The Role of Explainability in University Students' Willingness to Accept Learning Recommendations from Digital Human Tutors: A Literature Review

Xiao Lan
College of Art and Science, New York University, New York, NY 10003, USA
xl4842@nyu.edu

ABSTRACT

With the development of artificial intelligence, an increasing number of educational platforms have started using digital human tutors to provide students with personalized learning suggestions. Secondly, listed companies should first consider the benefits of digital transformation to governance. This is to strengthen the introduction of digital technologies in the fields of internal oversight, administration and information disclosure. By using digital tools to improve governance mechanism and reduce agency expenses, enterprises can give full play to the full power of digital value creation. Third, regulatory bodies can use digital technologies to better supervise listed companies. In this way, information disclosure will become clearer faster, and the ability to monitor and restrict the behavior of operators will become stronger. These efforts will help to create a good institutional environment, and digital conversion can play a good role in governance.

Keywords: Explainable AI; Digital Human Tutor; Student Acceptance; Technology Acceptance Model; Literature Review

1. INTRODUCTION

Artificial intelligence is now more and more common in online learning platforms and educational software. The number of these platforms is increasing, and AI's digital human tutor is used to help students learn. These systems can look at students' past achievements and give suggestion on chapter, exercise and study plan legislation that need to be reviewed. Some platform use virtual characters similar to people as interfaces, aiming to gain the experience of talking with real tutor [1][2]. In this paper, digital human tutor is a virtual agent run by AI that looks like a human being, and it is a kind of interaction with students, thus producing learning guidance suitable for everyone. Although this concept has similar terms, there are also differences. Pedagogical agents is a broader category. In order to help learning, you can put all the characters that appear on the screen, but not all of them move on AI, and you can also give a recommendation suitable for everyone [1]. Intelligent Tutoring Systems (its) pays attention to personalized guidance and feedback, but it doesn't necessarily use a human-like interface. General AI recommendation systems can give suggestion suitable for everyone, but there is no dialogue and communication of the body as a characteristic of digital human tutor.

This topic also comes from the great challenge of educational AI research. When students feel that the system does not know their current learning goals, they may not be committed to the recommendations given by AI. Pedagogical agents' research shows that the influence of students' motivation is obvious [1]. Many people also know that it is necessary to use the learning tool transparency of AI [3]. However, why do students follow or ignore these systematic learning suggestions? This problem has not been completely solved. From a broader perspective, students don't always listen to suggestions given by AI system. One possible reason is that many AI systems work like "black box". The system will say "do this", but it won't explain to me why. For example, if an app says "you should review Chapter 3" but doesn't tell us the reason, students may ignore it without knowing the meaning. This lack of transparency makes it difficult for students to trust the system and seriously accept suggestions [3].

Explainable AI (XAI) is considered as a way to solve this problem. XAI is committed to making the AI system easier to understand. In this way, the user knows the reasons for the system judgment [4][3]. In education, this is not only "reviewing Chapter 3", but also explaining that "the correct rate of your Chapter 3 question is only 40%, which is much lower than the average level, so you'd better re-examine this part first". With this explanation, students may be more likely to accept the recommendation.

However, the research so far has talked about the pedagogical agents in the field of education, the trust in automation and XAI respectively, but these fields are always investigated separately. In particular, how the ability to explain affects the mood of college students who are recommended to learn from digital human tutors is rarely concerned [5]. XAI research in the field of education mostly pays attention to the teacher's point of view and the structure of technology, but does not understand how the last user-students will react to the explanation of AI.

This paper tries to check this missing part in literature review. The purpose is to answer three questions. First, what are the main factors that affect whether university students are willing to accept learning recommendations from digital human tutors? Second, what role does XAI play in this process? Thirdly, based on what the current research tells us, what should the future research and product design focus on? The rest of the paper gives a review of literature in Section 2, an analysis framework from three new directions in Section 3, a detailed description of the role of XAI in Section 4, and a conclusion in Section 5.

2. LITERATURE REVIEW

This section will look at the existing research in the three areas involved in this study. The first part reviews the research on pedagogical agents in educational field. The second part discusses why people accept or reject the theory of new technology, and focuses on the Technology Acceptance Model and the trust in automation. The third part will discuss the research on XAI in the field of education so far.

2.1 Pedagogical Agents in Education

A digital assistant is a phantom agent. Pedagogical agent is a virtual character on the screen, which is made to help students learn in the digital environment. The research on this topic has been carried out in these 20 years, and there are many literatures to investigate whether these virtual characters can help students learn more effectively. Heidig and Clarebout [1] investigated many experimental studies on agents and found that agents can help students acquire more knowledge than learning without agents. However, with regard to motivation, the result is clear-there is no conclusive evidence that agents can further improve students' learning motivation. Schroeder, Adesope and Gilbert [2] used different methods to conduct meta-analysis, and summarized 43 studies with more than 3,000 participants. They found that agent has little statistical significance to learning, but it will have a good impact. An interesting result is that agents who use screen text work better than those who use voice explanation, and young students seem to benefit more than college students.

Schroeder and Adesope [6] later published another summary. This summary describes in more detail the students' feelings when learning with pedagogical agents. About motivation, hobbies, cognitive load and other elements. According to their results, students generally prefer to be in the school than not having an agent. On certain occasions, the agent also has a little clue to help drive. However, the results of cognitive load are mixed, and I don't know if the agent will add extra head burden to the learners.

It is worth noting that most of the previous studies are aimed at relatively simple virtual characters. In other words, it is an agent who displays information and responds according to the script. In the recent research, we changed according to each student's situation and began to investigate the agents who provided personal advice to AI [7]. But even after this change, most of the existing research still focuses on the technical part of making this system. From the students' point of view, few studies specifically ask what makes them want to accept or not accept the proposal of digital human tutor.

2.2 User Acceptance and Trust Theories

In order to understand why students accept or ignore the learning suggestions put forward by AI, it is good to look at the theories so far about how people react to new technologies.

Before reading these theories, it is good to find out the related but different things that appear in the research. Perceived usefulness means that users believe that using this system will improve their grades [8]. The so-called Trust is to make users believe that this system is reliable and can work for themselves [9]. Acceptance or usage intention refers to the user's clear intention to use or continue to use the system. Actual compliance refers to whether users actually follow the suggestions of the system in their lives. These things are related-trust affects perceived usefulness, and it affects acceptance, but it is not the same thing. On this paper, the result we are most concerned about is that students accept and listen to learning suggestions, which is not only the concept of trust, but also closer to the concept of acceptance or compliance.

The Technology Acceptance Model, that is, TAM made by Davis [8], is one of the most commonly used concepts in this field. The basic idea is simple: whether a person uses the new technology depends mainly on two things. The first one is perceived usefulness, that is, the degree to which that person believes that "this technology can help him do better". The second is perceived ease of use, which means the degree of "how hard it takes to use this". If students believe that AI tutor can really provide useful advice and the system is easy to communicate, they will be more likely to use it. TAM was originally made for workplace information systems, and later it began to be used in many other places, and educational technology is one of them.

Another related framework comes from Lee and See [9]. They wrote about their trust in automation. Their research focuses on how people decide whether to rely on automated systems. They explained the three foundations of trust: the first purpose is whether the user believes that the system has a good purpose. The second performance is whether the system can successfully complete its own work. The third process is how clear the system behavior is. The important point they say is that trust needs to be "adjusted" correctly. If users trust the system too much, they may listen to bad advice without asking questions. If you are too incredulous, useful suggestions will be ignored. The goal is not to maximize trust, but to help users cultivate a degree of trust that is consistent with what the system can actually do.

Add up these two theories and you will know that they are related to this study. The idea that students want to follow AI's learning advice comes from both the practical consideration of whether the tool is useful and easy to use, and the factors related to trust, that is, whether they believe that the system can make conclusive suggestions and how they make decisions. The second part-the transparency of decisions-is exactly what XAI wants to do well, which will lead to the next part.

2.3 XAI in Educational Contexts

Explainable AI has received growing attention in recent years, and its application in education is starting to be explored, though the research is still at a relatively early stage.

Khosravi et al. [3] made an important contribution by proposing what they called the XAI-ED framework. This framework identifies six key dimensions that should be considered when designing and studying explainable AI tools for education. One of the main arguments in their paper is that education is not the same as other domains where XAI has been applied, like healthcare or finance. In education, there are multiple stakeholders — students, teachers, and administrators — and the purpose of AI is not just to make accurate predictions, but to support the learning process. This means that what counts as a good explanation might be different in an educational context compared to, say, a medical diagnosis system.

Farrow [5] looked at XAI in education from a broader, socio-technical angle. His argument is that explainability is not a fixed concept — what counts as a satisfactory explanation depends on who is asking and what their background is. A technical explanation that makes sense to a data scientist might be completely unhelpful to a student or a teacher. He also raised an important caution: XAI should not be treated as a complete solution to the problems of AI in education. It is a useful starting point for making AI more transparent, but there are deeper questions about fairness, accountability, and power that explainability alone cannot answer.

On the empirical side, Conati et al. [10] conducted a study that is particularly relevant to this review. They worked with an Intelligent Tutoring System and added an explanation feature that told students why the system was giving them certain hints. Their survey results show that the students who read the instructions think the system is more credible, think the tips are more useful and want to use the system. But this study also found that not all students have the same reaction. Students with low Need for Cognition, which is related to how much people enjoy thinking, pay little attention to explanation. On the other hand, students with low Conscientiousness seem to get more good results through explanation. This shows that a design with the same description may be equally good for everyone. Away from students, Feldman-Maggor et al. [11] studied the influence of XAI on teachers' trust in using educational technology proposed by AI. They worked with 41 chemistry teachers and tried all kinds of explanations. Their important discovery is that combining the explanations of subject areas (for example, linking the proposal with the chemical ideas that teachers are familiar with) is more effective in cultivating trust than the explanations made of data that only show numbers and statistics. Although this study focuses on teachers rather than students, it gives us some useful ideas: it is very important to make instructions, and if it is related to the knowledge that users have already mastered, the instructions will become more credible.

It should be remembered that Conati et al. [10] and Feldman-Maggor et al. [11] studied different people, namely students and teachers. Both groups found that the trust would be higher after the explanation, but the structure might be different. Teachers will evaluate the recommendation of AI according to their professional knowledge, but students may not have enough subject knowledge to investigate whether the AI idea is correct. Therefore, in the absence of further investigation, we can't directly apply the research results targeted at teachers to students.

Overall, these studies show that XAI may help trust and acceptance in the educational AI system. However, it is not just a supplementary explanation that can automatically increase trust. It seems that the design of the explanation, the characteristics of the person who accepts the explanation and the place where it is used are all related. The next section proposes a framework for further organizing and arranging these different elements.

3. THREE-DIMENSIONAL FRAMEWORK OF INFLUENCING FACTORS

Based on the reviewed literature, this study proposes a way to organize the different factors that may influence whether students are willing to accept learning recommendations from digital human tutors. Drawing on findings from pedagogical agent research, technology acceptance theory, and XAI studies discussed in Section 2, the framework groups these factors into three categories: system-side factors, user-side factors, and interaction-side factors. This is not meant to be a definitive or exhaustive classification, but rather a way to make sense of the various findings scattered across different studies.

3.1 System-Side Factors

System-side factors are about the AI recommendation system itself — how it performs and what features it has.

One important factor is recommendation accuracy. If the system consistently suggests content that matches what the student actually needs, the student is more likely to trust it over time [9]. But accuracy alone does not seem to be enough. A recommendation could be correct based on the student's data, but if the student feels the suggested content is too basic, or not related to what they are preparing for, they may still choose to ignore it. Therefore, student acceptance may depend not only on objective recommendation accuracy, but also on perceived relevance.

Whether the system explains its recommendations is another system-side factor, and one that is central to this paper. As discussed in Section 2.3, studies like Conati et al. [10] and Shin [4] have found that providing reasons for AI decisions can make users more willing to trust and follow them. If there is no explanation, the system will remain black box, and students have little basis to judge whether they should follow the advice. Other system factors appearing in the literature include response speed and overall reliability. Even if the actual recommended content is good at runtime, a system with slow response or frequent errors may lose the trust of students. Although these factors are rarely studied in the specific case of AI education, they are often discussed in the broader technology acceptance literature [8].

3.2 User-Side Factors

User-side factors are related to what student have in interaction, that is, their attitude, personality and experience. One factor that has emerged in several studies is the attitude of student towards AI a long time ago. Although some student have a positive attitude towards using AI tools and think it is very helpful, others may feel uneasy and suspicious about letting algorithm guide their own learning. According to TAM [8], attitude plays an intermediary role between people's belief in technology and whether they actually use it. Therefore, student who already have a bad opinion of AI, even if the system runs well, are sometimes difficult to accept. Technology anxiety is a similar but slightly different factor. Some student are not opposed to AI in principle, but they will feel uneasy when dealing with new and unfamiliar technologies. This unpleasantness has reduced our desire to devote ourselves to AI tutoring systems, which is also consistent with the contents of the broader technology acceptance research [8]. On the other hand, students with more experience in using digital tools and/or students with higher level of digital literacy may find it easier to evaluate and follow AI's recommendation.

The motivation to study also seems to be very important. Students who work hard to improve their grades, or students who want to learn a course well, will mostly accept useful advice frankly, while students who are not motivated, no matter how good the advice, often don't listen. Finally, it also involves everyone's cognitive characteristics. The research of Conati et al. [10] is of great reference value in this respect. They found that students with low Need for Cognition tend to be less focused on the instructions given by the system. At the same time, the researchers also found that students with low Conscientiousness can learn more when they are seen and explained than serious students. Students are different-even if they use the same system, everyone's reaction may be completely different, and what is suitable for one type of student may not be suitable for other students.

3.3 Interaction-Side Factors

Interaction-side factors are about how the student and the digital human tutor interact with each other — the design of the interface, the agent's appearance, and the way information is communicated.

The appearance of the digital human is one such factor. Research on pedagogical agents has shown that visual design can influence how students feel about the learning experience [6]. Some studies suggest that more human-like agents can increase engagement, but this relationship is not always straightforward. If the virtual character looks very realistic but behaves in an unnatural way, it may make students uncomfortable rather than more trusting. Although this issue is not central in the reviewed studies, it shows why the design of the agent's appearance should be considered carefully.

How the digital human communicates also matters. Schroeder et al. [2] found that the way agents deliver information — whether through text or narration — can affect learning outcomes. In the context of learning recommendations, this raises the question of whether a conversational style (where the digital human talks to you like a tutor would) might be more effective than a simple list format (where suggestions are just displayed on screen). There is not much research on this specific comparison in educational AI, but the broader pedagogical agent literature suggests that communication style is worth paying attention to.

Perhaps most relevant to this paper is how explanations are presented. It is not just about whether an explanation exists, but about how it is designed and delivered. Feldman-Maggor et al. [11] showed that domain-specific explanations, framed using language and concepts familiar to the user, were significantly more effective than abstract, data-driven explanations. Leichtmann et al. [12] found that visual explanations can help users have just the right trust compared with the situation without visual explanations. For digital human tutors, this means that explanation design should be treated as part of interaction design, not as another technical function.

4. THE ROLE OF XAI IN STUDENT ACCEPTANCE

In the previous chapter, the influencing factors are classified into three ways. Among the elements mentioned, explainability is special. If you want to say why, it is because there is only one classification method that cannot be completely entered. Whether the system explains or not is a characteristic of the system. However, how to convey that description belongs to the distinction method of interaction. Moreover, how students react to the instructions is partly related to the user-side characteristics. Therefore, XAI involves the whole framework. In this chapter, we will see in detail what the research so far has said about the role of XAI.

4.1 Positive Effects of XAI Explanations

The most clear research so far is that it is easy to produce some good results if the reasons for the AI proposal are explained. For example, trust, feeling that you can use it, feeling that you want to use that system, etc., at least in some cases.

Conati [10] explained to the students who used the intelligent tutoring system why the system put forward the prompt, and the students who got the explanation had a high degree of trust and felt that they could use it. Using the system again was probably the most direct result related to this paper. Because this study investigates students' acceptance of AI learning suggestions, especially with and without explanations.

Shin [4] found something similar in another case. When he studied AI-driven recommendation services, he found that explainability had a good relationship with user trust. He also put forward the idea of "causability". This is the degree to which the user can fully understand the instructions given. From his results, it is not enough to give a description. If the user can't understand the description, it will not have an impact on trust.

In a wider context, XAI-ED framework of Khosravi et al. [3] thinks that transparency and trust are important reasons for putting XAI into educational AI system. That's what they advocate. When students and teachers see why the AI system makes this judgment, it becomes very simple to evaluate this judgment and decide whether to follow it.

4.2 Boundary Conditions and Limitations

However, it can be clearly seen from the contents of the research book that XAI is definitely not a good solution. Some studies have pointed out the important conditions that affect whether the explanation is actually useful.

The first condition is the kind of description and the method of making it. Feldman-Maggor et al. [11] found that domain-driven description using domain-specific languages and concepts is much more effective than data-driven description relying on numbers and algorithm details. In the scene of learning recommendation, the value of XAI may not only be to improve trust, but also to help students judge whether recommendation is suitable for their real learning needs.

The second condition is that each user is different. As mentioned earlier, Conati et al. [10] found that students with low cognitive desire pay less attention to explanation, and students with low integrity benefit more from explanation. This means that it may not be the best strategy to use the same interpretation method for everyone. Some students need a simpler explanation in order to attract attention, while others prefer a more detailed thinking process.

The third condition is related to what we mean when we say “improving trust”. Leichtmann et al. [12] gave us a detailed explanation on this point. In their study, XAI doesn’t simply make people trust system more. On the contrary, XAI helps users cultivate what they call “appropriate” trust. This means that the user’s trust level is now consistent with the actual trust level of the system. Especially for university students, this difference is very important. Because students are not only passive recipients of information, they often have their own study plans, preferences and their own evaluations. If XAI doesn’t help students evaluate the quality of specific recommendations, but simply makes them trust system more, it may lead to over-dependence, not better learning results. Finally, Farrow [5] gives us a broader view that we should remember. He argues that explainability is always relative. That varies according to who the person listens to the explanation, what that person already knows, and the context in which that person is located. Although it is very effective for university student studying mathematics, it is sometimes completely ineffective for students in completely different fields. Moreover, there is a bigger question than whether the explanation is effective: should AI judge a specific education? These are not questions that XAI can answer, but they are an important part of the whole.

5. CONCLUSION

This paper investigates the research so far and finds out what factor will affect the mood of university students who want to accept the recommendation from digital human tutors. Pay special attention to the role of XAI. Based on the investigation and research, a three-dimensional framework is made to divide these factor into three parts: system side, user side and interaction side. In review, if XAI’s explanation is well done and conforms to users, it will help to better adjust trust and make it more acceptable. However, the effect does not appear randomly, but depends on how the explanation is displayed, who the user is and the specific learning context. For example, domain-driven explanation is more effective than purely data-driven explanation [11]. In addition, students with different personality traits may have different reactions to descriptions [10]. To create effective explainability features for educational AI, we must carefully consider both the system and the users. You can’t just add the function of explanation, I hope it is suitable for everyone.

This study also has some limitations. Because we are conducting literature review, all the conclusions are based on what other researchers have found, and the number of studies that directly investigate the intersection of digital human tutors, XAI and student acceptance is still very small. The three-dimensional framework we launched has not been tested by the original data we collected. Therefore, this should not be a confirmed model, but should be used as a tool to organize information. It is also possible that in the process of searching for literature, some related studies have been missed. Up to now, most studies treat students as almost the same group, but student acceptance often changes according to the experience and learning goals before using their respective study habits and AI tools. Future research should pay more attention to this individual difference. In the actual education scene, it is of great value to conduct experimental research to investigate the influence of different kinds of XAI descriptions on student acceptance and confirm whether the framework launched this time can also be used in various student groups, disciplines and cultures. From the perspective of product design, developers who make educational AI should not only consider whether to add the description function, but also consider how to make the description in a way that is easy to understand and useful for students with various backgrounds and study habits.

REFERENCES

[1] Heidig, S., Clarebout, G. (2011) Do pedagogical agents make a difference to student motivation and learning? Educational Research Review, 6(1): 27–54.
[2] Schroeder, N. L., Adesope, O. O., Gilbert, R. B. (2013) How effective are pedagogical agents for learning? A meta-analytic review. Journal of Educational Computing Research, 49(1): 1–39.
[3] Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y.-S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., Gašević, D. (2022) Explainable artificial intelligence in education. Computers and Education: Artificial Intelligence, 3: 100074.
[4] Shin, D. (2021) The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies, 146: 102551.

[5] Farrow, R. (2023) The possibilities and limits of explicable artificial intelligence (XAI) in education: A socio-technical perspective. Learning, Media and Technology, 48(2): 266–279.
[6] Schroeder, N. L., Adesope, O. O. (2014) A systematic review of pedagogical agents' persona, motivation, and cognitive load implications for learners. Journal of Research on Technology in Education, 46(3): 229–251.
[7] Hostetter, J. W., Conati, C., Yang, X., Abdelshiheed, M., Barnes, T., Chi, M. (2023) XAI to increase the effectiveness of an intelligent pedagogical agent. In: Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents (IVA '23). ACM.
[8] Davis, F. D. (1989) Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3): 319–340.
[9] Lee, J. D., See, K. A. (2004) Trust in automation: Designing for appropriate reliance. Human Factors, 46(1): 50–80.
[10] Conati, C., Barral, O., Putnam, V., Rieger, L. (2021) Toward personalized XAI: A case study in intelligent tutoring systems. Artificial Intelligence, 298: 103503.
[11] Feldman-Maggor, Y., Cukurova, M., Kent, C., Alexandron, G. (2025) The impact of explainable AI on teachers' trust and acceptance of AI EdTech recommendations: The power of domain-specific explanations. International Journal of Artificial Intelligence in Education, 35: 2889–2922.
[12] Leichtmann, B., Humer, C., Hinterreiter, A., Streit, M., Mara, M. (2023) Effects of explainable artificial intelligence on trust and human behavior in a high-risk decision task. Computers in Human Behavior, 139: 107539.