Accuracy-Explainability Trade-off,” FAccT ’22, 2022.
29. Zana Buçinca et al., “Proxy Tasks and Subjective Measures Can Be Misleading in Evaluating Explainable AI Systems,” arXiv, 2020.
30. Bueff, Andreas; Papantonis, Ioannis; Simkute, Auste; Belle, Vaishak. (2025). Explainability in machine learning: a pedagogical perspective. Frontiers in Education. Demonstrates effectiveness of interactive tutorials and structured workbooks for teaching XAI
31. Conati, C.; Barral, O.; Putnam, V.; Rieger, L. (2019). Toward Personalized XAI: A Case Study in Intelligent Tutoring Systems. Shows how personalized explanations improve trust, perceived usefulness, and learning gains.