Algorithm Appreciation in Education: Educators Prefer Complex over Simple Algorithms

Kimberly Williamson
Cornell University
Ithaca, NY, USA
khw44@cornell.edu

Rene Kizilcec
Cornell University
Ithaca, NY, USA
kizilcec@cornell.edu

Sean Fath
Cornell University
Ithaca, NY, USA
sf562@cornell.edu

Neil Heffernan
Worcester Polytechnic Institute
Worcester, MA, USA
nth@wpi.edu

Abstract

Algorithm aversion among educators can pose a challenge to the adoption of AI tools in education, especially when complex algorithms are involved. This study investigates how providing explanations for a complex algorithm in an intelligent tutoring system (ITS) affects educators' attitudes, trust, and willingness to adopt the tool. In two randomized experiments (N = 570), we compare educator preferences between a simple heuristic algorithm and a complex (Bayesian Knowledge Tracing) algorithm, focusing on how explanations for the complex algorithm can improve attitudes and adoption. Surprisingly, we found that educators generally preferred the complex over the simple algorithm, and explanations did not improve attitudes or adoption intentions, even when educators had to explain the complex algorithm's predictions. The complex algorithm scored lower on informational fairness than the simple one, considering it is less transparent, and the explanation was insufficient to overcome this. Overall, the findings suggest that widespread algorithm aversion may have evolved into algorithm appreciation, at least in the context of widely used technologies like ITS.

CCS Concepts

• **Applied computing** → *Computer-assisted instruction*; *Interactive learning environments.*

Keywords

Algorithm Aversion, Algorithm Appreciation, AI Literacy, XAI

ACM Reference Format:
Kimberly Williamson, Rene Kizilcec, Sean Fath, and Neil Heffernan. 2025. Algorithm Appreciation in Education: Educators Prefer Complex over Simple Algorithms. In *LAK25: The 15th International Learning Analytics and Knowledge Conference (LAK 2025), March 03–07, 2025, Dublin, Ireland*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3706468.3706535

LAK 2025, Dublin, Ireland

ACM ISBN 979-8-4007-0701-8/25/03
https://doi.org/10.1145/3706468.3706535

1 Introduction

Trust is an important factor influencing the pace of artificial intelligence (AI) adoption in education because it is essential for accepting and effectively integrating AI-based educational technologies [24, 36]. Educators are reluctant to accept AI recommendations that contradict their prior knowledge about students, and they expect AI to be infallible even in subjective situations [28]. Factors that influence trust generally include perceived benefits, ease of use, transparency, and anxiety [2, 8]. In higher education, Aladi [1] highlights the lack of transparency, reliability issues, and ethical concerns as key factors slowing down the integration of AI technologies. These factors need to be addressed to understand how best to handle concerns of adoption and trust and successfully implement AI-powered tools in education.

Algorithm aversion has been proposed as an explanatory framework to understand how trust, transparency, and confidence issues might undermine the adoption of AI tools. Algorithm aversion posits that educators would be reluctant to use tools with algorithms, even though these tools' capabilities could benefit their instructional practices [12, 21]. Algorithm aversion suggests that educators would prefer simple algorithms — or seek human advice — over seemingly complex algorithms. The framework also suggests remedies to reduce algorithm aversion, such as increased transparency [35], giving users control over the algorithm's outcomes [7], and providing feedback mechanisms [39], to raise the likelihood of technological adoption. However, how to provide effective explanations for AI in education, explanations that raise algorithmic transparency in such a way that promotes technology adoption, remains an open question.

In this paper, we provide a background on algorithm aversion in education and present two experimental studies investigating the effects of providing detailed explanations for a complex algorithm on lowering algorithm aversion and encouraging tool adoption. Our overarching research question is: How does the presence of a detailed explanation for a complex algorithm affect educators' attitudes and intent to use an AI-powered tool? Study 1 employed a two (explanation) by two (visualization) between-subjects design with repeated measures to compare educators' attitudes and preferences for a complex algorithm versus a simple one when exposed to both algorithms. Study 2 compares three conditions (simple algorithm vs. complex algorithm with explanation vs. without explanation) in a between-subjects design, measuring educators' attitudes and preferences after explaining the algorithm to another person—i.e., simulating educators' need to explain and justify decisions informed by algorithms they use. Lastly, we discuss the implications of this research for how algorithm aversion among educators may have evolved, at least for AI-based learning analytic technology.

2 Background

2.1 Algorithm Aversion

Algorithm aversion refers to people's reluctance to use algorithmic decision aids despite their superior performance [12]. This phenomenon has been observed in various domains, including autonomous vehicles [20], HR decision-making [29], and forecasting [12]. In education, it has been studied with both in-service and pre-service teachers, who often prefer human advice over expert models [21].

Several factors influence algorithm aversion, including perceived task subjectivity [4], familiarity with algorithms [26], and trust in the system [19]. However, Kaufmann [21] found that commonly studied personality traits, such as openness and neuroticism, are unrelated to teachers' perceptions or behaviors regarding expert models using algorithms. Introducing AI tools for instructors requires considering the effects of algorithm aversion to increase adoption and use and identifying instructor-related factors that might influence initial levels of algorithm aversion. Researchers have proposed using various levels of algorithm transparency to help reduce algorithm aversion [7, 35]. Xu et al. [39] tested algorithm aversion reduction strategies with higher education administrators using an AI tool to assist with making credit articulation decisions. They found mixed results for reducing algorithm aversion, as allowing users to provide feedback about the algorithm increased aversion instead of reducing it. This study, like other educational research before it, has emphasized the importance of human-centered design for crafting effective technology strategies in education [3, 16]. More research is needed to understand what approaches can influence algorithm aversion, such as providing higher levels of algorithm transparency.

2.2 Measuring Algorithm Aversion

Various study designs have been used to assess algorithm aversion. One common approach is to compare human and algorithmic decisions [4, 12, 21]. Another approach is to assess participants' perceived competence of the algorithms based on the recommendations made [4]. Finally, researchers have examined participants' intention to use tools with algorithms [7, 12, 20]. This current study uses a blend of these approaches to study the differences in educators' perceptions between a simple heuristic and a more complex algorithm.

2.2.1 Attitudes. One way to measure perceptions of algorithms is to measure educators' attitudes about different algorithms. Prior research indicates that attitudinal constructs such as perceived confidence, accuracy, and trust are strongly associated with lower algorithm aversion and increased adoption of new technologies [4, 7, 12, 19, 28]. Likewise, educators' trust and confidence in education technology tools significantly influence their adoption and usage. Key factors affecting trust include perceived benefits, transparency, self-efficacy, and understanding of AI [28, 36]. Trust can be enhanced through professional development programs that explain AI decision-making processes [27] or other essential factors, including minimizing additional workload, increasing teacher ownership, and addressing ethical concerns [9]. Further research on educators' perspectives and needs can guide approaches to optimize AI use in education [24, 32]. This area of focus will require measuring educator attitudes related to trust and confidence, given their association with both algorithm aversion and technology adoption.

2.2.2 Algorithm Comparisons. In addition to understanding educators' perceptions of an algorithm, algorithm aversion also requires us to understand the differences in perceptions between decision-making agents. Commonly, comparisons are used to study the effects of algorithm aversion are investigated by providing two versions of decision-making agents to be directly compared, typically a human or an algorithm [4, 12, 21]. Participants are exposed to decisions made by both agents and asked to answer which decision-making agent they would prefer to use. While effective, this approach may not always be realistic or efficient in educational contexts, such as teachers' use of intelligent tutoring systems (ITS). Instead of comparing a human to an agent, in this context, we need to compare educators' heuristics for learning and the ITS algorithm in the ITS environment [18]. One way to simulate this comparison would be to study the difference between a simple heuristic algorithm and a more complex algorithm used for knowledge tracing.

Knowledge tracing (KT) algorithms have been widely used in education to model and track students' knowledge states during the learning process [10, 42]. These algorithms are employed in ITS to personalize instruction and improve student outcomes [17, 41]. KT has been applied to sequence educational content, leading to improved student performance and engagement compared to expert-designed approaches [11]. Two methods of KT, Bayesian Knowledge Tracing (BKT) and N-Consecutive Correct Responses (N-CCR), have been widely used [22]. N-CCR is based on a simple heuristic: once a student answers N questions correctly in a row, the student has demonstrated proficiency. For example, 3-CCR would determine proficiency after a student answered three questions right in a row. BKT has also been widely used in ITS for student modeling and performance prediction. BKT utilizes Bayesian statistics to maintain internal states of student proficiency. BKT is a two-state Hidden Markov Model where the unobserved hidden state being modeled is student learning, and for a given knowledge component, a student has a state of either learned or not learned [22]. BKT can use individualized parameters and personal priors to constantly update the hidden knowledge state. As students answer questions, the internal state can be updated based on accuracy, time, hints, and other question-related factors. Comparatively, BKT is a more complex algorithm than N-CCR, and fewer people are expected to know the inner workings of BKT initially. This research informs our first research questions.

RQ1: What are the differences in educator perceptions of tools using a simple versus a complex algorithm?

2.3 Explanations as an Intervention

Several studies have demonstrated that providing explanations can reduce algorithm aversion and stimulate intentions to adopt technology [23, 35, 37]. Yet research on algorithmic transparency in educational settings reveals a complex pattern of relationships between explanations, trust, and perceived fairness. Providing explanations for algorithms like BKT can increase trust, perceived accuracy, and user confidence [38]. Explainable AI (XAI) techniques have been increasingly applied in education to enhance trust and confidence in AI tools among educators. These techniques have the potential to significantly improve educators' trust and technology acceptance without increasing cognitive load too much [37]. However, challenges remain, such as tailoring explanations to different stakeholders and the relative nature of explicability across populations and domains [14]. In general, there is a need for more research to understand what types of explanations can reduce algorithm aversion towards education technology for different stakeholder groups. This research informs our second research question.

RQ2: How does explaining a complex algorithm influence educators' attitudes towards tools that use complex or simple algorithms?

We conducted a series of studies to better understand how explanations can influence algorithm aversion and technology adoption. We explore the effects of explanations on attitudes related to algorithm aversion and technology adoption, using a comparison between a simple algorithm (N-CCR) and a more complex one (BKT) with varying levels of explanation.

3 Study 1

In preparation for Study 1, we conducted a pilot study with teachers who actively use the ASSISTments platform (recruited via an email newsletter). Our study design and stimuli are inspired by the Skill Builder feature in the ASSISTments platform, and we were eager to understand how teachers using ASSISTments would react to algorithms. We used feedback from the pilot data to improve study materials to make them more realistic. Notably, even in the small pilot study (N = 39 educators), we saw evidence consistent with the results reported in Studies 1 and 2. Grounded in our research questions and the review of the literature, our goal was to test the following hypotheses in Study 1:

H1. Verbal and visual explanations of BKT lead participants to prefer it over N-CCR.

H2. Verbal and visual explanations of BKT lead participants to have increased ratings of (a) confidence, (b) understanding, (c) sophistication, (d) accuracy, (e) trust, (f) speed, and (g) use about the BKT algorithm.

3.1 Participants

We recruited 170 participants from Prolific. Participants received $1.70 for completing the 10-minute survey, advertised as seeking input on an Adaptive Teaching App. A power analysis conducted with G*Power [15] suggested a target sample size of 170, assuming 95% power to detect a medium effect size of 0.25 at a significance level (alpha) of 0.05. Two of the 170 participants were excluded: one due to a lack of teaching experience and another due to prior experience with the BKT algorithms. Analyses were conducted on the remaining 168 participants. Most participants had significant teaching experience: 44% taught for 11 or more years, 31% for 6-11 years, 23% for 2-5 years, and only 2% for less than a year. The majority identified with feminine pronouns (59%), 36% with masculine pronouns, 2% with non-binary pronouns, and 3% selected "Not Listed" or preferred not to answer the question. Most participants identified as White (81%), followed by Asian (9%), Black or African American (6%), Native Hawaiian or Pacific Islander (1%), and American Indian or Alaska Native (1%); 2% of participants indicated "Not Listed."

3.2 Procedure

Participants were given a narrative introducing a learning tool called Skill Builder that helps teachers determine when a student has learned a particular skill in any subject area. They were informed that Skill Builder uses an algorithm to determine if a student has learned a topic and adjusts the questions accordingly. Participants were then asked to review two algorithms and provide their opinions on both.

Next, they were presented with a description of the 3RR [1] algorithm and a sample report showing multiple students and their progress to proficiency (see Figure 1). Participants then answered questions about their attitudes towards the 3RR algorithm.

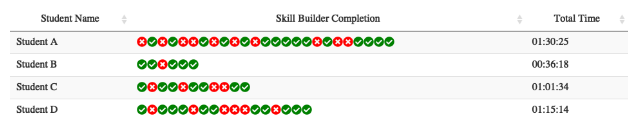


Figure 1: Sample report that participants were shown for Study 1. Participants in the Detailed Visualization condition saw percentages below every fourth question, indicating the student's proficiency.

Participants were randomly assigned to one of four conditions based on a 2x2 factorial design. The next page introduced the BKT algorithm, providing a description and sample learning progress visualization (the content of the explanation and visualization dependent on their assigned condition), followed by the same set of attitudinal questions. Finally, participants were asked to compare the two algorithms and provide a rationale for their preference.

3.3 Experimental Manipulation

In the no-explanation condition, participants received this one-sentence description of the BKT algorithm: "This algorithm determines that a student has mastered a skill once they reach a high probability of mastery based on their responses up to that point."

In the BKT explanation condition, participants additionally received the following information about the BKT algorithm:

The algorithm uses all of the following information to estimate the probability that a student has mastered a skill. If the probability is above 95%, the algorithm determines that the student has mastered the skill.

- *an* ***initial probability*** *that the student has mastered the skill based on their first answer (this will be higher if they answer correctly)*
- *a* ***guess probability*** *for multiple-choice questions (e.g., the chance of guessing correctly is 50% for a True/False question)*
- *a* ***slip probability*** *for answering incorrectly even though the student already mastered the skill (i.e., they accidentally get it wrong)*

[1]We will use the 3RR label for the N-CCR algorithm.

- *the **question difficulty** based on how many other students got it wrong before*
- *other **process information**, such as how many hints the student asked for or how much time it took them to answer the question*

In the BKT simple visualization condition, participants received a simple table depicting a sample student's learning progress mirroring the one shown in the 3RR algorithm condition. In the BKT detailed visualization condition, the same table was enhanced to show the estimated probability of proficiency on the report (Materials on OSF https://osf.io/7c5zt/?view_only=58f0a6314e9243e9acdbd1c844d43d7b).

3.4 Measures

We measured participants' attitudes towards each algorithm using six questions with 5-point unipolar response scales ('Not at all," 'Somewhat," 'Moderately," 'Very," 'Extremely'). We adapted this set of measures from prior work that used them to examine people's attitudes towards algorithms in a similar educational context [38].

Confidence: "How confident are you that the Skill Builder feature with this algorithm will help you teach your students?"

Understanding: "How well do you understand how this algorithm determines if a student has mastered a skill?"

Sophistication: "How sophisticated do you think this algorithm is for determining if a student has mastered a skill?"

Accuracy: "How accurate do you think this algorithm is at determining if a student has mastered a skill?"

Trust: "How much do you trust this algorithm to determine if a student has mastered a skill?"

Speed: "How much effort will it take you to gauge your students' mastery using Skill Builder problem sets with this algorithm?"

Usage Intention: "How likely are you to start using Skill Builder problems with this algorithm in your teaching practice?"

At the end of the survey, participants rated their general preference over the two algorithms in response to the following question (questions adapted from [38]): "You just learned about the 3 Right in a Row (3RR) algorithm and the Bayesian Knowledge Tracing (BKT) algorithm. Which algorithm would you prefer to use for Skill Builder problem sets in your teaching practice?" Response options were on a 7-point bipolar scale: 'Strongly prefer 3RR', 'Moderately prefer 3RR', 'Slightly prefer 3RR', 'Neither prefer 3RR nor BKT', 'Slightly prefer BKT,' 'Moderately prefer BKT,' 'Strongly prefer BKT.' Participants were invited to provide a rationale for their preference using an open-ended question: "Please tell us why you prefer the algorithm that you chose above."

3.5 Results

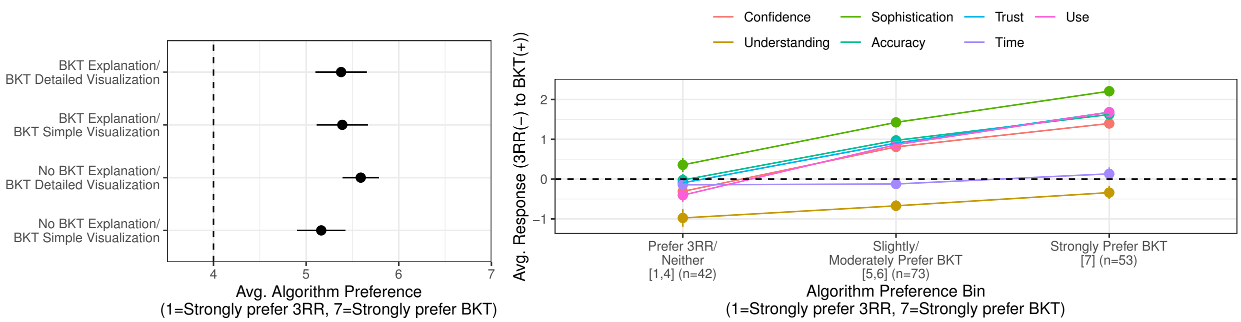
3.5.1 Effects of Explanations on Preference and Attitudes. We examined how algorithm preference varies across experimental conditions by fitting a linear regression model. We find no evidence in support of hypothesis H1, as algorithm preference did not significantly change as a result of adding explanations or visualizations ($F_{3,164} = 0.4454, p = 0.7209$). Figure 2a shows that educators, independent of their experimental condition, tended to prefer BKT on average. While we did not find a significant effect on preference, we observed that the detailed visualization (but not textual explanation) raised confidence in BKT ($b = 0.466, t_{164} = 2.583, p = 0.011$), trust in BKT ($b = 0.4216, t_{164} = 2.124, p = 0.035$), and its perceived accuracy ($b = 0.410, t_{164} = 2.086, p = 0.039$). Thus, we find that detailed visualizations can improve some attitudes related to algorithms. Overall, the finding that educators prefer the complex over the simple algorithm is surprising, considering prior work on algorithm aversion.

3.5.2 Relationship Between Attitudes and Preferences. To analyze the relationship between participants' attitudes and algorithm preference, we calculated the difference between their responses to the 3RR and BKT tools for each attitudinal measure. A negative difference indicated a preference for 3RR, while a positive difference indicated a preference for BKT. For example, a participant who rated the 3RR tool 4 for confidence and the BKT tool 2 would have a difference score of -2, indicating a higher confidence in 3RR. Figure 2b shows the average response on each measure at three levels of preference. Results indicate that all seven measures are positively correlated with preference (all Pearson's r, between 0.17 and 0.71, with all $p < 0.001$). Accuracy, confidence, use, and trust had the strongest correlations with algorithm preference ($r > 0.63$). These four constructs are particularly influential in determining educator's preference for the BKT algorithms. Together, the seven measures explain 60.1% of the variance in preferences ($F_{7,160} = 34.43, p < 0.001$).

While participants in this study generally preferred BKT, our attempts to improve this preference through explanations and visualizations were limited. However, we successfully influence confidence, accuracy, and trust, which are key predictors of algorithm preference. Given the unexpected findings, we wondered if our manipulations could have been more effective if participants had engaged in an activity that would challenge their perceptions of BKT. In the scenario we set up in Study 1, participants might choose the more sophisticated algorithm without hesitation and not react strongly to the explanations because they do not have to be accountable to the algorithmic recommendations or explain them to other stakeholders. To explore this, we changed the design in Study 2 to add an activity designed to reinforce the need to understand and trust the algorithm.

4 Study 2

In light of the surprising findings in Study 1, we updated the scenario-based design in Study 2 to add a realistic activity that would force educators to engage with the algorithm in ways that demand trust and explainability. Specifically, we created a situation in which the algorithm raised a concern about a student's progress, and we asked educators to explain the educational application with its algorithmic prediction to the student's parent/guardian. Our findings from the open-ended questions in Study 1 suggested that teachers might face trust and confidence issues when explaining algorithms to parents or students. Several participants in Study 1 expressed concerns about the complexity of explaining the BKT algorithm to parents and other teachers. Some participants preferred the 3RR algorithm, citing its perceived simplicity and ability to provide positive affirmations to students. These responses highlight the importance of understanding an algorithm when explaining it to others, as noted by Chaushi et al. [5]. These concerns motivated



(a) Average algorithm preference by condition for all 168 participants. Participants consistently preferred the BKT algorithm, regardless of the provided explanations or visualizations.

(b) The average responses on each attitudinal measure demonstrate a significant positive correlation with preference for the BKT algorithm. This trend is evident across three preference levels: Prefer 3RR to Neither, Slightly and Moderately Prefer BKT, and Strongly Prefer BKT.

Figure 2: Study 1 Results

us to design Study 2, which aimed to create a more realistic scenario where educators would need to explain the algorithm to a parent or guardian. This more realistic scenario allowed us to explore further the potential impact of explanations on attitudes and intentions.

In addition to the new study design, we also modified our dependent measures. Rather than using single-item measures, we adapted established scales to assess trust [34] and perceived competence [30]. We also introduced a new measure, informational fairness [34], to better understand users' comprehension of the algorithm. Still grounded in the same overarching research questions, we formulated the following two hypotheses for Study 2:

H3: Educators rate the application with the simple algorithm higher on (a) trust, (b) competence, (c) informational fairness, and (d) usage intention than the same application with the complex algorithm if no further explanation is provided.

H4: Adding an explanation for the complex algorithm increases educators' ratings of (a) trust, (b) competence, (c) informational fairness, and (d) usage intention for the application with the complex algorithm.

The study was preregistered on OSF at https://osf.io/cjadu. All code, data, and supplementary materials for this study are available on OSF.

4.1 Participants

We recruited a total of 300 participants from Prolific. Participants received $2.00 for completing the 10-minute survey, advertised as seeking input on an Adaptive Teaching App. Anyone who participated in Study 1 was excluded from participating in Study 2. All 300 participants were included in the analysis. Most participants had been teaching for 11+ years (46%), followed by 2-5 years (26%), then 6-11 years (23%), and finally 5% who taught for less than a year. The majority identified with feminine pronouns (70%), followed by masculine pronouns (26%), then by non-binary pronouns(1%). 3% of participants preferred not to answer or selected "Not Listed.". Most participants identified as White (86%), followed by Black or African American (7%), followed by Asian (6%), American Indian or Alaska Native (2%), followed by Native Hawaiian or Pacific Islander (1%). 3% of participants indicated "Not Listed."

4.2 Experimental Manipulation

Participants in Study 2 were randomly assigned to one of three conditions with different algorithms and levels of detail about the algorithm. Participants in the 3RR condition were told that the Skill Builder application determines a high level of proficiency once the student answers three questions correctly in a row. Participants in the BKT condition were told that the Skill Builder application uses BKT to determine a high level of proficiency. Participants in the BKT with Explanation condition were also told that the Skill Builder application uses BKT. However, they were provided with a more detailed explanation of BKT (with similar text from Study 1) and a more detailed visualization indicating the percentage of proficiency in their report.

4.3 Procedure

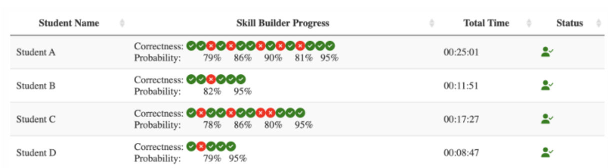


Figure 3: The Skill Builder sample report that was shown to participants in the BKT with Explanation condition. With the exception of the row indicating the Probability, the same report was shown in the other two conditions.

As in Study 1, participants were provided with a contextual narrative that introduced a general learning tool called Skill Builder that helps teachers determine when a student has learned a particular skill in any subject area. In contrast to Study 1, where participants were told they would see two versions of the application (with different algorithms), participants in Study 2 were only asked to give their thoughts on one version (i.e. a between-subjects design).

Participants were randomly assigned to one of three conditions to determine the algorithm they were told that Skill Builder uses. There were 101 participants in the 3RR condition, 104 in the BKT condition, and 95 in the BKT with Explanation condition. Next, participants were given more information about Skill Builder and an introduction to the algorithm (based on their condition assignment) that Skill Builder uses to determine proficiency. At the bottom of the page, they were shown a Skill Builder sample report that visualizes student performance. Figure 3 shows the report that participants were shown for the BKT with Explanation condition. The visualization shows that the student performance in all three conditions was the same. For example, Student A answered 15 questions in all three conditions to reach proficiency. Since we prioritized consistency across the conditions, the BKT proficiency percentages were not created by running the BKT algorithm as we had done in the previous studies (we still selected realistic values).

After the detailed descriptions, participants were introduced to Chris, a student in their class who, according to Skill Builder, has yet to reach proficiency. They were also shown a sample report of Chris's performance, showing that Chris has not reached proficiency. While Chris's performance metrics (Question Answers, Total Time, and Status) were the same across all three conditions, participants in the BKT with Explanation condition did see the proficiency percentages in their report. After reviewing the report, we asked the participants to prepare a note for parent-teacher conferences that addressed Chris's performance on the Skill Builder application, using the following instructions:

For the parent-teacher conference, you decide to write a note explaining Chris's progress with the goal of developing a collaborative plan to help Chris with practicing the skill. Write a note explaining the situation to Chris's parent(s)/guardian(s). In your note, you should:

- *Briefly introduce the Skill Builder tool.*
- *Explain how the tool assessed that Chris requires more practice with the skill.*
  - If choosing a specific skill (i.e., in math, writing, history) helps you write the note, feel free to pick one.

After finishing their note, the participants answered several questions about their perceptions of the Skill Builder application (see Measures) and demographics.

4.4 Measures

We assessed participants' attitudes towards the Skill Builder application using the following preregistered measures:

Usage Intention. Intention to use the Skill Builder application was assessed using the question: How likely or unlikely would you be to use Skill Builder for your subject area in your classroom practice? The participants rated the question on a 5-point Likert scale (Very Unlikely, Somewhat Unlikely, Neither Likely or Unlikely, Somewhat Likely, Very Likely).

Competence. Educators perceptions of competence of the Skill Builder tool was measured with a four-item scale. The four questions asked were: It is effective at recommending decisions on student proficiency; It lacks the expertise to estimate my student's proficiency; It fails to understand the student's level of proficiency; and It performs its role of recommending decisions on student proficiency well. Each question was rated on a 5-point Likert scale (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree). The average score across the four items (Cronbach's alpha = .81) provided a single competence score for each participant.

Informational Fairness. of the Skill Builder application was measured with a four-item scale. The four questions asked were: It explains the decision procedures thoroughly; Its explanations regarding procedures are reasonable; I cannot understand the process by which the decision was made; and I don't have enough information to judge whether the decision procedures are fair or unfair. Each question was rated on a 5-point Likert scale (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree). The average score across the four items (Cronbach's alpha = .82) provided a single informational fairness score for each participant.

Trust. Trust in the Skill Builder application was measured with a six-item scale. The six questions asked were: I trust that it makes high-quality decisions about student proficiency; I believe its decision-making procedure is unbiased; I know it is trustworthy based on my understanding of the decision-making procedure; I think I cannot trust it; It cannot be trusted to carry out student proficiency decisions faithfully; and In my opinion, it is not trustworthy. Each question was rated on a 5-point Likert scale (Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree). The average score across the four items (Cronbach's alpha = .89) provided a single trust score for each participant.

Note to Parents. A thematic analysis was conducted on the participants' responses. Responses were examined for recurring themes and tone, and similar themes were grouped to develop our distinctive themes.

4.5 Results

4.5.1 Impact of Explanation on Usage Intention and Attitudes. We first examine if the participants demonstrated initial algorithm aversion. We fitted a robust linear regression model to estimate the attitudinal and intention differences between participants in the 3RR and BKT conditions (Table 1). Contrary to our hypothesis (H3), participants in the 3RR condition rated the tool lower in perceived competence and trust, and there were no differences between the conditions for usage intention. Consistent with our hypothesis, informational fairness (understanding how the application works) was higher for the 3RR condition. While participants indicated the BKT tool was less understandable, the participants still had high ratings for perceived Competence and trust.

We then examine if providing explanations improved participants' perceptions of BKT (H4). We fitted another linear regression model to estimate the attitudinal and intention differences between participants in the BKT with and without explanation conditions (Table 2). Contrary to our hypothesis (H4), adding explanations to the presentation of the BKT algorithm did not improve any of the attitudinal or preference outcomes. Overall, there were no differences between the two BKT conditions. While not significant, the informational fairness ratings suggested that the BKT with Explanation condition may have helped with understanding the algorithm. The difference in Information Fairness and all other outcomes across conditions is shown in Figure 4.

4.5.2 Exploratory Measures. **Relationship Between Intention and Competence, Informational Fairness, and Confidence.** We examined the relationship between intention and the other attitudinal measures. Regression analysis revealed that competence, trust, and informational fairness collectively explained 36% of the variance in participants' intention to use the tool. This indicates that these factors are significant predictors of educators' intentions

Table 1: Results from robust linear regressions estimating the difference between the 3RR and BKT conditions in each dependent measure. The table shows regression coefficients with standard errors in parentheses.

	Dependent variable:			
	Intention	Informational Fairness	Competence	Trust
BKT	0.190	-0.412^{***}	0.370^{***}	0.236^{**}
	(0.131)	(0.134)	(0.103)	(0.105)
Intercept: *3RR condition*	3.743^{***}	3.621^{***}	3.327^{***}	3.488^{***}
	(0.098)	(0.093)	(0.080)	(0.084)
Observations	205	205	205	205
R^2	0.010	0.045	0.060	0.024
Residual Std. Error (df = 203)	0.936	0.957	0.737	0.752
F-statistic (df = 1; 203)	2.114	9.501^{***}	12.936^{***}	5.048^{**}

Note: Statistical significance of coefficients is indicated by $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 2: Results from robust linear regressions estimating the difference between the BKT and BKT with Explanation conditions in each dependent measure. The table shows regression coefficients with standard errors in parentheses.

	Dependent variable:			
	Intention	Informational Fairness	Competence	Trust
BKT with Explanation	−0.112	0.191	−0.058	0.191
	(0.135)	(0.134)	(0.097)	(0.134)
Intercept: *BKT condition*	3.933^{***}	3.209^{***}	3.697^{***}	3.209^{***}
	(0.087)	(0.096)	(0.066)	(0.096)
Observations	199	199	199	199
R^2	0.003	0.010	0.002	0.010
Residual Std. Error (df = 197)	0.947	0.945	0.684	0.945
F-statistic (df = 1; 197)	0.691	2.025	0.353	2.025

Note: Statistical significance of coefficients is indicated by $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

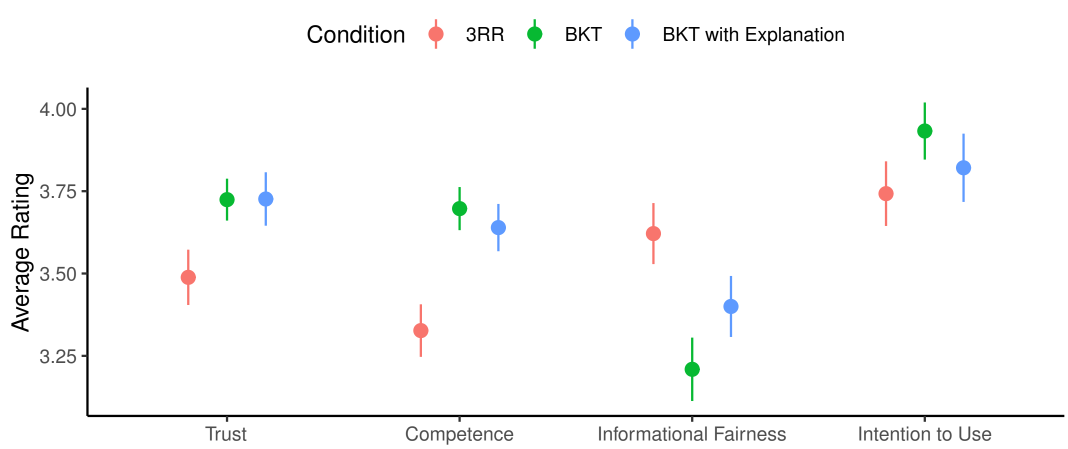


Figure 4: Averages (with standard errors) of attitudinal and usage intention measures by experimental condition. We find that trust and presumed competence were significantly lower in the 3RR than in both of the BKT conditions.

to adopt and use the tool; competence ($b = 0.581, p < 0.001$) and trust ($b = 0.234, p = 0.026$) stood out as most influential in this analysis. This indicates that higher levels of competence and trust are most strongly associated with a higher likelihood of using the application. In contrast, informational fairness showed no statistical relationship with usage intention ($b = 0.004, p = 0.943$).

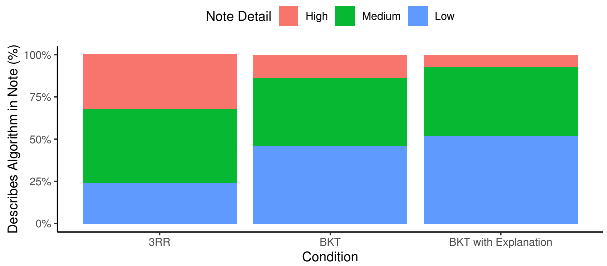


Figure 5: Likelihood of Providing Details about Algorithm in Note by Experimental Condition. Participants in the 3RR condition provided more detailed information about the algorithms in their notes to parents/guardians compared to those in the BKT conditions, regardless of whether an explanation was provided.

Likelihood of Providing Details about Algorithm in Note by Experimental Condition. To better understand participants' attitudes and experiences, we analyzed the notes they wrote to parents. We extracted two types of themes from this analysis: deductive themes related to the level of detail about Skill Builder and inductive themes that emerged from the content of the notes.

We categorized the notes into three levels of detail about the algorithm: Low Information, Medium Information, and High Information (Table 3). 40% of responses were tagged as Low Information, 41% as Medium Information, and 19% as High Information. Figure 5 visualizes the distribution of note details by condition. As expected, participants in the 3RR condition provided the most detail about the Skill Builder application. This finding and the informational fairness ratings suggest that 3RR was the easiest to understand and explain.

Outside of the three themes about algorithm details in the note, we identified 12 other inductive themes from the notes. The complete list of themes can be found on OSF. While helpful for providing a deeper interrogation into the participants' thoughts, most themes were irrelevant to our initial hypothesis (H3 and H4), except for one theme, which we called "Data from Reports." The Data from the Reports theme was coded anytime a participant referenced the visualized report in their note. This code came in the form of mentioning how many questions Chris answered correctly, adding in the proficiency percentages, or indicating how much time Chris took on the questions. It gives some evidence that the participants found the report understandable and clear. For example, one participant wrote, "[Chris] needs to obtain a higher percentage score. Skill Builder recorded scores of 52%, 64% and 46%," and another participant wrote, "It looks as if it's getting harder for [Chris] since his score has dropped so he will need more practice." About 30% of the participants included some aspect of the reports in their note. While the prevalence of this code did not vary by condition ($F_2 = 0.283, p = 0.754$), there was an interesting pattern with the informational fairness dependent measure. Those reporting information from the report in their note appeared to have higher information fairness ratings. Thus, including data from the report could be an additional measure of understanding and clarity of the data.

5 Discussion

In this paper, we conducted two studies to understand the effects of providing algorithm transparency in additional text and detailed visualizations of educators' attitudes and intentions to use AI-powered tools. In Study 1, we provided educators with two algorithms, a simple one (3RR) and a complex one (BKT), and they overwhelmingly preferred and had generally positive attitudes toward the tool with the complex algorithm regardless of whether they were provided transparency into the algorithm. In Study 2, we exposed each educator to only one algorithm but added an explanation task where educators needed to explain the tool to a parent or guardian. Replicating our finding from Study 1, we found that educators had positive attitudes toward the complex algorithm regardless of the level of transparency they received. Study 2 also indicated that understanding how the algorithm works (informational fairness) is not critical to predicting future usage intention of the tool. Revisiting our two RQs, we found across both studies that educators have positive attitudes towards both the simple and the complex algorithm (RQ1), and the addition of explanations has almost no effect on attitudes (RQ2). In light of these findings, we discuss whether algorithm aversion has dissipated.

5.1 Algorithm Appreciation

Algorithm aversion was theorized when initial research on people's attitudes toward using algorithms to assist them in decision-making tasks indicated a sense of resistance to tools that use algorithms [12]. However, recent research suggests that this trend has dissipated or reversed in what researchers call algorithm appreciation [25, 40]. Algorithm appreciation is the tendency to rely on algorithmic advice over human advice, which might help explain the results of the studies in this paper. While we initially hypothesized that educators would demonstrate algorithm aversion and prefer simple heuristics to assist them in decision-making tasks, we found that educators believe that a complex algorithm is better for evaluating mastery learning. Turel and Kalhan [35] framed algorithm aversion as an implicit bias or prejudice against AI and reasoned that, like other types of prejudices, exposure to AI would move people from aversion to appreciation. Exposure to AI frequently comes from an increased usage of generative AI (GenAI) applications. Educators have employed GenAI tools like ChatGPT and Bing Image Creator for lesson planning, brainstorming, and professional development [6, 33]. It may be that the recent explosion of GenAI tools geared toward educators has changed educators' perceptions of algorithms from aversion to appreciation. Even in these two studies, we saw evidence of a change in attitudes over time. We conducted Study 1 in December 2022 and Study 2 in August 2024. In that time, we saw an increase in average differences in attitudes and intentions towards BKT. For example, in Study 1, the average intent to use the tool with BKT was 2.899 ($se = 0.083$), but in Study 2, it was 3.880 ($se = 0.067$) for educators in both BKT conditions. Trust displayed the same pattern, 3.130 for Study 1 and 3.725 for Study 2. While an imperfect measure, this shift suggests that the

Table 3: Descriptions of the themes about algorithm details in the note.

Theme	Definition	Example
Low Information	No or very little specific details about Skill Builder.	"Hello, we have a new tool to help Chris identify where he needs help. It uses AI to determine where his weakest points are. Please improve the indicated areas."
Medium Information	There is mention of a tool that helps identify low proficiency areas along with either information about how it determines proficiency or acknowledgment of what happens after proficiency is achieved.	"Skill Builder is an online tool that helps students review content and master skills by answering questions and trying to build streaks of correct answers."
High Information	Skill Builder is explained with clear definitions of how it determines proficiency and what happens after proficiency is achieved.	"This year, I have began to use a tool call Skill Builder to help my students gain proficiency in their skills. So often in education, we tend to move on to new topics without ensuring that students have mastered the current ones. Skill Builder allows me to see that students are ready to move to new material after they can successfully answer three consecutive questions."

increased usage of GenAI may have led to more favorable attitudes toward algorithms in general.

5.2 AI Literacy

The growing prevalence of AI literacy initiatives specifically designed for educators may have played a significant role in shifting their attitudes toward algorithms. Research suggests that increased AI literacy correlates with educators' willingness to learn and use AI-powered tools [13]. There has been a focus on integrating AI literacy into teacher education programs, including technical skills, ethical considerations, and practical applications [2]. This could have contributed to an evolution from algorithm aversion to appreciation in the present context because educators with higher AI literacy are more likely to have positive attitudes and intentions regarding AI-based education technology. Our findings underscore recent calls in the literature to focus on psychological and social factors in technology adoption [9, 24]. Therefore, it is essential to provide ongoing training and support still to help educators develop the skills and knowledge needed to use AI tools effectively. While algorithm aversion might have had a limited impact, understanding the factors that facilitated this shift is crucial for designing and implementing AI-powered tools effectively in education. Tailoring these tools to individual needs and perceptions is essential for successful adoption.

5.3 Limitations and Future Research

This research responds to calls to better understand educators' needs and perspectives on AI-powered tools. Our results indicate a few possible future research areas, partly inspired by the limitations of our research designs. First, the data collected was from educators on Prolific [31] and is limited to the US, thus it may not be representative of all educators. Further, this study was designed for the context of ITS, however there are many more uses of AI in education.

We chose BKT as our complex algorithm, even though it is not a black-box algorithm. We favored the ability to explain the algorithm to novice users over a truly black-box algorithm. Our results show positive trends in understanding and informational fairness when additional explanations are provided, indicating that it was, on average, less understandable without any explanation. However, the gap is small, and running this type of study with a more complex algorithm may provide lower values for the complex algorithm without explanation, allowing for a greater space for the manipulation to work.

Another popular study design in this literature is about understanding AI attitudes and intentions where the AI gives wrong answers. We designed our stimulus to provide realistic recommendations that educators would agree with and want to adopt. Another study could look at the effects of AI on student performance, namely when it gives a mix of right and obvious wrong recommendations. We also encourage researchers to understand this phenomenon in real-world educational contexts. Although we were limited by the constraints of a 10-minute online experimental study, additional research should consider long-term observational studies of these attitudes with teachers exploring AI-powered tools in their actual classrooms.

Study 2 asks participants to explain the tool to a parent/guardian. The last area we encourage researchers to explore is experimenting with the context of how they need to explain the algorithm. The task of explaining the tool to a parent may have unintentionally influenced participants to report more positive views, as people are generally more likely to present a positive image of a classroom tool when communicating with parents. Future research could consider creating a narrative where the participant explains the algorithm and tool to a principal or school leader, along with their feedback about if and how to use the tool. This design might give more space for participants who disagree with the tool to indicate more negative sentiment.

References

[1] Clement Chimezie Aladi. 2024. IT Higher Education Teachers and Trust in AI-Enabled Ed-Tech: Implications for Adoption of AI in Higher Education. In *Proceedings of the 2024 Computers and People Research Conference*. ACM, Murfreesboro TN USA, 1–16. https://doi.org/10.1145/3632634.3655852

[2] Musa Adekunle Ayanwale, Owolabi Paul Adelana, and Tolulope Timothy Odufuwa. 2024. Exploring STEAM teachers' trust in AI-based educational technologies: a structural equation modelling approach. *Discover Education* 3, 1 (April 2024), 44. https://doi.org/10.1007/s44217-024-00092-z

[3] Simon Buckingham Shum, Rebecca Ferguson, and Roberto Martinez-Maldonado. 2019. Human-Centred Learning Analytics. *Journal of Learning Analytics* 6, 2 (July 2019). https://doi.org/10.18608/jla.2019.62.1

[4] Noah Castelo, Maarten W. Bos, and Donald R. Lehmann. 2019. Task-Dependent Algorithm Aversion. *Journal of Marketing Research* 56, 5 (Oct. 2019), 809–825. https://doi.org/10.1177/0022243719851788
[5] Blerta Abazi Chaushi, Besnik Selimi, Agron Chaushi, and Marika Apostolova. 2023. *Explainable Artificial Intelligence in Education: A Comprehensive Review*. Communications in Computer and Information Science, Vol. 1902. Springer Nature Switzerland, Cham, 48–71. https://doi.org/10.1007/978-3-031-44067-0_3
[6] Shengnan Chen, Qifang Liu, and Bin He. 2023. A Generative AI-based Teaching Material System Using a Human-In-The-Loop Model. In *2023 International Conference on Intelligent Education and Intelligent Research (IEIR)*. IEEE, Wuhan, China, 1–8. https://doi.org/10.1109/IEIR59294.2023.10391244
[7] Lingwei Cheng and Alexandra Chouldechova. 2023. Overcoming Algorithm Aversion: A Comparison between Process and Outcome Control. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–27. https://doi.org/10.1145/3544548.3581253
[8] Seongyune Choi, Yeonju Jang, and Hyeoncheol Kim. 2023. Influence of Pedagogical Beliefs and Perceived Trust on Teachers' Acceptance of Educational Artificial Intelligence Tools. *International Journal of Human–Computer Interaction* 39, 4 (Feb. 2023), 910–922. https://doi.org/10.1080/10447318.2022.2049145
[9] Mutlu Cukurova, Xin Miao, and Richard Brooker. 2023. *Adoption of Artificial Intelligence in Schools: Unveiling Factors Influencing Teachers' Engagement*. Lecture Notes in Computer Science, Vol. 13916. Springer Nature Switzerland, Cham, 151–163. https://doi.org/10.1007/978-3-031-36272-9_13
[10] Miao Dai, Jui-Long Hung, Xu Du, Hengtao Tang, and Hao Li. 2021. Knowledge Tracing: A Review of Available Technologies. *Journal of Educational Technology Development and Exchange* 14, 2 (2021), 1–20. https://doi.org/10.18785/jetde.1402.01
[11] Yossi Ben David, Avi Segal, and Ya'akov (Kobi) Gal. 2016. Sequencing educational content in classrooms using Bayesian knowledge tracing. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK '16*. ACM Press, Edinburgh, United Kingdom, 354–363. https://doi.org/10.1145/2883851.2883885
[12] Berkeley J Dietvorst, Joseph P Simmons, and Cade Massey. 2015. Algorithm aversion: people erroneously avoid algorithms after seeing them err. *Journal of experimental psychology: General* 144, 1 (2015), 114.
[13] Hua Du, Yanchao Sun, Haozhe Jiang, A. Y. M. Atiquil Islam, and Xiaoqing Gu. 2024. Exploring the effects of AI literacy in teacher learning: an empirical study. *Humanities and Social Sciences Communications* 11, 1 (May 2024), 559. https://doi.org/10.1057/s41599-024-03101-6
[14] Robert Farrow. 2023. The possibilities and limits of explicable artificial intelligence (XAI) in education: a socio-technical perspective. *Learning, Media and Technology* 48, 2 (April 2023), 266–279. https://doi.org/10.1080/17439884.2023.2185630
[15] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39, 2 (May 2007), 175–191. https://doi.org/10.3758/BF03193146 00000.
[16] Hongwen Guo, Matthew Johnson, Kadriye Ercikan, Luis Saldivia, and Michelle Worthington. 2024. Large-Scale Assessments for Learning: A Human-Centred AI Approach to Contextualizing Test Performance. *Journal of Learning Analytics* 11, 2 (Aug. 2024), 229–245. https://doi.org/10.18608/jla.2024.8007
[17] Yann Hicke. 2023. Knowledge Tracing Challenge: Optimal Activity Sequencing for Students. (2023). https://doi.org/10.48550/ARXIV.2311.14707
[18] Kenneth Holstein, Bruce M. McLaren, and Vincent Aleven. 2017. Intelligent tutors as teachers' aides: exploring teacher needs for real-time analytics in blended classrooms. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. ACM, Vancouver British Columbia Canada, 257–266. https://doi.org/10.1145/3027385.3027451
[19] Leanna Ireland. 2020. Who errs? Algorithm aversion, the source of judicial error, and public support for self-help behaviors. *Journal of Crime and Justice* 43, 2 (March 2020), 174–192. https://doi.org/10.1080/0735648X.2019.1655781
[20] Yuze Kang. 2022. Algorithm Aversion and Self-driving Cars:. Zhuhai, China. https://doi.org/10.2991/aebmr.k.220307.077
[21] Esther Kaufmann. 2021. Algorithm appreciation or aversion? Comparing in-service and pre-service teachers' acceptance of computerized expert models. *Computers and Education: Artificial Intelligence* 2 (2021), 100028. https://doi.org/10.1016/j.caeai.2021.100028
[22] K. Kelly, Yan Wang, Tamisha Thompson, and N. Heffernan. 2015. Defining Mastery: Knowledge Tracing Versus N- Consecutive Correct Responses. In *Educational Data Mining*.
[23] René F. Kizilcec. 2016. How Much Information?: Effects of Transparency on Trust in an Algorithmic Interface. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, San Jose California USA, 2390–2395. https://doi.org/10.1145/2858036.2858402
[24] René F. Kizilcec. 2024. To Advance AI Use in Education, Focus on Understanding Educators. *International Journal of Artificial Intelligence in Education* 34, 1 (March 2024), 12–19. https://doi.org/10.1007/s40593-023-00351-4
[25] Jennifer M. Logg, Julia A. Minson, and Don A. Moore. 2019. Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes* 151 (March 2019), 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005
[26] Hasan Mahmud and Najmul Islam. 2023. *The Role of Algorithm and Task Familiarity in Algorithm Aversion: An Empirical Study*. Lecture Notes in Computer Science, Vol. 14316. Springer Nature Switzerland, Cham, 3–13. https://doi.org/10.1007/978-3-031-50040-4_1
[27] Tanya Nazaretsky, Mutlu Cukurova, and Giora Alexandron. 2022. An Instrument for Measuring Teachers' Trust in AI-Based Educational Technology. In *LAK22: 12th International Learning Analytics and Knowledge Conference*. ACM, Online USA, 56–66. https://doi.org/10.1145/3506860.3506866
[28] Tanya Nazaretsky, Mutlu Cukurova, Moriah Ariely, and Giora Alexandron. 2021. Confirmation bias and trust: human factors that influence teachers' attitudes towards AI-based educational technology. In *CEUR Workshop Proceedings*, Vol. 3042.
[29] David T Newman, Nathanael J Fast, and Derek J Harmon. 2020. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. *Organizational Behavior and Human Decision Processes* 160 (2020), 149–167. https://doi.org/10.1016/j.obhdp.2020.03.008
[30] Jeroen Ooge, Shotallo Kato, and Katrien Verbert. 2022. Explaining Recommendations in E-Learning: Effects on Adolescents' Trust. In *27th International Conference on Intelligent User Interfaces*. ACM, Helsinki Finland, 93–105. https://doi.org/10.1145/3490099.3511140
[31] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. 2017. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70 (2017), 153–163. https://doi.org/10.1016/j.jesp.2017.01.006
[32] Fen Qin, Kai Li, and Jianyuan Yan. 2020. Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology* 51, 5 (Sept. 2020), 1693–1710. https://doi.org/10.1111/bjet.12994
[33] Pati Ruiz, Alessandra Rangel, and Merijke Coenraad. 2024. Using Generative AI to Support PK-12 Teaching and Learning: Developing Sample Lessons and More. In *Proceedings of the 55th ACM Technical Symposium on Computer Science Education V. 2*. ACM, Portland OR USA, 1800–1801. https://doi.org/10.1145/3626253.3635522
[34] Jakob Schoeffer, Niklas Kuehl, and Yvette Machowski. 2022. "There Is Not Enough Information": On the Effects of Explanations on Perceptions of Informational Fairness and Trustworthiness in Automated Decision-Making. In *2022 ACM Conference on Fairness, Accountability, and Transparency*. ACM, Seoul Republic of Korea, 1616–1628. https://doi.org/10.1145/3531146.3533218
[35] Ofir Turel and Shivam Kalhan. 2023. Prejudiced against the Machine? Implicit Associations and the Transience of Algorithm Aversion. *MIS Quarterly* 47, 4 (Dec. 2023), 1369–1394. https://doi.org/10.25300/MISQ/2022/17961
[36] Olga Viberg, Mutlu Cukurova, Yael Feldman-Maggor, Giora Alexandron, Shizuka Shirai, Susumu Kanemune, Barbara Wasson, Cathrine Tømte, Daniel Spikol, Marcelo Milrad, Raquel Coelho, and René F. Kizilcec. 2023. What Explains Teachers' Trust of AI in Education across Six Countries? (2023). https://doi.org/10.48550/ARXIV.2312.01627
[37] Deliang Wang, Cunling Bian, and Gaowei Chen. 2024. Using explainable AI to unravel classroom dialogue analysis: Effects of explanations on teachers' trust, technology acceptance and cognitive load. *British Journal of Educational Technology* (April 2024), bjet.13466. https://doi.org/10.1111/bjet.13466
[38] Kimberly Williamson and René F Kizilcec. 2021. Effects of Algorithmic Transparency in Bayesian Knowledge Tracing on Trust and Perceived Accuracy. *International Educational Data Mining Society* (2021).
[39] Lingrui Xu, Zachary A. Pardos, and Anirudh Pai. 2023. Convincing the Expert: Reducing Algorithm Aversion in Administrative Higher Education Decision-making. In *Proceedings of the Tenth ACM Conference on Learning @ Scale*. ACM, Copenhagen Denmark, 215–225. https://doi.org/10.1145/3573051.3593378
[40] Sangseok You, Cathy Liu Yang, and Xitong Li. 2022. Algorithmic versus Human Advice: Does Presenting Prediction Performance Matter for Algorithm Appreciation? *Journal of Management Information Systems* 39, 2 (April 2022), 336–365. https://doi.org/10.1080/07421222.2022.2063553
[41] Qiao Zhang and Christopher J Maclellan. 2021. Going Online: A Simulated Student Approach for Evaluating Knowledge Tracing in the Context of Mastery Learning. *International Educational Data Mining Society* (2021).
[42] Aayesha Zia, Jalal Nouri, Muhammad Afzaal, Yongchao Wu, Xiu Li, and Rebecka Weegar. 2021. A step towards Improving Knowledge Tracing. In *2021 International Conference on Advanced Learning Technologies (ICALT)*. IEEE, Tartu, Estonia, 38–39. https://doi.org/10.1109/ICALT52272.2021.00019