Personalization over Privacy? Implications of the Privacy-Personalization Trade-Off on Future Use of Intelligent Tutoring Systems

Adrienn Toth(✉), Linda Fanconi, Noé Zufferey, and Verena Zimmermann

ETH Zurich, Zurich, Switzerland
{adrienn.toth,linda.fanconi,noe.zufferey,verena.zimmermann}@gess.ethz.ch

Abstract. The low adoption of Intelligent Tutoring Systems (ITS) and other AIED tools is partly due to privacy concerns when handling student data. Previous research investigated the privacy-personalization trade-off, where users contrast the envisioned benefits against the amount of personal data required. However, related AIED research is still in its initial stages. We analyzed how usability and privacy-related factors influence university students' decision to adopt one of three versions of an ITS, differing in the amount of data collected and—as a result—the availability and depth of educational features. We found that even though students perceived differences in privacy risk across the versions, this did not influence intention to use. Conversely, perceived usefulness of the ITS and its personalized features proved to be a strong predictor of future adoption. We deepen these findings with qualitative results and conclude with recommendations for AIED researchers and developers.

Keywords: Privacy · Usability · Privacy-Personalization Trade-Off

1 Introduction

Intelligent Tutoring Systems (ITS) have been proven to have a positive effect on learning outcomes [14], through personalized features such as immediate feedback, scaffolding hints, and evaluation of student skills over time [16]. However, adoption by higher education institutions has been slow and the need to use rich student interaction data raises privacy concerns [23]. These concerns differ greatly between students and contexts [22], and students have varying expectations toward AI in education (AIED) systems and the related amount of data collection [23]. Also, students need to be taught AI literacy [27] for them to exercise agency over their data responsibly [10]. To that end, we present and evaluate a prototype of ITS that provides students with varying educational

E. G. Blanchard et al. (eds.), AIED 2026, LNCS 16586, pp. 327–336, 2027.
https://doi.org/10.1007/978-3-032-29773-0_24

features depending on the amount of data collected. The system also transparently informs users about what data is collected and the features that have been disabled due to limited access to certain data types. Through this feedback, the system aims to enhance transparency and increase users' AI and privacy literacy.

Previous studies based on the Technology Acceptance Model (TAM) [6] showed the importance of perceived usefulness and perceived ease of use as well as risk perception [7] of a system as predictors for future adoption. Specifically, the privacy-personalization trade-off [1] describes how the benefits one could gain from sharing their data can conflict with the related perceived privacy risks. In the field of AIED, studies analyzing student privacy concerns have mainly relied on self-reports from surveys and interviews instead of experimental methods. We aim to fill this gap by exposing students to a first-hand experience with three versions of an ITS prototype differing in the amount of data collected and the availability of certain educational features. Our goal is to understand how students evaluate the usability and the privacy risk of these different versions, as well as which of these factors have an influence on their intention to use each version. Therefore, we defined the following research question: *How do students evaluate the usefulness, ease of use, and perceived privacy risk of different ITS versions requiring more or less user-data collection, and how do these factors affect their intention to use these versions?*

To answer this question, we use Structural Equation Modeling (SEM) to model the assumed relationships. To investigate potential personal factors behind their perceptions, we collected data on personal traits such as technological affinity [8] and general online privacy concerns [15]. Finally, we performed a qualitative analysis of open-ended responses to deepen our understanding of how the privacy-personalization trade-off influences student adoption of AIED tools.

2 Background: Privacy in AIED Tools

Ethical and societal aspects of AI, such as privacy, are part of most AI literacy frameworks, but usually portrayed as the last step after acquiring technical and algorithmic knowledge [17]. University students typically expect to exercise agency over data shared with AIED tools [22], however, they often lack the required understanding of responsible personal data management [10,27]. Therefore, we aim to enhance an ITS with design features that make data privacy implications clear through short explanations and enabling or disabling certain features depending on students' willingness to share their data.

Prior work showed that privacy-related decisions not only differ between users, but also depend on contextual factors, such as recipient or data type [18]. In AIED, students also have differing views about the data they are willing to share [22], and how they evaluate the privacy of certain features, which in turn is a significant predictor of their adoption [21]. Both students and faculty consider privacy as important as the analytics aspect of learning analytics [24]. Yet, even though students express a strong desire for data privacy, they do not always act accordingly [24]. This is partly due to the privacy-personalization trade-off [9], referring to the fact that enjoying the benefits of personalization usually means sacrificing part of one's privacy. Prior work explored this trade-off for AIED based on survey and interview data, therefore not exposing participants to an authentic experience where their attitudes regarding data sharing could be observed in a specific context. In our study, users directly interacted with different variants of a prototype that allows for capturing their interaction experience.

The widely used Technology Acceptance Model (TAM) [6] contains two main predictors of future intention to use a system: perceived usefulness and perceived ease of use. Usability studies have expanded the model with the construct of risk perception and found significant relationships between perceived risk and intention to use [7]. Differences in affinity for using new technologies are also considered an important factor in predicting user experience and future use [8], which applies to the educational field [3]. In our study, we thus use and expand the TAM model with the following constructs: perceived privacy risk, general online privacy concerns, and technological affinity.

3 Methods

We conducted in-person lab sessions, with $N = 126$ university students, during which they were asked to use different versions of an ITS prototype. Supplementary materials (i.e., screenshots, codebooks, participant demographics, and detailed results) can be found online. Our study has been approved by the Ethics Committee of our institution, and preregistered.

The ITS Prototype. The prototype used for the study was based on CTAT - short for Cognitve Tutor Authoring Tools [25]. We created three versions with adaptations in data collection and the availability of educational features:

- **Tutor A:** This version only recorded essential data from students that let the system operate, such as correctness of answers. Also, students could request a hint for a task, but no additional educational features were enabled.
- **Tutor B:** Tutor B additionally collected and saved data about the development of student's skills, such as *identifying trends*. The hint feature was more customizable, allowing the student to select the preferred type of hint.
- **Tutor C:** Tutor C further contained a chatbot with access to the student's previous interaction to provide personalized support. The chatbot was prompted to provide relevant guidance instead of direct solutions.

For students to understand how the educational features and privacy implications of each version differ, we adopted the same screen structure for all versions. Disabled features (e.g., disabled LLM chatbot in Tutor A) were covered by a short text explaining the reason for disabling it. For our study, we used educational tasks related to visualization literacy. The questions followed the structure of the Visualization Literacy Assessment Test [13]. Each task consisted of one graph along with three to four related questions (multiple choice or short-text answer) that tested students' ability to interpret the graph based on the three levels of graph comprehension by Curcio [5], namely *read the data, read between the data*, and *read beyond the data*, followed by an open-ended question asking them to describe the graph's main insights in 50–100 words. An example task across the three versions is shown in Fig. 1.

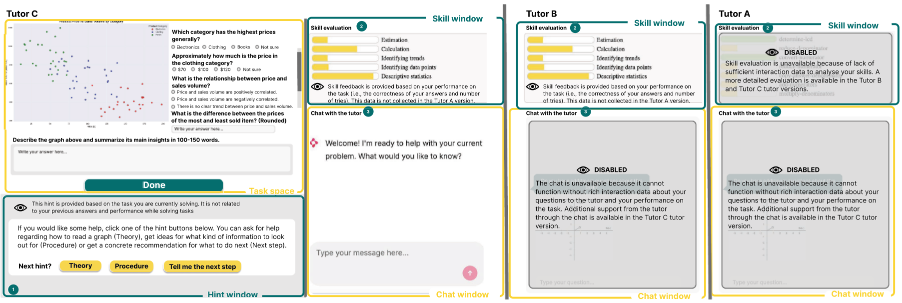


Fig. 1. An illustration of the three tutor versions. Features on the right depend on the available student data. Hint window (1): The students received pre-written hints that help them interpret the graph. Skill feedback (2): Disabled in Tutor A. In Tutors B and C, it provides feedback based on task performance. Chat window (3): Disabled in Tutors A and B. Tutor C uses GenAI to provide additional support. The privacy implications are explained in dedicated information boxes (marked with an eye sign).

Recruitment. The sample consisted of $N = 126$ bachelor students enrolled at ETH Zurich or University of Zurich. The minimum sample size was determined based on a simulation with estimated effect sizes between .15-.35 for all paths, alpha error probability of .05 and power of .8. Among participants, 75 were men, 49 women, one non-binary participant, and one preferred not to disclose their gender. The average age was 21.3 (SD = 2.8), ranging from 18 to 35 years. Every participant was compensated CHF 25 per hour.

Procedure. The experiment took place in person at our institution's lab. First, students (1) interacted with one of the tutor versions by solving a visualization literacy task, directly followed by (2) a survey asking them about their experience with that version using validated questionnaires and open-ended questions. We measured the following constructs: Perceived Usefulness (PU) [6], Perceived Ease of Use (PEoU) [6], Perceived Privacy Risk (PPR) [2], and Intention to Use (ItU) [26]. Additionally, participants answered three open-ended questions, namely (Q1) what they liked or disliked about the version at hand, (Q2) the types of data they think was collected about them, and (Q3) whether and in what context they think they would use that ITS version in the future. Steps (1) and (2) were repeated for each of the three tutor versions. Finally, (3) participants completed a final survey with additional questionnaires related to online privacy concerns (IUIPC) [15], technological affinity (ATI) [8], and provided demographic information. Students were assigned to one of two groups: interacting with the versions in "increasing" (Version A→B→C) (N = 62) or "decreasing". (Version C→B→A) order (n = 64).

3.1 Analysis

We used Multilevel Structural Equation Modeling (MLSEM), where Level 1 (within-subjects) represents the interactions with different tutor versions and Level 2 (between-subjects) represents the individual students. Level 1 variables are therefore those measured by the three version-related surveys, and Level 2 variables were those measured in the exit survey. Figure 2 summarizes the path analysis of the MLSEM model on both the within- and between-subjects level.

Data preparation. Univariate normality was assessed using skewness and kurtosis; all items fell within acceptable ranges (skewness ±2, kurtosis ±7; [11]) except for one item (awa1 of the IUIPC). This item was not removed from the dataset due to being part of a validated scale. To evaluate the necessity of a multilevel approach, we calculated the Intraclass Correlation Coefficients (ICCs) for all Level 1 items. The ICCs ranged from .22 to .50, indicating that a relevant proportion of the total variance (22%50%) was at the between-subjects level. These values exceed the common threshold of .05 or .10 for multilevel modeling [12], confirming that MLSEM was required to account for the nested structure of the data. We additionally performed Multilevel Confirmatory Factor Analysis (MCFA) to verify the construct validity of our measures.

Model Specification. The outcome variable in our model was ItU as our goal was to understand how (perceived) usefulness and privacy concerns influence students' decision to use an ITS. Based on our hypotheses, PU, PEoU, and PPR were modeled as predictors of ItU. Based on previous research [28], we also modeled PU as a mediator between PEoU and ItU. In terms of the between-subjects variables, we expected the effect of ATI on ItU to be mediated by PU and PEoU. Similarly, the effect of IUIPC was expected to be mediated by PRR. We used the R lavaan package [19] to perform our analysis. Because one item exceeded the skewness threshold, we used the Robust Maximum Likelihood (MLR) estimator in our MLSEM, which is robust to non-normality.

Qualitative Analysis. We analyzed answers to open-ended questions using an inductive thematic analysis approach [4]. The qualitative answers were analyzed by two of the authors. For Q1 and Q3 (i.e., what they liked/disliked about the tool, and if they intended to use it), the authors started coding 10% of the dataset. The two created codebooks were then compared, and the researchers agreed on the final code structure. Finally, the researchers independently coded the full dataset, after which interrater reliability was calculated. Remaining disagreements were solved through discussion.

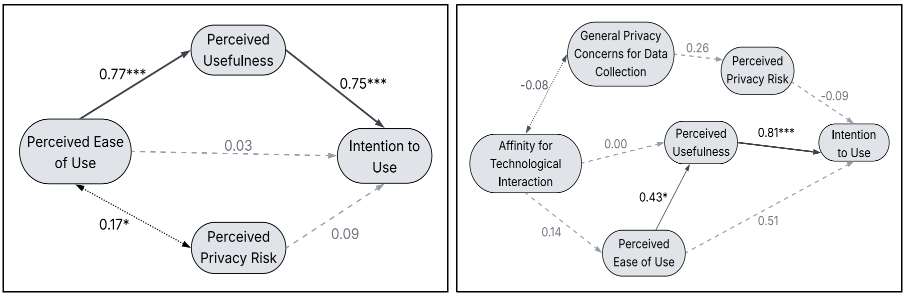
4 Results

Factor Analysis. Construct validity was assessed via a two-level MCFA. On Level 1 (within-subjects level), all items exhibited significant loadings on their respective constructs ($p < .001$). On Level 2 (between-subjects level), the measurement model for individual traits (ATI and IUIPC) was similarly supported. The factor loadings of IUIPC show that only the *Collection* subscale (items 710) achieved acceptable factor loadings (¿.7). We thus decided to only include the four *Collection* items in the final model.

Model Fit. The MLSEM yielded a good fit to the data (Robust CFI = .944; Robust TLI = .939; Robust RMSEA = .035). The SRMR was .093 at the within-subjects level and .171 at the between-subjects level.

Within-Subjects Paths (Level 1). At the within-subjects level, which captures different perceptions across tutor versions, PU was a strong and significant predictor of ItU ($\beta = .75, p < .001$). Whereas the direct path from $PEoU$ to ItU was not significant ($\beta = .03, p = .738$), $PEoU$ demonstrated a significant influence on ItU indirectly by significantly predicting PU ($\beta = .77, p < .001$). The path from PPR to ItU was not significant ($p > .05$), suggesting that perceptions of privacy risk do not influence intention to use different ITS versions.

Between-Subjects Paths (Level 2). At the between-subjects level, which captures differences between individual students, the pattern for ItU remained consistent: PU strongly predicted ItU ($\beta = .81, p < .001$), accounting for 72.6% of the variance in ItU ($R^2 = .726$). Similar to the within-level results, trait-level $PEoU$ significantly predicted trait-level PU ($\beta = .43, p = .020$). Also in line with the within-subjects model, PPR did not prove to be a significant predictor of ItU on the trait level either. Interestingly, the direct paths of technological affinity (ATI) and data collection concerns (IUIPC Collection) did not reach statistical significance ($p > .05$), suggesting that the immediate perceptions of the tutor (PU and $PEoU$) are more important drivers of adoption than personal traits.



$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Fig. 2. Results of the MLSEM model on the within- and between-subjects level.

Exploratory Analysis. To investigate (1) how *PPR* differed across tutor versions, and (2) whether the second-order groups demonstrated any unintended significant differences when rating Level 1 items we conducted ANOVAs with *Version* and *Order* as predictors. For *PU*, *PEoU*, and *ItU*, only the effect of *Version* proved to be significant, while in the case of *PPR*, both the effects of *Version* and *Order* were significant. The significant effect of *Order* was due to the group starting with Tutor A demonstrating generally higher privacy concerns towards all versions (see Fig. 1 in Supplementary Material). Our analysis also shows that participants of both groups clearly perceived an increasing privacy risk from Tutor A to C as intended with the prototype design.

Qualitative Results. All key results are summarized in detail in the online materials. The valence of the responses provided for Q1, i.e., what students liked/disliked, was slightly more often positive (n = 109) than negative (n = 95), and often mixed (n = 127) due to the dual nature of the question. Another n = 47 were neutral or unclear responses. Tutor A received 31 positive responses, Tutor B received 38, and Tutor C received 40, which aligns with the questionnaire scores. This may be due to the availability of the chatbot, e.g., *"The three tips [...] weren't very helpful, while the LLM agent did provide some useful insight."*. The most commonly mentioned aspects that students liked/disliked were the *educational features* (n = 217) and the *User Interface (UI)* (n = 128), with the former being mentioned more often in a positive light (66 positive and 87 mixed), while the UI was more often associated with negative responses (40 negative and 70 mixed). Other popular categories were the *educational content* (n = 78), *usefulness/functionality* (n = 72) and *understandability* (n = 48).

Regarding Q2 (data types students think have been collected), most responses mentioned *educational data*, such as results and skills (n = 259), followed by interaction data, such as clicks (n = 190), e.g., *"It did not collect personal information, it only collected my understanding and reading of the provided graph."* (about Tutor C). All other data types were mentioned less frequently: personally identifiable data 33 times, contextual data (such as date, time, location) 30 times, and personal characteristics (such as style of writing) 29 times.

Responses to Q3 show that students would be interested in using at least one of the tutor versions for their studies. The most frequently mentioned use case for the tutor was *independent studying* (N = 73), with many (N = 62) mentioning that they would only use it for *specific subjects* (mainly mathematics or statistics) or *exam preparation* (N = 35). A lot of participants who answered negatively did not provide additional explanation, but one common reply indicated low expected *effectiveness* (N = 45) from the tool.

5 Discussion and Conclusion

Our study investigated how university students' intention to use ITS differs based on the privacy-personalization trade-off when presented with multiple ITS versions. We found that even though students are clearly aware of the differences between versions in terms of privacy risk, this does not predict their intention to use a specific version in the future, which aligns with a recent related study [20]. The main aspect influencing students' intended adoption is how useful they perceive a specific version of the tool. We additionally found that perceived ease of use is a predictor of perceived usefulness, which is in line with other findings from usability research [6,28]. Based on our findings, we conclude the following:

Perceived Usefulness is the Key to Motivate Students to Use AIED. Both our model and our qualitative results prove the importance of the tool's effectiveness and the quality of the provided educational features. In contrast, data privacy was rarely mentioned as a deciding factor for future use. It is therefore important to make students aware of the educational benefits - while still designing with privacy and other ethical aspects in mind - when the goal is to motivate future adoption.

Transparent Privacy-Related Feedback and First-Hand Experience with an ITS support students' **awareness of the data collection implications** of using its personalized educational features. Participants clearly associated a higher privacy risk with tutor versions that collected more data. When asking them to list the types of data they think the system has access to, most of them thought of those directly identifiable from the screen, such as educational and interaction data. Whereas this is definitely a first step toward successful user notice, the privacy-related feedback should ideally be complemented with additional AI and digital literacy training.

Finally, the **responsibility to ensure student privacy remains with AIED designers**, even if student agency is supported by a system. We saw that user experience and learning benefits are enough for students to sacrifice their data privacy, even when they associate privacy risks with an ITS. AIED designers should thus minimize the privacy-personalization trade-off.

References

1. Awad, N., Krishnan, M.: The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Q. (2006). https://doi.org/10.2307/25148715
2. Ayalon, O., Toch, E.: Evaluating users' perceptions about a system's privacy: differentiating social and institutional aspects. In: SOUPS 2019. USENIX Association (2019). https://www.usenix.org/conference/soups2019/presentation/ayalon
3. Bahari, M., Arpaci, I., Azmi, N.F.M., Shuib, L.: Predicting the intention to use learning analytics for academic advising in higher education. Sustainability (2023). https://doi.org/10.3390/su152115190
4. Braun, V., Clarke, V.: Reflecting on reflexive thematic analysis. Qual. Res. Sport, Exerc. Health (2019). https://doi.org/10.1080/2159676X.2019.1628806
5. Curcio, F.R.: Comprehension of mathematical relationships expressed in graphs. J. Res. Math. Educ. (1987). https://doi.org/10.2307/749086
6. Davis, F.D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q. (1989). https://doi.org/10.2307/249008
7. Featherman, M.S., Pavlou, P.A.: Predicting e-services adoption: a perceived risk facets perspective. Int. J. Hum Comput Stud. (2003). https://doi.org/10.1016/S1071-5819(03)00111-3

8. Franke, T., Attig, C., Wessel, D.: A personal resource for technology interaction: Development and validation of the affinity for technology interaction (ATI) scale. Int. J. Hum. Comput. Interact. (2019). https://doi.org/10.1080/10447318.2018.1456150
9. Garcia-Rivadulla, S.: Personalization vs. privacy: an inevitable trade-off? IFLA J. (2016). https://doi.org/10.1177/0340035216662890
10. Greenhalgh, S.P., DiGiacomo, D.K., Barriage, S.: Platforms, perceptions, and privacy: ethical implications of student conflation of educational technologies. ILS (2023). https://doi.org/10.1108/ILS-03-2023-0030
11. Hair, J.F., Jr., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis, 7th edn. Pearson Prentice Hall, Upper Saddle River (2010)
12. Hox, J.J., Moerbeek, M., van de Schoot, R.: Multilevel Analysis: Techniques and Applications. 3rd edn. (2017). https://doi.org/10.4324/9781315650982
13. Lee, S., Kim, S.H., Kwon, B.C.: Vlat: development of a visualization literacy assessment test. IEEE Trans. Vis. Comput. Graph. (2017). https://doi.org/10.1109/TVCG.2016.2598920
14. Létourneau, A., Deslandes Martineau, M., Charland, P., Skelling-Desmeules, Y., Lepage, V.: A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. Sci. Learn. (2025). https://doi.org/10.1038/s41539-025-00320-7
15. Malhotra, N.K., Kim, S.S., Agarwal, J.: Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. ISR (2004). https://doi.org/10.1287/isre.1040.0032
16. Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, S.R., Rakhshan, M., Keikha, L., Ghazi Saeedi, M.: Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. Int. Learn. Envir. (2021). https://doi.org/10.1080/10494820.2018.1558257
17. Nahar, N., Howard, D., Akeren, K., Ngele, E., Hollins, P., Prescott, G.: Cultivating ai literacy in higher education students: a four-step conceptual framework. In: AIED. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-98420-4_2
18. Nissenbaum, H.: A Contextual Approach to Privacy Online. Daedalus (2011). https://doi.org/10.1162/DAED_a_00113
19. Rosseel, Y.: lavaan: an R package for structural equation modeling. J. Stat. Softw. (2012). https://doi.org/10.18637/jss.v048.i02
20. Rücker, M.T., Büchting, C., Kosch, T.: Understanding the effect of risk perception on the acceptance and use of large language models among university students. ACM Hum.-Comput. Interact. (2025). https://doi.org/10.1145/3757693
21. Schumacher, C., Ifenthaler, D.: Features students really expect from learning analytics. Comput. Human Behav. (2018). https://doi.org/10.1016/j.chb.2017.06.030
22. Stinar, F., Bosch, N.: Surveying contextualized student data sharing preferences for educational ai. In: AIED 2025. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-98417-4_25
23. Toth, A., Roch, N., Zufferey, N., Zimmermann, V.: "i would share it, but..." exploring ways to optimize the privacy-personalization trade-off in intelligent tutoring systems. In: AIED. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-98462-4_9
24. Tsai, Y.S., Whitelock-Wainwright, A., Gašević, D.: The privacy paradox and its implications for learning analytics. In: Lak 2020. LAK 2020, ACM (2020). https://doi.org/10.1145/3375462.3375536
25. University, C.M.: Cognitive Tutor Authoring Tools (CTAT) (2025). https://www.cmu.edu/simon/open-simon/toolkit/tools/learning-tools/ctat.html

26. Venkatesh, V., Morris, M.G., Davis, G.B., Davis, F.D.: User acceptance of information technology: toward a unified view1. MIS Q. **27**(3) (2003). https://doi.org/10.2307/30036540
27. Vincoli, M., Scholz, N., Nagashima, T.: Multidimensional student agency in learning with ai: a conceptual framework and design implications. In: AIED 2025. Springer, Cham (2025). https://doi.org/10.1007/978-3-031-98459-4_16
28. Xue, L., Mahat, J., Ghazali, N., Shi, X.: Technology acceptance model (TAM) in artificial intelligence in higher education (AIHED): a systematic literature review. J. Inst. Res. South East Asia (2025). https://doi.org/10.1177/21582440251409441