Surveying Contextualized Student Data Sharing Preferences for Educational AI

Frank Stinar[1[0009−0000−3152−6281]] and Nigel Bosch[1[0000−0003−2736−2899]]

University of Illinois Urbana–Champaign, Champaign, IL, USA
{fstinar2,pnb}@illinois.edu

Abstract. The integration of artificial intelligence (AI) into educational systems provides numerous opportunities for personalized learning while also creating concerns about ethical data usage. This study investigates students' perceptions of how educational AI systems should handle academic, personal, and demographic information. We survey college students from the United States ($N = 128$) to explore their views on permissible data types, performance thresholds for using demographics, and shifts in opinion when personally impacted. We found that most students reject the usage of political affiliation, yet many still accept using some sensitive information in models such as demographics ($p < .005$, 95% CI [0.763, 0.893]). Personal stakes in the system result in shifts in data-sharing preferences. Additionally, some individual differences, such as gender, correlate with willingness to share sensitive information. These findings emphasize the importance of involving students as stakeholders to design ethical and equitable educational AI systems.

Keywords: student perspectives · stakeholders · learning analytics

1 Introduction

The usage of machine learning and artificial intelligence (AI) systems in education has grown significantly over the last decade. These systems provide new opportunities to enhance learning through such as adaptive tutoring [25] and intervention design [14]. However, these systems require student data (often vast amounts) to be effective. This has led to many questions about the ethical and practical implications of using student data for these systems [31, 36].

Obtaining consent in educational settings has long been an important facet of AI in education, particularly when it involves decisions that directly affect students. While teachers, school administrators, and parents are often asked to provide consent, research suggests that students may be less involved in consent processes than it seems [34, 41]. The lack of student engagement in the consent process raises important questions about students' agency and autonomy in these decisions (or even if current informed consent processes can be adequately transformed to handle large amounts of education data from computer-based learning [18]).

In the field of AI in education, there have been many calls to involve students more in the decision-making process of implementing AI systems into learning environments [29]. These systems often rely on diverse types of information, such as gaze data [15], academic data [21], student behaviors [7], or sometimes demographic information collected from students [28]. However, it is unclear if students accurately understand how and why their data are collected [18, 22].

Recognizing the possible gap in student perspectives, researchers have begun to query students to understand their concerns with their data being used in educational systems [24, 26, 27, 40]. However, there remains a need to understand if student concerns are relative to specific types of data or the group of students that the data collected could serve.

In this paper, we work to understand the complexity of student data-sharing concerns by examining students' general opinions on using sensitive information within AI models for educational applications given different levels of stakes in the model decisions. To answer preregistered research questions [1], we created a survey that queried students on their opinions of different types of data that could be used in educational AI systems given scenarios that are increasingly related to—or contrasting with—the students' own self-described identity. By better understanding the complexity of student opinions, we contribute to a growing understanding of student opinions on data collection for AI in education.

Research Question 1: How much do students think that machine learning systems should be aware of their own information? We anticipate that students will have nuanced opinions regarding the types of data machine learning systems should utilize. Specifically, we predict that students will generally permit the use of non-sensitive educational information (e.g., academic performance or attendance), while finding sensitive demographic-based features, such as ethnicity or gender, less permissible to use. We predict this due to related research in procedural fairness, wherein certain types of information are more protected than others [23].

Furthermore, we anticipate the emergence of three distinct groups of students based on their responses: (i) students who permit no data to be used by machine learning systems, (ii) students who permit all types of data to be used, and (iii) students who selectively allow certain types of data and systematically reject others.

Research Question 2: Do students believe there are performance thresholds in machine learning systems that justify the use of sensitive features of students? We expect that the majority of students will agree that there is no level of improvement within machine learning system performance that would justify using a student's own sensitive and demographic information. Our hypothesis is based on the idea that students will prioritize ethical and privacy concerns over potential algorithmic gains [6, 26, 35]. While students could recognize that such information could improve machine learning performance, we predict that they view the trade-off as unacceptable due to concerns about the fairness of the system, systemic biases, and the possible misuse of sensitive information [16].

[1] `https://osf.io/cms5e`

Research Question 3: When students are given a machine learning model that impacts their own scores, do their opinions change on what types of features can be justified? We hypothesize that when a machine learning model directly impacts a student, that student will change their opinions on what types of information that model can use. Specifically, we expect students to become more pragmatic and flexible in their views when a model's decisions have a direct impact on their own scores. Due to students having decisions dependent on model predictions, students may value the performance gains more than their own ethical concerns. However, we do not anticipate the change in opinions will be consistent across all participants. We suspect some students will remain consistent in rejecting the usage of sensitive features due to their own values in privacy or fairness [35].

Research Question 4: Do individual difference measures between students impact how much data they are willing to give to machine learning algorithms meant to better education? To explore this question, we examine how the specific individual differences of students (i.e., political affiliation and ethical ideology) shape students' willingness to provide personal and educational information with machine learning systems. We predict that students' political leanings will influence the information they are willing to share. Particularly, we anticipate more liberally leaning students will be more inclined to permit a wider range of information. In contrast, students who self-identify as conservative are predicted to allow fewer types of information. We expect this because, for example, research has shown that Republicans in the U.S. have become less willing to share some of their sensitive information post-COVID-19 [12].

Drawing on Forsyth's Ethics Position Questionnaire (EPQ), we predict that ethical positions will influence students' preferences for information sharing [10]. Specifically, students with higher idealistic tendencies—i.e., students who prioritize universally positive outcomes—will exhibit more caution regarding information sharing. We expect this since they may perceive fundamental risks or ethical dilemmas in sharing types of sensitive or demographic information. In contrast, more relativistic students —i.e., those who view moral decisions as dependent on the context—will be more flexible and pragmatic in their willingness to share information (especially if there are tangible benefits such as improved educational outcomes).

2 Background

The integration of machine learning into education has led to debate regarding the ethical implications and practical limitations of these technologies. We position our research within the advent of AI-powered algorithms for education and ethical concerns of using student data.

Modern educational practices have been shaped by machine learning-powered educational algorithms. These algorithms have been designed to analyze student data to improve learning outcomes and personalize instruction [5]. For example, algorithms have been used to predict student grades [21], mastery [20], or even if students are at risk of dropping out [33]. Recent work highlights the potential for these systems to positively transform education. Sha et al. identified dropout in Massive Open Online Courses by analyzing student engagement through Hidden Markov Models [33]. These recent studies emphasize the positive gains possible from machine learning-driven educational systems.

Alongside improving education, researchers are increasingly concerned with the ethics of applying these systems in practice as they harness large amounts of, often sensitive, student data. The use of such algorithms has raised critical questions about the types of data they should be allowed to utilize. Non-sensitive academic information such as academic performance or attendance are often permissible in educational machine learning since they directly relate to the outcome [2]. However, sensitive demographic features, such as ethnicity, gender, or socioeconomic status, are more contentious due to concerns about discrimination and privacy violations [26, 36].

To responsibly use student data, education researchers turned to insights from related fields such as algorithmic fairness [3] and data anonymization [38,42] to uphold student protections. For example, researchers have conducted audits of educational systems to identify bias [43], while others have applied strategies to mitigate the biases in their systems [4, 13, 17]. These efforts emphasize that ethical considerations must be within the design of educational AI systems.

While applying ethical frameworks and the technical advancements in protecting student data are integral to safe educational systems, involving a broader range of stakeholders in the design of these systems is important to ensure that they are equitable, transparent, and aligned with the values of students and educators. Considering many countries integrating AI into their education curriculum [1, 8], involving stakeholders can build trust in educational technologies which are closely tied to perceptions of fairness [27]. Recently, Sanusi et al. found that students, teachers, and policymakers had different conceptions of AI [32]. While there have been multiple studies measuring teachers' opinions of AI and contextualizing how teachers and AI ought to interact [9, 19, 44], understanding student opinions of AI is in its infancy. In early work thus far, Li et al. determined that a student's propensity to consent to educational data usage is based on trust in the system, concerns regarding personal data use, and understanding of the perceived usefulness of their data [24]. Nazaretsky et al. also measured student trust in adopting AI-powered educational technology [27].

Alongside student trust in AI, research has begun in dissecting the opinions students have about the types of data that are collected in educational technologies [40]. We further study the nuance students have in permitting different data types in educational AI systems by examining if students are consistent in their responses given different levels of stakes in the system.

3 Method

Our study utilized a randomized experimental design to investigate students' perceptions of fair artificial intelligence practices in educational settings. Students completed an online survey to measure individual differences, ethical beliefs, and general perceptions of machine learning systems for education.

We recruited participants through Prolific, an online recruitment platform with excellent data quality and participant selection capabilities [30]. To be eligible for our study, participants needed to be students currently enrolled in a college or university within the United States. We collected 135 student responses through Prolific, with each student compensated $11.25 USD upon completion. Seven responses were removed due to incorrect responses on attention checks within the study, resulting in a final sample of 128 students.

Students first completed a background survey to measure demographics and individual difference measures (i.e., race, gender, age, university, and political affiliation) and ethical ideology (from Forsyth's Ethics Position Questionnaire [10]). Students provided open-ended responses about their gender and race. Questions at later stages of the survey were adapted based on students' responses to the background survey questions. Throughout the remainder of the survey, the students were introduced to various hypothetical educational AI scenarios. We measured Likert-type responses and yes/no measures to understand the students' perceptions of how much machine learning systems should be aware of their own self (in terms of potentially personal characteristics) and the students' own opinions on how much demographic information is permissible in machine learning models for education.

Two of the questions were tailored to students' own responses. That is, we manipulated a variable to provide students with questions that were about individuals within their own demographic group membership, and outside of their group membership. This approach was designed to examine how in-group and out-group dynamics influence students' responses. This is based on prior research showing that individuals exhibit biases when engaging with content related to their own identities [39]. These questions asked students to choose between three options for a machine learning-driven educational tutor. Each option is given below:

- *Tutor 1*: This tutor improves student scores by two letter grades on average by automatically adapting its feedback based on data on grades, age, family income, race, and gender.
- *Tutor 2*: This tutor improves student scores by one letter grade on average by automatically adapting its feedback based on performance in other classes.
- *Tutor 3*: This tutor improves student scores by half of one letter grade on average. This tutor doesn't use any other data from the students. The tutor gives each student the same feedback.

To ensure that we measured comprehension without the influence of prior knowledge or any real-world biases, the last section of the survey involved a short quiz based on a nonsense, Seussian world. Furthermore, we informed the students that they would finish the participation early if they performed well on the quiz to motivate the students to do well on the quiz. Before the quiz, Students were instructed to determine which types of information educational AI systems could use (i.e., the individual's past grades, past grades of other students, demographic information, political affiliation, and information regarding time spent during the learning session). Then, after the quiz, students were instructed that the quiz would be graded based on the types of information they allowed to be used by educational AI systems and they were prompted to again determine which types of information educational AI systems could use. However, student quiz answers were simply graded based on accuracy, and those who failed to score at least 50% on the quiz were given an extra question at the end of the survey about using student data in educational systems.

Thus, our study consisted of a five-question introductory survey, followed by the EPQ [10], and 12 items on student opinions on data usage in educational machine learning systems.

3.1 Data Analysis

The data collected was analyzed with parallel factor analysis, significance tests, Kruskall–Wallis tests, Wilcoxon tests, and Mann–Whitney tests.

The data consisted of a diverse set of students, including 79 men, and 49 women, and race/ethnicity responses including white (49), Black (40), Asian (14), Latinx or Hispanic (14), and multiracial (11). Our students were less diverse in terms of their responses to the EPQ [10], where 90 students were considered more idealistic than relativistic, 34 were more relativistic, and 4 students were equally idealistic and relativistic. By splitting the responses of the political ideology scale into three ranges, we had 72 students identify as liberal, 28 identify as conservative, and 28 self-identified as independent (not leaning liberal or conservative).

For RQ1, *how much do students think that machine learning systems should be aware of their own information*, we conducted an exploratory factor analysis on the questions referenced in Table 1 to determine which questions map onto which constructs from our survey. We used parallel factor analysis (Figure 1) and found three distinct factors. The first factor (F1), consisting of items related to levels of information allowed to improve educational models, measures a student's opinion on the types of information that are permissible within educational AI support. The second factor (F2) includes the intergroup and intragroup items; thus, F2 measures students' opinions on group-specific modeling outcomes impacting the information that can be used in AI support. Finally, the last factor (F3) includes the three items that ask if there is any increase in accuracy that justifies the usage of certain types of information (one item, aip5 is not included due to low loadings onto each of the three factors). Thus, the F3 factor measures if a student feels that there is a level of support that justifies using certain types of sensitive information.

For RQ2, *do students believe there are performance thresholds in machine learning systems that justify the use of sensitive features of students*, we performed Z-tests to determine if students believed different types of features should be treated differently in terms of use. We specifically looked at the questions, ai_added_1, ai_added_2, and ai_added_3, because these questions each ask

Table 1. A brief table summarizing the 12 survey items. aip6 and aip7 have a manipulated variable based on the students' reported gender.

Item	Description	Type
aip1	Indicate how much you agree with the following statement. It is ethical to provide students with support that is determined by artificial intelligence algorithms.	Likert
aip2	Indicate how much you agree with the following statement. It is ethical to use information about past academic performance for each student.	Likert
aip3	Indicate how much you agree with the following statement. It is ethical to use information about each student's demographics.	Likert
aip4	Imagine your school wants to improve the grades of students in math courses. Your school currently has a bias in math scores where women have historically received worse scores on math exams. To improve this the school developed an artificial intelligence algorithm that helped by automatically scoring tests, partly based on demographic features of each student. It is ethical to include demographic information in this case.	Likert
aip5	There is an artificial intelligence algorithm that can predict a student's grades after their first two weeks of classes. The algorithm is 70% accurate. It was recently found that if the program began using demographic data about the students, the accuracy would increase. How much would accuracy have to be to justify using demographic data?	Range
aip6	It is determined that the underperformance of [ingroup gender] is due to inequalities in resources available at local schools. Which computer-based tutor should be used to help students succeed?	Radio
aip7	It is determined that the underperformance of [outgroup gender] is due to inequalities in resources available at local schools. Which computer-based tutor should be used to help students succeed?	Radio
ai_added_1	Is there any increase in accuracy that would justify using demographic data in an AI-powered tutoring system?	Yes/no
ai_added_2	Is there any increase in accuracy that would justify using past grades in an AI-powered tutoring system?	Yes/no
ai_added_3	Is there any increase in accuracy that would justify using political affiliations in an AI-powered tutoring system?	Yes/no
pre_LS	Select which types of data an AI-powered tutor can use.	Checkboxes
post_LS	Select which types of data an AI-powered tutor can use.	Checkboxes

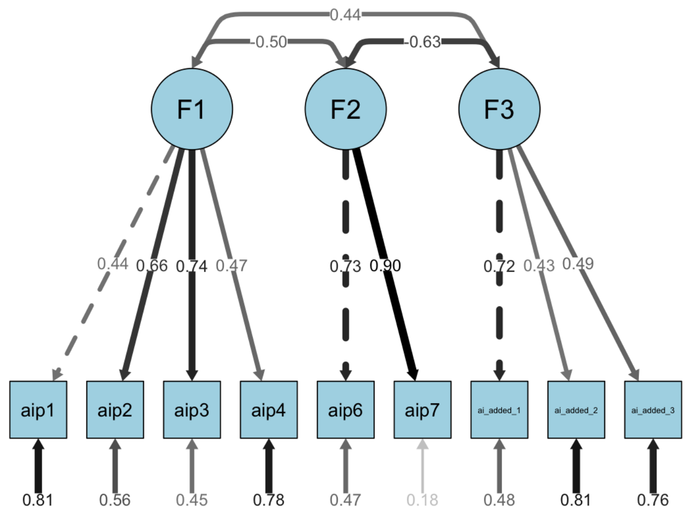

Fig. 1. Structural equation model of how the items outlined in Table 1 map onto three factors. Each item loads with at least a weight of .4. F1 denotes a factor related to levels of information allowed to improve educational models. F2 is a factor measuring a student's opinions on group-specific modeling outcomes impacting the data that is permissible. F3 measures if there is a level of support that justifies the usage of data in general. The arrows pointing up to each of the survey items represent the item variances.

if there is any increase in accuracy that justifies the usage of either demographic data, past grades, or political affiliation information. Considering each of these items were yes/no questions, we conducted two-tailed Z-tests with a null hypothesis of 0.5 to determine if students significantly agreed that there is or is not any increase in accuracy that justifies using these data.

To examine RQ3, *when students are given a machine learning model that impacts their own scores, do their opinions change on what types of features can be justified*, we asked students to select which types of features are permitted then provided a short learning session about a nonsense world to them. That is, a description of a made-up world that had specific terms for objects in that world that students could be quizzed on (e.g., birds are now called flutes). Afterward, we asked the students to select which types of features should still be permitted with the understanding that they would be graded on their responses to questions from the short learning session based on the types of features they allow. We used a chi-squared test on their responses before and after the learning session to determine if there was a significant change.

Finally, to analyze RQ4, *do individual difference measures between students impact how much data they are willing to give to machine learning algorithms* *meant to better education*, we used the Mann–Whitney U test to compare the relationships between the individual difference measures and participant responses to F2, the intergroup and intragroup responses factor.

4 Results

For the analyses of RQ1, *How much do students think that machine learning systems should be aware of their own information?*, we examined Spearman correlations between the items, factors, and responses to the introductory survey. In agreement with our hypotheses, using the Kruskall–Wallis test, we found three significant relationships. The first was the relationship between F2 and student gender ($H(1) = 6.39$, $p = .012$), indicating that women were more likely to increase the amount of sensitive information allowable for machine learning models in members of other groups. The second significant relationship was between F3 and student gender ($H(1) = 8.13$, $p < .001$), indicating that women were often more consistent than men when given scenarios related to a group they were in and a group they were not a member of. The last significant relationship was between item ai_added_3 (further described in Table 1) and gender ($H(1) = 8.13$, $p < .001$). Thus, there was a relationship between the gender of students and whether they had unique opinions on information allowed in models based on group membership. Similarly, there was also a relationship between gender and a perceived increase in accuracy that justifies using different types of sensitive information in educational machine learning models. We found that women were more likely to say that there is some level of improvement in educational models that would allow more sensitive types of information to be used.

Table 2 presents the results of RQ2, *Do students believe there are performance thresholds in machine learning systems that justify the use of sensitive features of students?* We found that students thought there are improvements in performance within educational models that *do* justify the usage of both demographic and past grades; however, there is not an improvement that justifies the usage of a student's political affiliation. This result opposes our hypothesis since we expected students would not permit demographic information to be used.

Table 2. Results of Z-tests for RQ2. We found each test had significant results with students agreeing that information about academics and demographics could be used if there was a sufficient improvement in model performance. Students also agreed that there is not a case where political affiliation should be used even if it improves model performance.

Item	**Z-score**	**p-value**	**Mean**	**95% CI**
ai_added_1	2.828	.0047	.625	[0.541, 0.709]
ai_added_2	7.424	$< .0001$	.828	[0.763, 0.893]
ai_added_3	-5.480	$< .0005$	.258	[0.182, 0.334]

For RQ3, *When students are given a machine learning model that impacts their own scores, do their opinions change on what types of features can be justified?*, a χ^2 test revealed a significant difference in the answers to the *pre_LS* item before the learning session and the *post_LS* item after the learning session ($\chi^2 = 127.6$, $p < .001$).

For RQ4, *Do individual difference measures between students impact how much data they are willing to give to machine learning algorithms meant to better education?*, we found one significant relationship between F2 and student gender ($\chi^2 = 9.79$, $p = .006$). However, we found no significance between gender and the intergroup and intragroup responses. There were no significant relationships for the other individual difference measures (EPQ responses, ethnicity, age, etc.).

5 Discussion

We examined students' perceptions of machine learning systems in education, with a focus on the ethical considerations of data usage—the ultimate goal of which is to provide insights into how students view the inclusion of certain individual differences and sensitive features in educational AI systems.

For RQ1, we found that students do in fact have nuanced views on which types of information machine learning systems can utilize. We specifically found three factors that were related to different considerations students make when determining if information can be utilized. Thus, students are not in consensus on which data types should be allowed for educational models.

Furthermore, for RQ2, we found that, generally speaking, students did believe there is some level of model improvement that would justify using features we expected might be ethically ambiguous. While students overwhelmingly agreed that using past academic information was permissible, students also had a consensus towards allowing demographic information if the model had sufficient accuracy improvement. This finding opposes our hypothesis, as students feel that there are cases where more sensitive information can still be used. However, students also agreed that there was not a level of improvement that justified using political affiliation information. We speculate this is due to the fact that it is not as clear that a student's political affiliation could be a proxy for educational outcomes.

When students were asked to consider a machine learning system that directly impacted their own scores after a brief learning session, a significant proportion of students revised their opinions on which information is permissible to use (RQ3). This shift suggests that having a personal stake in the system could lead students to prioritize personal accuracy over more abstract ethical concerns (e.g., a "self-serving bias" [11]).

Finally, through RQ4, we found few individual differences that actually played a significant role in shaping student data-sharing preferences. Specifically, we found that student gender was significantly linked to factors measuring student opinions in two scenarios: when models produce outcomes specific to either their own group (ingroup) or another group (outgroup) (F2), and when there is a certain level of justification supporting the use of data (F3). This shows that students who identify as women more often allow machine learning algorithms to use more sensitive information when the algorithm supports those not in their own group (i.e., students who are not women).

The results point to one overall finding: students have diverse opinions on which types of information educational machine learning systems should have access to; however, they agree that specific cases can permit more sensitive information to be used. Factor analysis showed that students do not have widespread agreement on which features should generally be used in educational machine learning systems. Despite this, RQ2 and RQ3 both illustrate cases in which students do agree on certain types of features being permissible in certain situations or if real-world stakes are on the line. Finally, RQ4 describes that gender might be correlated to the amount of sensitive information that a student is willing to provide to educational machine learning systems.

These results suggest novel considerations that should be taken by both educators and developers of educational systems. We find that students disagree systematically on which types of data are permissible. Thus, there is no "right" solution when taking student perspectives into consideration. Furthermore, it is important to consider student perspectives on data permissibleness in context, given that the stakes make a difference in their responses.

5.1 Limitations and Future Work

Given that this study focuses on student perspectives, it is important to acknowledge the limitations due to the student population we used for data collection and the survey design. The results have some sampling bias since the student sample we studied consists of only college students in the United States and only those on Prolific. Thus, our findings might not be generalizable to younger students or students in different geographical regions. Also, the nonsense learning session we devised to ensure novelty might not be as persuasive to students as a more serious learning topic, where the stakes (and effects) might be even larger.

We situate the paper as a part of a growing movement of treating students as stakeholders in the design of educational machine learning and AI systems [24,27, 40]. From our study, there is future work to better understand if other individual differences and demographics (e.g., socioeconomic status or disability) correlate with types of data that are considered permissible. Furthermore, more research can be done to understand if other hierarchical groupings of students have unique opinions on data usage in educational systems. That is, do students in specific classes, classrooms, schools, or regions have unique data usage opinions?

This study was also focused on asking students questions regarding data usage within hypothetical educational models. To further empower students in educational system design, future research can be done on specific, widely used, educational systems [37]. to provide specific insights to the students who use those systems. Furthermore, longitudinal studies could be designed to assess if students' opinions evolve with their familiarity and experience using different systems.

6 Conclusion

In our paper, we explored students' perceptions of data usage within AI systems in education. We found that students hold differing views on the types of data that should be shared in educational AI systems. While there is no consensus on the data types that can generally be used, students do agree that the context in which the educational AI systems are applied matters when considering the types of data that can be shared.

The results describe the importance of balancing the ethical concerns of AI, involving students and stakeholders, and the practical benefits of AI-powered education. Students' willingness to permit the usage of sensitive data when there are significant improvements in educational outcomes underscores the importance of context when taking student perspectives into account. Furthermore, our study identified gender as the only factor we observed that was related to data-sharing perspectives, rather than political affiliation or ethical ideology as we had expected in the preregistered hypotheses.

The variation in opinions among students suggests that there is no one-size-fits-all approach to data usage in educational AI systems, and thus, educators and researchers ought to take context into account when examining the diverse perspectives of students. Future research should continue to explore how differing contexts and individual difference measures of students influence opinions on data permissiveness in educational AI. Through a better understanding of student perspectives, we can begin to include them in the implementation of educational AI systems to support more ethically aligned and socially responsible technologies that uphold the values of students.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Aruleba, K., Dada, O.A., Mienye, I.D., Obaido, G.: Demography of Machine Learning Education Within the K12. In: Abraham, A., Madureira, A.M., Kaklauskas, A., Gandhi, N., Bajaj, A., Muda, A.K., Kriksciuniene, D., Ferreira, J.C. (eds.) Innovations in Bio-Inspired Computing and Applications. pp. 467–474. Springer International Publishing, Cham (2022)
2. Baker, R.S., Esbenshade, L., Vitale, J., Karumbaiah, S.: Using Demographic Data as Predictor Variables: a Questionable Choice. Journal of Educational Data Mining **15**(2), 22–52 (2023). https://doi.org/10.5281/zenodo.7702628
3. Baker, R.S., Hawn, A.: Algorithmic Bias in Education. International Journal of Artificial Intelligence in Education **32**(4), 1052–1092 (2022). https://doi.org/10.1007/s40593-021-00285-9
4. Belitz, C., Lee, H., Nasiar, N., Fancsali, S.E., Ritter, S., Almoubayyed, H., Baker, R.S., Ocumpaugh, J., Bosch, N.: Hierarchical Dependencies in Classroom Settings Influence Algorithmic Bias Metrics. In: Proceedings of the 14th Learning Analytics and Knowledge Conference. pp. 210–218. ACM, Kyoto Japan (2024). https://doi.org/10.1145/3636555.3636869

5. C. Romero, S. Ventura: Educational Data Mining: A Review of the State of the Art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **40**(6), 601–618 (2010). https://doi.org/10.1109/TSMCC.2010.2053532
6. Chen, H.T., Chen, W.: Couldn't or Wouldn't? The Influence of Privacy Concerns and Self-Efficacy in Privacy Management on Privacy Protection. Cyberpsychology, Behavior, and Social Networking **18**(1), 13–19 (2015). https://doi.org/10.1089/cyber.2014.0456
7. Cock, J.M., Marras, M., Giang, C., KÃ¤ser, T.: Generalisable methods for early prediction in interactive simulations for education. In: Mitrovic, A., Bosch, N. (eds.) Proceedings of the 15th International Conference on Educational Data Mining. pp. 183–194. International Educational Data Mining Society, Durham, United Kingdom (2022). https://doi.org/10.5281/zenodo.6852968
8. Eguchi, A., Okada, H., Muto, Y.: Contextualizing AI Education for K-12 Students to Enhance Their Learning of AI Literacy Through Culturally Responsive Approaches. KI - Künstliche Intelligenz **35**(2), 153–161 (2021). https://doi.org/10.1007/s13218-021-00737-3
9. Felix, C.V.: The Role of the Teacher and AI in Education. In: Sengupta, E., Blessinger, P., Makhanya, M.S. (eds.) International Perspectives on the Role of Technology in Humanizing Higher Education, Innovations in Higher Education Teaching and Learning, vol. 33, pp. 33–48. Emerald Publishing Limited (2020). https://doi.org/10.1108/S2055-364120200000033003
10. Forsyth, D.R.: A taxonomy of ethical ideologies. Journal of Personality and Social Psychology **39**(1), 175–184 (1980). https://doi.org/10.1037/0022-3514.39.1.175
11. Forsyth, Donelson R.: International Encyclopedia of the Social Sciences (2nd edition). Reference Reviews **22**(7), 17–19 (2008). https://doi.org/10.1108/09504120810905060
12. Goetzen, A., Dooley, S., Redmiles, E.M.: Ctrl-Shift: How Privacy Sentiment Changed from 2019 to 2021. Proceedings on Privacy Enhancing Technologies **2022**(4), 457–485 (2022). https://doi.org/10.56553/popets-2022-0118
13. Hu, Q., Rangwala, H.: Towards fair educational data mining: A case study on detecting at-risk students. In: Proceedings of the 15th International Conference on Educational Data Mining. pp. 431–437. International Educational Data Mining Society (2020)
14. Hur, P., Lee, H., Bhat, S., Bosch, N.: Using Machine Learning Explainability Methods to Personalize Interventions for Students. In: Proceedings of the 15th International Conference on Educational Data Mining. pp. 438–445. International Educational Data Mining Society, Durham, United Kingdom (2022)
15. Hutt, S., Mills, C., Bosch, N., Krasich, K., Brockmole, J., D'Mello, S.: "Out of the Fr-Eye-ing Pan": Towards Gaze-Based Models of Attention during Learning with Technology in the Classroom. In: Proceedings of the 25th Conference on User Modeling, Adaptation and Personalization. pp. 94–103. ACM, Bratislava Slovakia (2017). https://doi.org/10.1145/3079628.3079669
16. Hysa, X., D'Arco, M., Kostaqi, J.: Misuse of Personal Data: Exploring the Privacy Paradox in the Age of Big Data Analytics. In: Visvizi, A., Troisi, O., Grimaldi, M. (eds.) Big Data and Decision-Making: Applications and Uses in the Public and Private Sector, pp. 43–57. Emerald Publishing Limited (2023). https://doi.org/10.1108/978-1-80382-551-920231004
17. Jiang, L., Belitz, C., Bosch, N.: Synthetic Dataset Generation for Fairer Unfairness Research. In: Proceedings of the 14th Learning Analytics and Knowledge Conference. pp. 200–209. ACM, Kyoto Japan (2024). https://doi.org/10.1145/3636555.3636868

18. Jones, K.M.L.: Learning analytics and higher education: a proposed model for establishing informed consent mechanisms to promote student privacy and autonomy. International Journal of Educational Technology in Higher Education **16**(1), 24 (2019). https://doi.org/10.1186/s41239-019-0155-0
19. Kim, J.: Leading teachers' perspective on teacher-AI collaboration in education. Education and Information Technologies **29**(7), 8693–8724 (2024). https://doi.org/10.1007/s10639-023-12109-5
20. Koedinger, K.R., Stamper, J.C., McLaughlin, E.A., Nixon, T.: Using Data-Driven Discovery of Better Student Models to Improve Student Learning. In: Lane, H.C., Yacef, K., Mostow, J., Pavlik, P. (eds.) Artificial Intelligence in Education. pp. 421–430. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
21. Koprinska, I., Stretton, J., Yacef, K.: Predicting Student Performance from Multiple Data Sources. In: Conati, C., Heffernan, N., Mitrovic, A., Verdejo, M.F. (eds.) Artificial Intelligence in Education. pp. 678–681. Springer International Publishing, Cham (2015)
22. Korir, M., Slade, S., Holmes, W., Héliot, Y., Rienties, B.: Investigating the dimensions of students' privacy concern in the collection, use and sharing of data for learning analytics. Computers in Human Behavior Reports **9**, 100262 (2023). https://doi.org/10.1016/j.chbr.2022.100262
23. Leventhal, G.S.: What Should Be Done with Equity Theory? In: Gergen, K.J., Greenberg, M.S., Willis, R.H. (eds.) Social Exchange: Advances in Theory and Research, pp. 27–55. Springer US, Boston, MA (1980). https://doi.org/10.1007/978-1-4613-3087-5_2
24. Li, W., Sun, K., Schaub, F., Brooks, C.: Disparities in Students' Propensity to Consent to Learning Analytics. International Journal of Artificial Intelligence in Education **32**(3), 564–608 (2022). https://doi.org/10.1007/s40593-021-00254-2
25. Martin, F., Chen, Y., Moore, R.L., Westine, C.D.: Systematic review of adaptive learning research designs, context, strategies, and technologies from 2009 to 2018. Educational Technology Research and Development **68**(4), 1903–1929 (2020). https://doi.org/10.1007/s11423-020-09793-2
26. Mutimukwe, C., Viberg, O., Oberg, L.M., Cerratto-Pargman, T.: Students' privacy concerns in learning analytics: Model development. British Journal of Educational Technology **53**(4), 932–951 (2022). https://doi.org/10.1111/bjet.13234
27. Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., Käser, T.: The critical role of trust in adopting AI-powered educational technology for learning: An instrument for measuring student perceptions. Computers and Education: Artificial Intelligence **8**, 100368 (2025). https://doi.org/10.1016/j.caeai.2025.100368
28. Nguyen, T.M.: Learning approaches, demographic factors to predict academic outcomes. International Journal of Educational Management **30**(5), 653–667 (2016). https://doi.org/10.1108/IJEM-06-2014-0085
29. Pardo, A., Siemens, G.: Ethical and privacy principles for learning analytics. British Journal of Educational Technology **45**(3), 438–450 (2014). https://doi.org/10.1111/bjet.12152
30. Peer, E., Rothschild, D., Gordon, A., Evernden, Z., Damer, E.: Data quality of platforms and panels for online behavioral research. Behavior Research Methods **54**(4), 1643–1662 (2022). https://doi.org/10.3758/s13428-021-01694-3
31. Rubel, A., Jones, K.M.L.: Student privacy in learning analytics: An information ethics perspective. The Information Society **32**(2), 143–159 (2016). https://doi.org/10.1080/01972243.2016.1130502

32. Sanusi, I.T., Agbo, F.J., Dada, O.A., Yunusa, A.A., Aruleba, K.D., Obaido, G., Olawumi, O., Oyelere, S.S., Centre for Multidisciplinary Research and Innovation (CEMRI): Stakeholders' insights on artificial intelligence education: Perspectives of teachers, students, and policymakers. Computers and Education Open **7**, 100212 (2024). https://doi.org/10.1016/j.caeo.2024.100212
33. Sha, L., Fincham, E., Yan, L., Li, T., Gašević, D., Gal, K., Chen, G.: The Road Not Taken: Preempting Dropout in MOOCs. In: Wang, N., Rebolledo-Mendez, G., Matsuda, N., Santos, O.C., Dimitrova, V. (eds.) Artificial Intelligence in Education. pp. 164–175. Springer Nature Switzerland, Cham (2023)
34. Shaw, T., Cross, D., Thomas, L.T., Zubrick, S.R.: Bias in student survey findings from active parental consent procedures. British Educational Research Journal **41**(2), 229–243 (2015). https://doi.org/10.1002/berj.3137
35. Siemens, L., Althaus, C., Stange, C.: Balancing Students' Privacy Concerns While Increasing Student Engagement in E-learning Environments. In: Wankel, C., Blessinger, P. (eds.) Increasing Student Engagement and Retention in e-learning Environments: Web 2.0 and Blended Learning Technologies, Cutting-Edge Technologies in Higher Education, vol. 6 Part G, pp. 339–357. Emerald Group Publishing Limited (2013). https://doi.org/10.1108/S2044-9968(2013)000006G014
36. Slade, S., Prinsloo, P.: Learning Analytics: Ethical Issues and Dilemmas. American Behavioral Scientist **57**(10), 1510–1529 (2013). https://doi.org/10.1177/0002764213479366
37. Steenbergen-Hu, S., Cooper, H.: A Meta-Analysis of the Effectiveness of Intelligent Tutoring Systems on College Students' Academic Learning. Journal of Educational Psychology **106**, 331 (2013). https://doi.org/10.1037/a0034752
38. Stinar, F., Xiong, Z., Bosch, N.: An approach to improve k-anonymization practices in educational data mining. Journal of Educational Data Mining **16**(1), 61–83 (2024). https://doi.org/10.5281/zenodo.11056083
39. Tajfel, H., Turner, J.C.: The Social Identity Theory of Intergroup Behavior. Political psychology: Key readings., Psychology Press, New York, NY, US (2004). https://doi.org/10.4324/9780203505984-16
40. West, D., Luzeckyj, A., Searle, B., Toohey, D., Vanderlelie, J., Bell, K.R.: Perspectives from the stakeholder: Students' views regarding learning analytics and data collection. Australasian Journal of Educational Technology **36**(6), 72–88 (2020). https://doi.org/10.14742/ajet.5957
41. Wolfenden, L., Kypri, K., Freund, M., Hodder, R.: Obtaining active parental consent for school-based research: a guide for researchers. Australian and New Zealand Journal of Public Health **33**(3), 270–275 (2009). https://doi.org/10.1111/j.1753-6405.2009.00387.x
42. Yacobson, E., Fuhrman, O., Hershkovitz, S., Alexandron, G.: De-identification is Insufficient to Protect Student Privacy, or What Can a Field Trip Reveal? Journal of Learning Analytics **8**(2), 83–92 (2021). https://doi.org/10.18608/jla.2021.7353
43. Yanagiura, T., Yano, S., Kihira, M., Okada, Y.: Examining Algorithmic Fairness for First- Term College Grade Prediction Models Relying on Prematriculation Data. Journal of Educational Data Mining **15**(3), 1–25 (2023). https://doi.org/10.5281/zenodo.10117682
44. Yau, K.W., CHAI, C.S., Chiu, T.K.F., Meng, H., King, I., Yam, Y.: A phenomenographic approach on teacher conceptions of teaching Artificial Intelligence (AI) in K-12 schools. Education and Information Technologies **28**(1), 1041–1064 (2023). https://doi.org/10.1007/s10639-022-11161-x