Investigating the dimensions of students' privacy concern in the collection, use, and sharing of data for learning analytics

Journal Item

Investigating the dimensions of students' privacy concern in the collection, use, and sharing of data for learning analytics

Maina Korir, Sharon Slade, Wayne Holmes, Yingfei Héliot, Bart Rienties

PII: S2451-9588(22)00096-3

DOI: https://doi.org/10.1016/j.chbr.2022.100262

Reference: CHBR 100262

To appear in: *Computers in Human Behavior Reports*

Received Date: 1 December 2021

Revised Date: 28 July 2022

Accepted Date: 11 December 2022

Please cite this article as: 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* (2023), doi: https://doi.org/10.1016/j.chbr.2022.100262.

Investigating the dimensions of students' privacy concern in the collection, use, and sharing of data for Learning Analytics

MAINA KORIR (CORRESPONDING AUTHOR)

Institute of Educational Technology, The Open University, Walton Hall, Milton Keynes, MK7 6AA

School of Computer Science and Technology, University of Bedfordshire, Polhill Avenue, Bedford, MK41 9EA

Maina.korir@beds.ac.uk

SHARON SLADE

Earth Trust, Abingdon, Oxfordshire, UK

WAYNE HOLMES

UCL Institute of Education, University College London, London, UK

YINGFEI HÉLIOT

Surrey Business School, University of Surrey, Guildford, Surrey, UK

BART RIENTIES

Institute of Educational Technology, The Open University, Walton Hall, Milton Keynes, MK7 6AA

Keywords: privacy concern, contextual integrity, Learning Analytics

Acknowledgements: This work was supported and funded by the Leverhulme Trust, Open World Learning

Investigating the dimensions of students' privacy concern in the collection, use, and sharing of data for Learning Analytics

FIRST AUTHOR'S NAME, INITIALS, AND LAST NAME

First author's affiliation, an Institution with a very long name

SECOND AUTHOR'S NAME, INITIALS, AND LAST NAME

Second author's affiliation, possibly the same institution

THIRD AUTHOR'S NAME, INITIALS, AND LAST NAME

Third author's affiliation, possibly the same institution

FOURTH AUTHOR'S NAME, INITIALS, AND LAST NAME

Fourth author's affiliation, possibly the same institution

Abstract

The datafication of learning has created vast amounts of digital data which may contribute to enhancing teaching and learning. While researchers have successfully used learning analytics, for instance, to improve student retention and learning design, the topic of privacy in learning analytics from students' perspectives requires further investigation. Specifically, there are mixed results in the literature as to whether students are concerned about privacy in learning analytics. Understanding students' privacy concern, or lack of privacy concern, can contribute to successful implementation of learning analytics applications in higher education institutions. This paper reports on a study carried out to understand whether students are concerned about the collection, use, and sharing of their data for learning analytics, and what contributes to their perspectives. Students in a laboratory session (n = 111) were shown vignettes describing data use in a university and an e-commerce company. The aim was to determine students' concern about their data being collected, used, and shared with third parties, and whether their concern differed between the two contexts. Students' general privacy concerns and behaviours were also examined and compared to their privacy concern specific to learning analytics. We found that students in the study were more comfortable with the collection, use, and sharing of their data in the university context than in the e-commerce context. Furthermore, these students were more concerned about their data being shared with third parties in the e-commerce context than in the university context. Thus, the study findings contribute to deepening our understanding about what raises students' privacy concern in the collection, use and sharing of their data for learning analytics. We discuss the implications of these findings for research on and the practice of *ethical* learning analytics.

1 INTRODUCTION

Higher education institutions (HEIs) are keen to adopt learning analytics (LA) to inform teaching and learning, yet widespread uptake in institutions remains to be seen (Joksimović, Kovanović, & Dawson, 2019; Herodotou, Naydenova, Boroowa, Gilmour, & Rienties, 2020). LA has been applied with a view to influencing issues such as reducing the number of students who do not complete their studies (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014; Herodotou, Naydenova, Boroowa, Gilmour, & Rienties, 2020), identifying activities in learning design that may contribute to students completing and passing a module (Rienties & Toetenel, 2016), as well as providing timely feedback and intervention (Pardo, Jovanovic, Dawson, Gašević, & Mirriahi, 2019). LA is defined as the "measurement, collection, analysis and reporting of data about learners and their contexts for the purposes of understanding and optimising learning and the environments in which it occurs" (Long & Siemens, 2011, p. 34). Since the first international conference was held in 2011, the field of LA has experienced steady growth with journals (e.g., Journal of Learning Analytics), conferences (e.g., Learning Analytics and Knowledge Conference), and communities (e.g., Society for Learning Analytics and the Learning Analytics Community Exchange) disseminating research findings for academia and practice.

Concurrent to the development of LA in higher education, privacy in LA presents an opportunity for further innovation and adoption of LA applications where all stakeholders can be engaged in the development process (Gasevic, Dawson, & Jovanovic, 2016). One way to think about privacy in LA is to view it as "freedom from unauthorized intrusion: the ability of an individual or a group to seclude themselves or the information about them, and thus to express themselves selectively" (Ferguson, Hoel, Scheffel, & Drachsler, 2016, p. 11). Thus, as an example, enabling students to control how they are perceived by the use of their data suggests novel features for LA applications.

There is ongoing research on students' perceptions of privacy in LA. In a survey with 1,647 students in the USA, Vu, Adkins and Henderson (2019) informed their study participants what data is collected, who has access to it, and how it might be used. Students in their study indicated a lack of concern about the use of their data. Findings from a survey with 286 UK-based students (Slade, Prinsloo, & Khalil, 2019) were that students in that study accepted institutional use of their data to benefit their learning. In contrast, 330 students who took part in a laboratory study in Germany (Ifenthaler & Schumacher, 2016) were unwilling to share all the data that can be used for LA. Furthermore, a moderated forum discussion with 35 UK-based students (Slade & Prinsloo, 2014) identified students' concerns about surveillance or tracking. Thus, there are mixed results in the literature as to whether students are concerned about the use of their data for LA.

We use contextual integrity (Nissenbaum, 2010) to guide the study research and interpret the study findings. Contextual integrity is a way of understanding whether privacy is likely to be violated in the use of individuals' information. It posits that concerns about collection and use of data can vary in different contexts. As a result, in this study, we contribute new insights into the dimensions of students' privacy concern in LA by comparing privacy concern in LA and e-commerce. Comparing students' perspectives of the use of their data in these two distinct contexts allowed us to better understand the dimensions of their privacy concern (or lack thereof). Our findings show that students are not concerned about the collection and use of their data for LA when compared to the e-commerce context. However, they express more concern about their data being shared with third parties in the LA context than in the e-commerce context. Thus, these findings suggest the need for researchers, in examining students' perspectives of privacy in LA, to further unpack the concept of privacy to identify the specific dimensions that may and may not be of concern to students.

2 RELATED WORK

Students' privacy concern in Learning Analytics

Our understanding about whether students are concerned about privacy in LA begins with an examination of how aware students are about the use of their data. Here, research shows students reporting that they are unaware about how higher education institutions (HEIs) use their data for LA (Jones, et al., 2020; Sun, Mhaidli, Watel, Brooks, & Schaub, 2019), and in some cases, students report being unable to recall giving consent for their data to be used for LA (Jones, et al., 2020; Tsai, Whitelock-Wainwright, & Gašević, 2020; Falcao, Ferreira, Rodrigues, Diniz, & Gašević, 2019). This knowledge gap might help to explain heightened privacy concern when students are made aware of LA (Jones, et al., 2020).

Once researchers address this knowledge gap in their studies and inform students about the use of their data for LA, research has focused on students' corresponding attitudes and preferences. Research suggests a desire for institutional transparency about the use of student data (Slade & Prinsloo, 2014), and a desire to be informed about the use of their data and to be able to provide their consent (Slade & Prinsloo, 2014; Sun, Mhaidli, Watel, Brooks, & Schaub, 2019). Additionally, there is an interest to be compared to other students anonymously and to be able to access LA dashboards confidentially so as not to reveal the information to other students (Roberts, Howell, & Seaman, 2017). Finally, research has noted students' interest to have control over the use of their data (Slade, Prinsloo, & Khalil, 2019; Sun, Mhaidli, Watel, Brooks, & Schaub, 2019; Tsai, Whitelock-Wainwright, & Gašević, 2020), and control over who has access to their data (Jones, et al., 2020). Some authors have suggested that amending perceptions of control over data use could contribute to higher student acceptance of LA (Ifenthaler & Schumacher, 2016). Some research also suggests that students are comfortable sharing their data with third parties (Tsai, Whitelock-Wainwright, & Gašević, 2020).

There are multiple factors that may influence students' willingness to share data for LA, such as their trust in the HEI, concern about data collection, and comfort with instructors' use of their data (Li, Sun, Schaub, & Brooks, 2021); control over the data, length of time a student has been at a HEI, students' use of the Internet and social media, what they expect to gain from the use of their data (Ifenthaler & Schumacher, 2019); and their acceptance of LA (Ifenthaler & Schumacher, 2016).

High levels of trust in the university have been observed (Slade, Prinsloo, & Khalil, 2019) compared to e-commerce and social media companies (Jones, et al., 2020). As such, students' relative lack of concern about the use of their data for LA (Falcao, Ferreira, Rodrigues, Diniz, & Gašević, 2019; Jones, et al., 2020) seems unsurprising. However, research findings also suggest that students are both positive about LA as potentially beneficial to them and at the same time concerned, for instance, about third parties handling student data, the lack of transparency about the use of data, and the need for students to control the use of their data (Nevaranta, Lempinen, & Kaila, 2020); who has access to their information, the possible loss of student responsibility for their own learning, and that LA invades students' privacy (Roberts, Howell, Seaman, & Gibson, 2016). Additionally, student acceptance of LA seems to be tied to certain conditions being met, for instance, that the data is used for what students deem as legitimate purposes (Tsai, Whitelock-Wainwright, & Gašević, 2020), or they are made aware of what data is collected, who has access to the data, and how it is used (Vu, Adkins, & Henderson, 2019). While students may prefer certain types of LA over others (Arnold & Sclater, 2017), they may be unwilling to share data with their lecturers (Ifenthaler & Schumacher, 2019) or be reluctant to share personal data and data about their use of the virtual

learning resources (Ifenthaler & Schumacher, 2016). These particular types of data are especially relevant for predictive LA (Kuzilek, Hlosta, Herrmannova, Zdrahal, & Wolff, 2015).

Even with students' acceptance of LA, they express interest in remaining responsible for their learning (Falcao, Ferreira, Rodrigues, Diniz, & Gašević, 2019; Knox, 2017), and might question the accuracy or completeness of the data as some learning activities are not included (Knox, 2017; Roberts, Howell, Seaman, & Gibson, 2016). Furthermore, it could be that they think that some elements of learning should remain private, and therefore not be included in LA (Knox, 2017). Students may also express concern about surveillance (Slade & Prinsloo, 2014; Schumacher & Ifenthaler, 2018), a stance noted to be in conflict with their interest in receiving personalised support (Slade & Prinsloo, 2014). They might expect to provide their data in exchange for a service from the HEI (Jones, et al., 2020; Tsai, Whitelock-Wainwright, & Gašević, 2020), however, other research (Slade, Prinsloo, & Khalil, 2019) found only a small difference between those who accepted the exchange of their data for a service and those who did not.

Theoretical Background – Contextual Integrity

This study is underpinned by contextual integrity (Nissenbaum, 2010), which is an approach to understand privacy and identify possible privacy violations in the use of individuals' information. Nissenbaum argues that social life is governed by norms of information flow, that is, what type of information is passed on from one entity to another, and under what conditions. These norms are identified from various sources including culture, law, history, and convention among others. Nissenbaum's work identifies two norms: of appropriateness and of flow. Norms of appropriateness govern what personal information can or cannot be revealed in a given context, for example one might feel free to talk about politics with immediate family and close friends but not colleagues. Norms of flow govern the movement of information from one party to another, for example, one can tell their doctor things about their health status, and not expect this information to be shared with others (apart from other health professionals). Contextual integrity is used to identify when privacy is breached and to understand why this is the case: it would be violated if either the norms of appropriateness or flow are breached. To identify privacy violations, contextual integrity identifies: i) the context, for example, where data is collected and where it is used, ii) the actors involved, namely, senders and recipients of information and the information subjects, iii) the attributes or information types, and iv) the transmission principles guiding the flow of information between different actors.

Researchers have applied contextual integrity to LA, in various ways. Heath (2014) uses it to analyse data use scenarios for learning analytics and identifies where potential privacy violations might arise. It has also been used in empirical literature on privacy in LA primarily to explain the research findings. For example, participants in Tsai, Whitelock-Wainwright and Gašević's study (2020) were found to conceptualise privacy using contextual integrity. To determine the appropriateness of data sharing, study participants considered the data that was to be shared, who was involved (e.g., the tutor and the student) and the type of relationship between the tutor and student. Research findings also indicate that practices of institutional data use are misaligned with students' expectations (Jones, et al., 2020). The authors call for effort to re-align these two to achieve contextual integrity. Ifenthaler and Schumacher's work (2016) also supports contextual integrity in the LA context as students in their study did not want data which had been freely shared in one context (social media) then used in a different context (LA). In the present study, contextual integrity was applied to identify the norms of appropriateness and flow held by students and to better understand the study results.

Research questions

Based on the literature analysed in the previous section, we can surmise that the picture on students' perspectives of privacy in LA is not clear. Despite some research demonstrating students' acceptance of LA and a lack of concern about privacy, other studies suggest a more nuanced landscape. Thus, our study sought to unpack and better understand privacy in LA by focusing on the dimensions of privacy in the collection, use and sharing of data for LA. We decided to compare students' privacy concerns in the LA context to the e-commerce context. This decision was motivated by literature suggesting that students had higher levels of trust in their university than in e-commerce and social media companies (Slade, Prinsloo, & Khalil, 2019; Jones, et al., 2020).

In light of the research summarised in this section, we identified the following research questions for our study: **RQ1:** To what extent are students concerned about the collection, use and sharing of their data for learning analytics, and compared to e-commerce? **RQ2:** To what extent are students' general privacy concerns and behaviours related to their concern over the collection, use and sharing of student data for learning analytics? and **RQ3:** What issues contribute to students' concern or lack of concern over data collection, use and sharing for learning analytics?

3 METHOD

This section describes how the study was carried out with students in a laboratory session. The study received ethical approval from the university's human research ethics committee *[HREC number omitted for review]*.

Setting and participants

This study was conducted at the business school of a UK university. Students were studying a Masters' module in Organisational Behaviour. A total of 143 students were registered for the Organizational Behaviour module and 111 took part in the laboratory session. Of these, the majority were female (n = 90, 81%). Two students did not indicate their gender. The average age was 23.1 (SD = 1.9), and the ages ranged from 21 to 31 years. Three students did not indicate their age.

The GLOBE country cluster system (Mensah and Chen, 2013) was used to categorise students according to their region of origin as there were several countries with only one or two students (see Table A1 in the Appendix). Most students were from the Confucian Asian (73 – 65.7%), Anglo (16 – 14.4%), and Southern Asian (10 – 9.01%) clusters. The large proportion of students with international backgrounds is typical for postgraduate courses in the field of business and management[1].

Study design and procedure

Masters' students studying Organisational Behaviour (n=111) took part in a laboratory session and follow-up semi-structured interviews during which they answered questions about their general privacy behaviour and privacy concerns. In addition, they answered several questions based on two vignettes. The study questions are detailed in Section 3.3. After answering the privacy questions and questions focusing on the vignettes, students participated in group discussions to enhance their learning on the topic being studied as seen in educational research (Rienties & Héliot, 2018) and learning analytics studies (Pijeira-Díaz, Drachsler, Järvelä, & Kirschner, 2016; Knight, et al., 2017). The design of Study 2 is shown in Figure 1.

[1] Data obtained from the UK Higher Education Statistics Agency - https://www.hesa.ac.uk/data-and-analysis/students/what-study

➔ **Insert Figure 1 about here**

Vignettes were used to guide the discussion with students. This section first provides a description of the vignettes followed by the study protocol.

3.1.1 Vignettes

Vignettes depict situations in short story form to which study participants are invited to respond (Finch, 1987). They enable actions to be explored in a given context in a distanced and less personal way (Barter & Renold, 1999). Vignettes have been used extensively, both in privacy and human computer interaction research (Xu & Teo, 2004; Naeini, et al., 2017), and in education research (Rienties & Héliot, 2018).

Students were shown two vignettes to explore whether they were concerned about the collection and use of their data, comparing e-commerce and LA contexts. The first vignette shown to participants was based on Amazon, an American technology company offering its services in many countries around the world. This study focused only on its e-commerce services. The first vignette, which was read out loud to students in the laboratory session, is shown below:

Amazon is an e-commerce company that a number of you might be familiar with. It provides a personalised user experience, suggesting potentially relevant purchases based on your browsing and purchasing history. Please answer the questions that follow about the Amazon vignette.

The second vignette was based on a student-facing learning analytics dashboard (SFLAD). It described a hypothetical situation where student-facing LA was introduced to students at the university. Students were shown screenshots based on [*blinded for peer review*], a predictive LA system which has been adopted on a large scale at [*blinded for peer review*]. While this system [*blinded for peer review*] predicts students at risk of failing or of not submitting the next assessment, this study focused on the student activity recommender feature which recommends resources that students are yet to interact with and that will help them prepare for the next assessment (Kuzilek, Hlosta, Herrmannova, Zdrahal, & Wolff, 2015). The text of the second vignette, which was also read out loud to students in the laboratory session, is shown below:

The University plans to roll out dashboards to help students keep track of their learning progress in individual modules and courses of study. The dashboards will be created using individual student data, data from their peers, and data from students who took the module in the past. Individual student data will include their performance on various assessments, their attendance to the classes, as well as their personal data provided at registration. You will now see a screenshot of the proposed system. Please review the screenshot, imagining you are the student referred to, and answer the following questions.

Both vignettes were designed to be realistic, relevant, and easy for the students to relate to. Amazon was considered familiar to students as it offers incentives that are specifically relevant for students including reductions on book prices as well as free subscription to a next-day delivery service for one year. This assumption was later verified in the study as all students stated that they had an account with Amazon. Similarly, the SFLAD vignette was relevant to the students as it described how they could keep track of their learning progress and receive personalised recommendations for learning resources.

The two vignettes shared similar characteristics as they focused on the provision of personalised services for students. They were of differing lengths, and while the Amazon vignette was realistic, the SFLAD vignette was

hypothetical, since at the time of carrying out this research there were no known plans to unilaterally introduce LA at this particular university.

3.1.2 Study protocol

Students in the study attended a one-hour interactive laboratory session. They were briefed on the study and provided with an information sheet to review before the laboratory session began. The information sheet contained details about the study and informed students about their rights, including that they could withdraw from the study with no negative effect on their course participation or grades.

Students provided their consent to participate in the study. They then filled out the privacy questions (Buchanan, Paine, Joinson, & Reips, 2007) using an online survey tool from JISC. Students then engaged with the Amazon and SFLAD vignettes and were prompted to answer several questions to assess their privacy concerns with various uses of data in each context.

The free version of the PollEV software was used to collect data during the interactive part of the laboratory session. PollEV has been used successfully to improve student engagement in lectures and classrooms (Kappers & Cutler, 2014) and was therefore appropriate to use in the study. As students engaged with the vignettes, their responses to the different questions were displayed as graphs on the screen. Students then had a brief discussion session with their peers to: (i) share their thoughts on collection and use of data as described in the vignettes, (ii) reflect on why they thought the way they did and find out what members of their group thought and why, and (iii) explore whether they and their peers had similar or different personality profiles. The latter two steps were linked to students' learning for the Organisational Behaviour module. Finally, the students were debriefed, and further discussions were held to relate the work carried out in the laboratory session to their learning on personality and organisational data practices for the Organisational Behaviour module (Rienties & Héliot, 2018).

Shortly after the laboratory session, students who took part in the study were sent a personalized privacy profile as feedback. The privacy profile showed students their scores in response to the questions and provided them with additional reading resources on privacy if they wanted more information.

Study instruments

The online privacy concern questionnaire (Buchanan, Paine, Joinson, & Reips, 2007) was used to determine students' privacy behaviour and general privacy concern. The questionnaire is divided into three scales: general caution, technical protection, and privacy concern. It has also been used in numerous studies, see, for example, (Coles-Kemp & Kani-Zabihi, 2010; Woodruff, Pihur, Consolvo, Brandimarte, & Acquisti, 2014; Lee, Wong, Oh, & Chang, 2019).

The questions for the interactive laboratory session were adapted from work by Slade, Khalil, and Prinsloo (2019). These questions focus on students' perspectives of data collection and use for LA. Examples of the questions used in the study are shown in **Error! Reference source not found.**, while all the questions are shown in the Appendix.

➔ **Insert Table 1 about here**

3.1.3 Follow-up interviews

After data from the laboratory session was analysed, 41 students (out of the 50 who had volunteered to participate) were contacted for the follow-up interviews. These 41 students were selected to meet two criteria: first,

that they had responded to most of the study questions, and second, that they represented different privacy segments based on their responses to the privacy index questionnaire. The aim of the follow-up interviews was to gain deeper insights into the motivation for students' individual responses to the questions in the laboratory session. The interview schedule is shown in the Appendix.

It was not possible to interview all students as some did not respond to the invitation or were no longer able to participate due to end of year holidays followed by an examination period. In total, four students were interviewed. The follow up interviews did not aim to obtain a representative sample, instead the focus was on obtaining insights into students' motivations. The findings are discussed in Section 4.4. Given that only four students took part, the insights from the follow-up interviews are preliminary, and point out areas for further investigation with a larger group of participants.

3.1.4 Data analysis

T-tests were used to analyse the data to answer the first research question, comparing participants' concern across the e-commerce and LA contexts. In addition, correlation tests were used to analyse the data to answer the second research question as to how the study variables related to each other.

3.1.5 Missing data

A total of 111 students attended the laboratory session. All 111 students filled out the online privacy questionnaire and provided their demographic data. However, some data was not collected during the interactive laboratory session. Several issues contributed to the missing data. First, the free version of the PollEV software allowed a maximum of 40 participants per session. As the first two groups of students had slightly over 40 participants, some students were unable to take part in the poll. Second, since the laboratory session was scheduled for one hour, there was little opportunity to wait for extended periods of time after each question for all students to respond, and it was challenging to keep track of who was yet to respond to a question. Third, there was data loss during the data download process. Finally, some data was lost as students used different identifiers across the two data collection tools (PollEV and JISC online surveys), and thus their data across the two data sets could not be combined for analysis. Missing values were replaced with the mean value calculated from participants' responses to a question following best practice recommendations (Groves, et al., 2009).

4 RESULTS

Factor analysis

Principal component analysis with direct Oblimin rotation was carried out on the questions focused on collection and use of data in the university context. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .659. Bartlett's test of sphericity was significant (chi-square = 101.713; df = 15 p<.001). Two components were identified, explaining 57% of the variance. The first component had an eigenvalue of 2.28 (corresponding to 38.1% of the variance), the second component had an eigenvalue of 1.17 (corresponding to 19.5% of the variance). The first factor was related to comfort with data use and data sharing and the second factor to comfort with benefits in exchange for tracking. The factors and components are shown in **Error! Reference source not found.**.

➔ **Insert Table 2 about here**

Similarly, factor analysis using principal component analysis with direct Oblimin rotation was carried out on the questions focused on collection and use of data in the Amazon context. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .594. Bartlett's test of sphericity was significant (chi-square = 88.802; df = 15; p<.001). Three components were identified, explaining 71.5% of the variance. The first component had an eigenvalue of 2.06 (corresponding to 34.3% of the variance), the second component had an eigenvalue of 1.16 (corresponding to 19.3% of the variance), while the third component had an eigenvalue of 1.08 (corresponding to 18% of the variance). The first factor was related to students' comfort with benefits for tracking with no data sharing, the second to comfort with benefits in exchange for tracking and with identifiable data sharing, and the third to comfort with data use and anonymised data sharing. The factors and components are shown in **Error! Reference source not found.**.

➔ **Insert Table 3 about here**

The factor analysis of the questions across the two contexts identified different factors. This pointed to the need to further validate the scale and is highlighted as one of the limitations of the work. In presenting the remaining results in this section, only the factors identified from the questions focused on collection and use of data in the university context will be used further.

The extent of students' privacy concern in the collection, use and sharing of data for learning analytics

The first research question was "to what extent are students concerned about the collection and use of their data for learning analytics, and compared to e-commerce?" The mean values from participants' responses to the questions on collection and use of their data in both the Amazon and university scenarios were obtained, as seen in **Error! Reference source not found.**.

➔ **Insert Table 4 about here**

Students seemed to be more comfortable with the university rather than Amazon carrying out the following activities: the collection of their personal data (mean = 3.93), sharing of their personal and online activity data with third parties in an anonymised format (mean = 2.81), and in a personally identifiable way (mean = 2.08), and being offered specific benefits in exchange for being tracked online (mean = 2.71). This could result from the collection of personal data in the educational context being more familiar to them. However, students were observed to be less comfortable with their data being shared with third parties by the university compared to Amazon (mean = 1.93). This result may be because students are unaware who the third parties are and may think that they are influential entities, such as future employers.

In the Amazon context, students were more comfortable with Amazon offering them specific benefits in exchange for being tracked on the condition that their data was not shared with third parties (mean = 3.68) and least comfortable with Amazon sharing their personal and online activity data in a personally identifiable way with third parties (mean = 1.82). There was a small difference between the two contexts with respect to students' comfort with specific benefits in exchange for being tracked online, thus their comfort levels in both contexts were comparable in this instance.

A single scale was derived from participants' responses separately for the Amazon questions (Cronbach's alpha = 0.59; Mean = 2.61; SD = .46) and the university questions (Cronbach's alpha = 0.67; Mean = 2.77; SD = .49). A Shapiro-Wilk's test revealed that the data was not normally distributed (W = .940 Amazon; W = .936 University; p < .000). Therefore, a Wilcoxon's signed-rank test was carried out to determine if there were any differences in the median values between participants' responses to the Amazon and university questions. A statistically significant difference was observed (Z = -3.463; p < .001), suggesting that overall, participants were more comfortable with the collection and use of data in the university context than in the Amazon context. This might be because students have greater trust in the university than Amazon, or that they are more familiar with the practice of the collection and use of their data in the university context.

Furthermore, students indicated that they were comfortable with sharing their data with their tutors so that the tutors could support them better. While 18 students (17.1%) disagreed with their data being shared with their tutors, 60 students (57.2%) indicated that they were comfortable with this practice, while 26 students (24.8%) were neutral. This finding aligns with those from other research (Vu, Adkins, & Henderson, 2019) which suggest that students are comfortable with the collection and use of their data where the recipient and the purposes are known and the use is related to their learning.

The relationship between students' general privacy concern and behavior and privacy concern in the collection, use and sharing of data in learning analytics

The second research question was "to what extent are students' general privacy concerns and behaviour related to their concern about the collection and use of student data for learning analytics?"

'Hiding a bank PIN when using cash machines/making purchases' and 'shredding personal documents when disposing of them' were the most practiced activities on the general caution scale (mean = 4.05 and 3.23, respectively), while 'reading a website's privacy policy' and 'reading licence agreements fully before agreeing to them' were the least practiced activities (mean =2.07 and 1.99, respectively). The most practiced technical protection activities for participants in the study were 'watching for ways to control what one is sent online' and 'using pop-up window blockers' (mean = 3.39 and 3.19, respectively), while the least practiced technical protection activities were 'checking one's computer for spyware' and 'removing cookies' (mean = 2.65 and 2.59, respectively). In both responses, it might be the case that participants did not know what cookies or spyware were or did not know how they could be removed. Finally, with the privacy concern scale, the activities leading to the highest privacy concerns related to 'someone intercepting a credit card while one is buying something on the Internet', or 'one being mischarged when buying something on the Internet using the credit card' (mean = 3.96 and 3.92, respectively), and the activities with the least concern involved 'information about one being found on an old computer' and 'someone gaining access to the student's electronic medical records' (mean = 3.32 and 3.05, respectively). Thus, participants' responses to the general caution, technical protection, and privacy concern scales were as expected.

Shapiro-Wilk's tests showed that the general caution, technical protection, and privacy concern scales were normally distributed while the scales identified from a factor analysis of the university-related questions - comfort with data use and data sharing (Factor 1) and comfort with benefits for tracking (Factor 2) - were not. Therefore, Spearman correlations were used on all the scales. The results alongside mean, standard deviation and normality results for the different study scales are shown in **Error! Reference source not found.**.

➔ **Insert Table 5 about here**

The results suggested that students who carried out technical protection activities also adopted activities related to general caution ($r = .449$; $p < .01$). Those students who had high privacy concerns undertook more general caution activities ($r = .200$; $p < .05$). Finally, those who were comfortable with data use and data sharing for LA were also comfortable receiving benefits in exchange for tracking in the university context ($r = .361$; $p < .01$). Thus, these results suggest that students' general privacy concerns and behaviour are distinct from concern over the collection, use and sharing of data for learning analytics.

Issues contributing to students' (lack of) concern about collection, use and sharing of data for learning analytics

Follow-up semi-structured interviews were carried out with 4 students who took part in the initial study. There were two male and two female students, with an average age of 23.5. A thematic analysis of their responses highlighted three relevant themes which we discuss in this section.

4.1.1 Relationship with the university and corresponding (lack of) trust

Students' relationship with the university influenced how they perceived institutional use of student data. Both Participant 1 and Participant 2 were willing to share data based on their *relationship* with the university. Participant 1 expected that the university would have and therefore would use students' data by virtue of the student-university relationship. In fact, for this student, this seemed to be a foregone conclusion:

> *"And for the [University name] part, I mean, I'm their student. They are supposed to have my data. I don't have a problem with that at all." (Participant 1)*

The relationship between the student and the university was noted to contribute to the student developing trust in the university. Participant 2 was more supportive of personalised services from the university compared to Amazon:

> *"I think it's because [University name] is something that is really close to me right now. Amazon, I'm not. So, I would want to believe that I can rely on my institute more. And obviously when it comes to my privacy and everything. But Amazon is not something I'm connected to." (Participant 2)*

Participant 3 and 4 also stated that they trusted the university to handle their data appropriately and not to students' detriment. Consequently, Participant 3 stated that they were more comfortable being tracked online by the university than by Amazon. However, the student expressed that there were limits to the influence they expected the university to have:

> *"So, when I think about tracking is, I don't know, maybe on my location, or what I do, what I search on Amazon, so yeah, it's like I'm being watched or something. That's what it comes to mind. So that's why I totally disagree with Amazon... Yes, I agree with the university to offer me something based on what I do. Not of course, to intervene me in my personal life. So yes, again, I trust the university more on that and not obviously the Amazon platform." (Participant 3)*

In contrast, Participant 1 expressed mistrust that the university would handle student data appropriately, given that students rarely read the data use policies where details of data use would ordinarily be provided:

> *"Because sometimes I don't think that they might share data anonymously. I don't think they do that. Even though they might be anonymous, at the end, they might know that this data are for me. So there's no anonymity at the end… I think that's most of the cases when we agree on the terms and conditions, we never read about it. We never do it. So, you don't know what you have signed. And in most cases, the small prints are the ones that they say, we might sell your data anonymously with third parties. So that's why I say that disagree. Because I never do that. I never read the terms and conditions, but I accept them. So I mean, in the backstage, you don't know what really happens. They might tell you something, but it might be otherwise." (Participant 1)*

4.1.2 Data access and control

Participants expressed an interest in having control over third parties' access to their data. Participant 2, for example, wanted control over which third parties could access her data, rather than have the university make this decision. The participant's stance could have been motivated by a lack of information about the third party's identity and how it would use students' data:

> *"…because I would not want that to go to the third party, because I'm, like I said, it's restricted to one particular institute, and I would want it that way. If I really want access to another third party, I will go there. I wouldn't want someone else to give my information there… if it's a third party, I don't know the party. I don't know what my information is going to be used [for]." (Participant 2)*

4.1.3 Benefits and trade-offs

While the potential benefits of the collection, use and sharing of student data for learning analytics were observed to play a role in enabling students' acceptance, students also indicated an awareness of the need to provide their data to access these benefits. Both Participant 3 and Participant 4 referred to the benefits they stood to gain from sharing their data with the university. While Participant 4 anticipated that other interaction would become more convenient as a result, Participant 3 felt more comfortable sharing data with the university as he perceived that it would provide more functional benefits:

> *"I think Amazon could have more data from me, but the only thing that they can do is suggest me things to buy. So, to give more money. While the university can offer me a different kind of service, more quality of my studies, or yeah. So I think if I was to give information to these two platforms, or something, the university could use them more widely to offer me something better. While Amazon wants my money actually. So yeah, so I will be more comfortable to give more information to the university." (Participant 3)*

Finally, students' experience in other situations outside of the learning context may have contributed to the perception that they shared their data in return for some service from the university, as suggested by Participant 2:

"So I think that works for me, because how else will I ever work in an institution? I mean, no, no company or nobody in person is ever going to be able to help me without giving some input like that particular company should know what are the things that I'm looking for? What are the searches that I have? What are the things I'm like I want right now? And if I give that data, only, then they will be able to help me with what I want. … So I think the trade-off would if it benefits me, I will be okay with that trade-off. Because its' not only going to give me what I really want, but it is also going to provide me with benefits." (Participant 2)

5 LIMITATIONS

While fifty students signed up to participate in the follow-up interviews, demonstrating their interest in the issues under investigation, only four students were available to attend, providing preliminary insights from the qualitative data collected for this study. Consequently, future research with more students is needed to identify further insights and to determine if these insights are shared by students depending on their stand on privacy. Future studies can incorporate open-ended questions for participants to respond to during the study to minimize the drop-out rate. While acknowledging this limitation, it is noted that the responses from the four students have enabled relevant and noteworthy insights to be identified.

As students engaged with the vignettes, their responses to the different questions were displayed as graphs on a screen visible to all students in the laboratory. While the PollEV software used to collect data from students displays these responses anonymously (that is, they are not linked on the display to the individual participating students), having aggregated responses on the screen could have influenced students' responses to later questions. To mitigate this in future work, the results can be shown to students after they have answered all questions.

While validated questionnaires (Buchanan, Paine, Joinson, & Reips, 2007) were used to examine students' general privacy concerns and behaviour, the questionnaires used to examine privacy concern in LA, despite being taken from existing research (Slade, Prinsloo, & Khalil, 2019) were observed to load onto multiple factors and different factors across the e-commerce and university contexts. This demonstrates the need for further work to develop validated questionnaires to examine privacy issues in LA, such as the students' expectations of LA questionnaire (Whitelock-Wainwright, Gašević, Tejeiro, Tsai, & Bennett, 2019), or questionnaires used in the work of Ifenthaler and Schumacher (2016).

This study only had input from students from a single university. Furthermore, they were all postgraduate students pursuing a single programme of study. Future work can include students from a variety of higher education institutions and programmes of study in order to gain a broader perspective. In particular, distance learning programmes and universities are more heavily reliant on uses of student data to track their progress. Further studies might also focus on differences in student attitudes between distance learning and campus-based programmes.

6 DISCUSSION AND CONCLUSION

Regarding RQ 1, a key finding of this study was that students in the study were significantly more comfortable with the collection and use of data in the university context rather than the Amazon context. Thus, these findings point to students' lack of concern to share their data for LA. Furthermore, students in the study were less comfortable with the university sharing their data with third parties compared to Amazon. At first glance, this appears a little counterintuitive. Students have suggested higher trust in their university than for external bodies, yet express greater concern in their university's potential data sharing practices. This may be explained by referring

to contextual integrity (Nissenbaum, 2010). For example, it may be that many are already aware of data sharing in a commercial context, but less aware, and potentially therefore, more disturbed by, data sharing in an educational context. Additionally, students might have been concerned that they did not know who would have access to their educational practices and data records. They may have considered that the third parties were potential employers, and as such, would have wanted to know what was shared with them, given the potential to influence their future employment prospects. As these concerns could be due to students' lack of knowledge about the details of LA implementations, greater transparency by universities is recommended, to clarify this information for students. Based on the work of Vu, Adkins and Henderson (2019) this level of transparency can be expected to have a positive impact on students' willingness to share data for LA. Initiatives to enhance institutional transparency to students regarding use of their data for LA should take into account lessons learned from similar initiatives. For example, students might not read university data use policies, therefore other potentially more effective approaches should be identified and implemented. Another consideration is that the point in time when data use information is shared with students is crucial. At the start of the term, for example, students might be focused on completing registration activities and therefore be unable to pay close attention to information on the use of their data. Consequently, universities could explore consent reminders at other times, and give students opportunities to make changes.

Regarding RQ 2, students' general privacy concerns and behaviour were seen as distinct from their concern over collection, use, and sharing of their data for LA. Those students who were comfortable with data use and data sharing for learning analytics were also comfortable receiving benefits in exchange for tracking in the university context. This result suggests that universities emphasise benefits students stand to gain, given the use of their data for learning analytics. However, given ethical considerations about presenting an accurate picture to students, information on potential risks and how these can be mitigated should also be provided. Further work is needed to better understand additional factors (other than privacy concern) that may contribute to students' unwillingness to share their data for learning analytics.

Regarding RQ 3, the qualitative data suggested that the relationship between the student and the university could lead them to trust the university to use student data for students' benefit. Where mistrust was expressed, this was not necessarily due to the university's inaction, rather it was expressed due to students' lack of awareness (for instance, by not engaging with the content of the privacy policies). This emphasises the need for students to take up available opportunities to be made aware about how their data is used. At the same time, as discussed previously, universities need to ensure that this information is given to students in ways that make it accessible to ensure students are truly informed. Finally, students' perceptions of what they stood to gain from sharing their data for learning analytics also seemed to play a role in minimising their concern about sharing their data for learning analytics.

Overall, this study's findings shed further light on the dimensions of privacy and students' specific concerns around the collection, use and sharing of their data for LA. These findings are aligned to contextual integrity (Nissenbaum, 2010) that comfort with data use and data sharing might be influenced by the context in which data is collected and used. The results demonstrated students' comfort with the university using their data for LA compared to Amazon *using* their data. Additionally, students' discomfort with their data being *shared* with third parties can be seen as examples of their norms of appropriateness and flow, that is how they expect their data to be used, and who they expect to have access to their data. These findings emphasise that the context where data is collected and where it is used is an important component in understanding students' data use preferences and what

practices might stand out to them as unusual or unacceptable, and thus what they might perceive as violating their privacy.

Concerns about who has access to students' data, for example, third parties in general and future employers as a specific example highlights an informational norm that there is an expectation that data can or cannot be shared, or that it is shared under certain constraints. This highlights a possible need for opportunities for these informational norms to be shared with or captured by the HEI. Thus, it is important to consider how HEIs can identify these informational norms from students, and how these can be used in the design and development of LA, while considering personnel and other resource constraints that HEIs operate under.

Following on from this work, it is recommended that HEI data use transparency initiatives to include information whether student data is shared with third parties, and what this means. For example, it might be the case that only anonymised data is shared and informing students about this can help ease their concerns, or as is usual, that HEIs share student data only as part of a service agreement, for example, with regard to marketing.

Furthermore, the results from this study emphasise the need to unpack privacy as a concept into specific dimensions for study, in this way bringing greater clarity to research findings on privacy concern in LA.

REFERENCES

Arnold, K. E., & Sclater, N. (2017). Student Perceptions of their Privacy in Leaning Analytics Applications. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, (pp. 66–69).

Barter, C., & Renold, E. (1999). The Use of Vignettes in Qualitative Research. *Social research update, 25*, 1–6.

Buchanan, T., Paine, C., Joinson, A. N., & Reips, U.-D. (2007). Development of Measures of Online Privacy Concern and Protection for Use on the Internet. *Journal of the American Society for Information Science and Technology, 58*, 157-165.

Coles-Kemp, L., & Kani-Zabihi, E. (2010). On-line Privacy and Consent: A Dialogue, Not a Monologue. *Proceedings of the 2010 New Security Paradigms Workshop* (pp. 95–106). ACM. doi:10.1145/1900546.1900560

Cormack, A. N. (2016). A Data Protection Framework for Learning Analytics. *Journal of Learning Analytics, 3*, 91-106.

Falcao, T. P., Ferreira, R., Rodrigues, R. L., Diniz, J., & Gašević, D. (2019). Students' perceptions about learning analytics in a Brazilian higher education institution. *2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT)*, *2161*, pp. 204–206.

Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest editorial: Ethics and privacy in learning analytics. *Journal of learning analytics, 3*, 5–15.

Finch, J. (1987). The Vignette Technique in Survey Research. *Sociology, 21*, 105–114.

Gasevic, D., Dawson, S., & Jovanovic, J. (2016). Ethics and Privacy as Enablers of Learning Analytics. *Journal of Learning Analytics, 3*, 1–4.

Groves, R. M., Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology* (2nd ed.). John Wiley & Sons.

Heath, J. (2014). Contemporary Privacy Theory Contributions to Learning Analytics. *Journal of Learning Analytics, 1*, 140-149.

Herodotou, C., Naydenova, G., Boroowa, A., Gilmour, A., & Rienties, B. (2020). How Can Predictive Learning Analytics and Motivational Interventions Increase Student Retention and Enhance Administrative Support in Distance Education? *Journal of Learning Analytics, 7*, 72–83.

Hoel, T., & Chen, W. (2016). Privacy-driven design of learning analytics applications–exploring the design space of solutions for data sharing and interoperability. *Journal of Learning Analytics, 3*, 139–158.

Iachello, G., & Hong, J. (2007). End-User Privacy in Human–Computer Interaction. *Foundations and Trends® in Human–Computer Interaction, 1*, 1–137. doi:10.1561/1100000004

Ifenthaler, D., & Schumacher, C. (2016). Student Perceptions of Privacy Principles for Learning Analytics. *Educational Technology Research and Development, 64*, 923-938.

Ifenthaler, D., & Schumacher, C. (2016). Student Perceptions of Privacy Principles for Learning Analytics. *Educational Technology Research and Development, 64*, 923–938.

Ifenthaler, D., & Schumacher, C. (2019). Releasing personal information within learning analytics systems. In *Learning Technologies for Transforming Large-Scale Teaching, Learning, and Assessment* (pp. 3–18). Springer.

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early alert of academically at-risk students: An open source analytics initiative. *Journal of Learning Analytics, 1*, 6–47.

Joksimović, S., Kovanović, V., & Dawson, S. (2019). The journey of learning analytics. *HERDSA Review of Higher Education, 6*, 27–63.

Jones, K. M., Asher, A., Goben, A., Perry, M. R., Salo, D., Briney, K. A., & Robertshaw, M. B. (2020). "We're being tracked at all times": Student perspectives of their privacy in relation to learning analytics in higher education. *Journal of the Association for Information Science and Technology*.

Kappers, W. M., & Cutler, S. L. (2014). Poll Everywhere! Even in the Classroom: An Investigation into the Impact of Using Poll Everywhere in a Large Lecture Classroom. *The ASEE Computers in Education (CoED) Journal, 6*, 21.

Knight, S., Rienties, B., Littleton, K., Mitsui, M., Tempelaar, D., & Shah, C. (2017). The relationship of (perceived) epistemic cognition to interaction with resources on the internet. *Computers in Human Behavior, 73*, 507–518.

Knox, J. (2017). Data Power in Education: Exploring Critical Awareness with the "Learning Analytics Report Card". *Television & New Media, 18*, 734–752. doi:10.1177/1527476417690029

Kobsa, A. (2007). Privacy-enhanced Personalization. *Communications of the ACM, 50*, 24–33. doi:10.1145/1278201.1278202

Krosnick, J. A. (2018). Questionnaire Design. In D. L. Vannette, & J. A. Krosnick (Eds.), *The Palgrave Handbook of Survey Research* (pp. 439–455). Springer International Publishing. doi:10.1007/978-3-319-54395-6_53

Kumaraguru, P., & Cranor, L. F. (2005). *Privacy Indexes: A Survey of Westin's Studies.* Carnegie Mellon University, School of Computer Science, Institute for Software Research International.

Kuzilek, J., Hlosta, M., Herrmannova, D., Zdrahal, Z., & Wolff, A. (2015). OU Analyse: analysing at-risk students at The Open University. *Learning Analytics Review*, 1–16.

Lee, H., Wong, S. F., Oh, J., & Chang, Y. (2019). Information Privacy Concerns and Demographic Characteristics: Data from a Korean Media Panel Survey. *Government Information Quarterly, 36*, 294-303. doi:https://doi.org/10.1016/j.giq.2019.01.002

Li, W., Sun, K., Schaub, F., & Brooks, C. (2021). Disparities in Students' Propensity to Consent to Learning Analytics. *International Journal of Artificial Intelligence in Education*. doi:10.1007/s40593-021-00254-2

Lietz, P. (2010). Research into Questionnaire Design. *International Journal of Market Research, 52*, 249–272.

Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review, 46*, 30–40.

Mittelmeier, J., Edwards, R. L., Davis, S. K., Nguyen, Q., Murphy, V. L., Brummer, L., & Rienties, B. (2018). 'A double-edged sword. This is powerful but it could be used destructively': Perspectives of Early Career Education Researchers on Learning Analytics. *Frontline Learning Research, 6*, 20–38.

Naeini, P. E., Bhagavatula, S., Habib, H., Degeling, M., Bauer, L., Cranor, L. F., & Sadeh, N. (2017). Privacy Expectations and Preferences in an IoT World. *Thirteenth Symposium on Usable Privacy and Security*, (pp. 399–412).

Nevaranta, M., Lempinen, K., & Kaila, E. (2020). Students' Perceptions about Data Safety and Ethics in Learning Analytics. *Proceedings of the Conference on Technology Ethics*, (pp. 23–37).

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life.* Stanford University Press.

Olson, J. S., Grudin, J., & Horvitz, E. (2005). A Study of Preferences for Sharing and Privacy. *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (pp. 1985–1988). ACM. doi:10.1145/1056808.1057073

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology, 50*, 128–138. doi:https://doi.org/10.1111/bjet.12592

Pijeira-Díaz, H. J., Drachsler, H., Järvelä, S., & Kirschner, P. A. (2016). Investigating collaborative learning success with physiological coupling indices based on electrodermal activity. *Proceedings of the Sixth international conference on Learning Analytics & Knowledge*, (pp. 64–73).

Preibusch, S. (2013). Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human-Computer Studies, 71*, 1133–1143.

Rienties, B., & Héliot, Y. (2018). Enhancing (In)Formal Learning Ties in Interdisciplinary Management Courses: A Quasi-Experimental Social Network Study. *Studies in Higher Education, 43*, 437–451. doi:10.1080/03075079.2016.1174986

Rienties, B., & Toetenel, L. (2016). The Impact of Learning Design on Student Behaviour, Satisfaction and Performance: A Cross-institutional Comparison Across 151 Modules. *Computers in Human Behavior, 60*, 333–341.

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education. *Technology, Knowledge and Learning, 22*, 317–333.

Roberts, L. D., Howell, J. A., Seaman, K., & Gibson, D. C. (2016). Student Attitudes toward Learning Analytics in Higher Education: ``The Fitbit Version of the Learning World". *Frontiers in Psychology, 7*.

Schumacher, C., & Ifenthaler, D. (2018). Features Students Really Expect from Learning Analytics. *Computers in Human Behavior, 78*, 397–407.

Slade, S., & Prinsloo, P. (2014). Student Perspectives on the Use of their Data: Between Intrusion, Surveillance and Care. *Challenges for Research into Open & Distance Learning: Doing Things Better ? Doing Better Things*, (pp. 291–300).

Slade, S., Prinsloo, P., & Khalil, M. (2019). Learning Analytics at the Intersections of Student Trust, Disclosure and Benefit. *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 235–244). ACM. doi:10.1145/3303772.3303796

Sun, K., Mhaidli, A. H., Watel, S., Brooks, C. A., & Schaub, F. (2019). It's My Data! Tensions Among Stakeholders of a Learning Analytics Dashboard. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 594:1–594:14). ACM. doi:10.1145/3290605.3300824

Taylor, H. (2003). Most People Are ``Privacy Pragmatists'' Who, While Concerned About Privacy, Will Sometimes Trade It Off for Other Benefits. *Most People Are ``Privacy Pragmatists'' Who, While Concerned About Privacy, Will Sometimes Trade It Off for Other Benefits*. Harris Interactive.

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2020). The privacy paradox and its implications for learning analytics. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, (pp. 230–239). doi:10.1145/3375462.3375536

Tsai, Y.-S., Whitelock-Wainwright, A., & Gašević, D. (2020). The privacy paradox and its implications for learning analytics. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, (pp. 230–239).

Vu, P., Adkins, M., & Henderson, S. (2019). Aware, But Don't Really Care: Students' Perspective on Privacy and Data Collection in Online Courses. *Journal of Open Flexible and Distance Learning, 23*, 42-51.

Westin, A. F., & Maurici, D. (1998). *E-commerce and Privacy: What Net Users Want.* Privacy and American Business Hackensack, NJ.

Whitelock-Wainwright, A., Gašević, D., Tejeiro, R., Tsai, Y.-S., & Bennett, K. (2019). The student expectations of learning analytics questionnaire. *Journal of Computer Assisted Learning, 35*, 633–666.

Woodruff, A., Pihur, V., Consolvo, S., Brandimarte, L., & Acquisti, A. (2014). Would a Privacy Fundamentalist Sell Their DNA for $1000...If Nothing Bad Happened as a Result? The Westin Categories, Behavioral Intentions, and Consequences. *10th Symposium On Usable Privacy and Security (SOUPS 2014)* (pp. 1–18). Menlo: USENIX Association.

Worcester, R. M., & Burns, T. R. (1975). Statistical Examination of Relative Precision of Verbal Scales. *Journal of the Market Research Society, 17*, 181–197.

Xu, H., & Teo, H.-H. (2004). Alleviating Consumers' Privacy Concerns in Location-Based Services: A Psychological Control Perspective. *International Conference on Information Systems (ICIS).*

APPENDIX

Table A1: Descriptive statistics of students' cultural backgrounds and nationalities

Cluster	No. of students	Percentage	Countries and no. of students for each
Confucian Asian	73	65.7	China (67), Taiwan (5), Hong Kong (1)
Anglo	16	14.4	UK (13), USA (3)
Southern Asia	10	9.01	India (5), Malaysia (2), Thailand (2), Vietnam (1)
Eastern Europe	6	5.41	Greece (5), Slovak (1)
Germanic Europe	2	1.8	Austria (1), Netherlands (1)
Sub-Saharan Africa	2	1.8	Nigeria (1), Tanzania (1)
Latin Europe	1	0.9	Italy (1)
Middle East	1	0.9	Turkey (1)

Table A2: Questions on general caution

Questions on general caution	Response options
Do you shred/burn your personal documents when you are disposing of them?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you hide your bank card PIN number when using cash machines / making purchases?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you only register for websites that have a privacy policy?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you read a website's privacy policy before you register your information?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you look for a privacy certification on a website before you register your information?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you read license agreements fully before you agree to them?	Never, Rarely, Sometimes, Very often, Always, Not applicable

Table A3: Questions on technical protection

Questions on technical protection	**Response options**
Do you watch for ways to control what people send you online (such as check boxes that allow you to opt-in or opt-out of certain offers)?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you remove cookies?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you use a pop-up window blocker?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you check your computer for spyware?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you clear your browser history regularly?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Do you block messages/emails from someone you do not want to hear from?	This question was excluded as it was not relevant for the study context

Table A4: Questions on privacy concern

Questions on privacy concern	**Response options**
In general, how concerned are you about your privacy while you are using the Internet?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned about online organisations not being who they claim they are?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned that you are asked for too much personal information when you register or make online purchases?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned about online identity theft?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned about people online not being who they say they are?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned that information about you could be found on an old computer?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned who might access your medical records electronically?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned about people you do not know obtaining personal information about you from your online activities?	Never, Rarely, Sometimes, Very often, Always, Not applicable
Are you concerned that if you use your credit card to buy something on the Internet your credit card	Never, Rarely, Sometimes, Very often, Always, Not applicable

number will be obtained/intercepted by someone else?	
Are you concerned that if you use your credit card to buy something on the internet your card will be mischarged?	Never, Rarely, Sometimes, Very often, Always, Not applicable

Table A5: Questions on the Amazon vignette

Questions on the Amazon vignette	**Response options**
Have you signed up for an Amazon account?	Y/N
I feel comfortable that Amazon can offer me a better service (e.g., offers based on my buying or search patterns) by collecting my personal data?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that Amazon shares my personal and online activity data, in a personally identifiable way, with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that Amazon shares my personal and online activity data, in an anonymised format, with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online and assures me that my data will not be shared with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online on condition that my data will be shared with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree

Table A6: Questions on the university vignette

Questions on the university vignette	**Response options**
I would feel comfortable that my personal and online activity data is shared with my tutor to help him/her to improve support to me	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University can offer me a better service (e.g., alerts on potential problems or recommendations of learning resources) by collecting my personal data?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University shares my personal and online activity data, in a personally identifiable way, with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University shares my personal and online activity data, in an anonymised format, with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University offers me specific benefits in exchange for tracking me online?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University offers me specific benefits in exchange for tracking me online and assures me that my data will not be shared with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree
I feel comfortable that the University offers me specific benefits in exchange for tracking me online on condition that my data will be shared with third parties?	Totally disagree, Disagree, Neutral, Agree, Totally agree

Follow up interview schedule

This interview is a follow up to the Organizational Behaviour (OB) lab where we focused on personality and privacy using two scenarios – personalised recommendations on things you can purchase from Amazon and using student data for a student learning dashboard and to improve learning at the University (this was a hypothetical scenario). I would like to discuss your responses to the OB lab questions and understand more about your perspective on how organizations use customer data.

Settling in

To get us settled in, could you tell me briefly about your experience as a graduate student at the University?

Probe: Build on what they mention of interest/relevance to settle in

Questions on responses in the lab session

In the lab session we looked at two scenarios – Amazon and a student facing learning dashboard. I will recap the questions and remind you of your response. I will then invite you to tell me more about your response.

The first question I would like to focus on asked "I feel comfortable that Amazon can offer me a better service (e.g., offers based on my buying or search patterns) by collecting my personal data?" and "I feel comfortable that the University can offer me a better service (e.g., alerts on potential problems or recommendations of learning resources) by collecting my personal data?" You stated [insert student's answer]. Could you tell me more about your responses to these questions?

Next, we asked "I feel comfortable that Amazon shares my personal and online activity data, in a personally identifiable way, with third parties?" and "I feel comfortable that the University shares my personal and online activity data, in a personally identifiable way, with third parties?" You stated [insert student's answer]. Could you tell me more about your responses to these questions?

Next, we asked "I feel comfortable that Amazon shares my personal and online activity data, in an anonymised format, with third parties?" and "I feel comfortable that the University shares my personal and online activity data, in an anonymised format, with third parties?" You stated [insert student's answer]. Could you tell me more about your responses to these questions?

Next, we asked "I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online?" and "I feel comfortable that the University offers me specific benefits in exchange for tracking me online?" You stated [insert student's response]. Could you tell me more about your responses to these questions?

We asked "I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online and assures me that my data will not be shared with third parties?" and "I feel comfortable that the University offers me specific benefits in exchange for tracking me online and assures me that my data will not be shared with third parties?" You stated [insert student's answer]. Could you tell me more about your responses to these questions?

We asked "I feel comfortable that Amazon offers me specific benefits in exchange for tracking me online on condition that my data will be shared with third parties?" and "I feel comfortable that the University offers me specific benefits in exchange for tracking me online on condition that my data will be shared with third parties?" You stated [insert student's answer]. Could you tell me more about your responses to these questions?

There are a number of ways that people define or think about privacy. Could you tell me what you think privacy is in the specific context where your data is used for the learning dashboard and to improve your learning?

Controlling use of data

Is there data that you would not want to be used in preparing the learning analytics dashboard and to improve your learning?

Probe: could you tell me more about why you would want to exclude some data from use?

Probe: Could you tell me more about why you would not want to exclude some data from use?

Benefits of the learning dashboard

Do you think there are benefits to you personally if you use the student learning dashboard?

Thank you and wrap up

Do you have anything to add that we have not talked about?

Thank you for your time.

Table 1: Examples of questions used in the study

Scale	N items	Example item	Response scale	M	SD	Alpha
General caution	6	Do you shred/burn your personal documents when you are disposing of them?	[1] Never - [5] Always	2.82	.86	.755
Technical protection	5	Do you watch for ways to control what people send you online (such as check boxes that allow you to opt-in or opt-out of certain offers)?	[1] Never - [5] Always	2.95	.8	.665
Privacy concern	10	In general, how concerned are you about your privacy while you are using the Internet?	[1] Never - [5] Always	3.58	.66	.836
Concern about data collection and use - Amazon	6	I feel comfortable that Amazon can offer me a better service (e.g., offers based on my buying or search patterns) by collecting my personal data?	[1] Totally disagree - [5] Totally agree	2.61	.457	.590
Concern about data collection and use - University	6	I feel comfortable that the University shares my personal and online activity data, in a personally identifiable way, with third parties?	[1] Totally disagree - [5] Totally agree	2.77	.485	.668

Table 1: Factor analysis results for comfort with collection and use of data - University

Factor	Items and loading	Proportion variance	Alpha	Mean	SD
Factor 1 – comfort with data use and data sharing	… can offer me a better service by collecting my personal data (.80); … shares my data in a personally identifiable way (.73); … shares my data in an anonymised format (.69)	38.1%	.61	2.94	.54
Factor 2 – comfort with benefits for tracking	… offers benefits for tracking (.84); offers specific benefits for tracking and data shared (.82); offers specific benefits for tracking and data not shared (.49)	19.5%	.61	2.59	.63

Table 1: Factor analysis results for comfort with collection and use of data - Amazon

Factor	Items and loading	Proportion variance	Alpha	Mean	SD
Factor 1 – comfort with benefits for tracking and no data sharing	… offers benefits for tracking (.808); offers specific benefits for tracking and data not shared (.918)	34.3%	.70	2.86	.60
Factor 2 – comfort with benefits for tracking and identifiable data sharing	… shares my data in a personally identifiable way (.842); offers specific benefits for tracking and data shared (.598)	19.3%	.39	1.93	.56
Factor 3 – comfort with data use and anonymised data sharing	… can offer me a better service by collecting my personal data (.549); shares my data in an anonymised format (.934)	18%	.42	2.71	.65

Table 1: Mean and Standard Deviation Results for Individual Items in the Amazon and University Scales

	Amazon		**University**	
I feel comfortable that ….	**Mean**	**Std. dev.**	**Mean**	**Std. dev.**
… can offer me a better service (e.g., offers based on my buying or search patterns) by collecting my personal data	3.11	.84	3.93	.64
… shares my personal and online activity data, in a personally identifiable way, with third parties	1.82	.72	2.08	.76
… shares my personal and online activity data, in an anonymised format, with third parties	2.31	.8	2.81	.77
… offers me specific benefits in exchange for tracking me online	2.68	.9	2.71	.93
… offers me specific benefits in exchange for tracking me online and assures me that my data will not be shared with third parties	3.68	.82	3.14	.95
… offers me specific benefits in exchange for tracking me online on condition that my data will be shared with third parties	2.04	.71	1.93	.63

Table 1: Mean, standard deviation, normality results and Spearman correlations for study scales

	Mean	SD	Shapiro-Wilk	Sig.	1	2	3	4	5
General caution	2.82	.819	.980	.086	1				
Technical protection	2.95	.795	.979	.080	.449**	1			
Privacy concern	3.58	.657	.989	.483	.200*	.117	1		
Comfort with data use and sharing_F1	2.94	.635	.920	.000	.067	.128	-.017	1	
Comfort with benefits for tracking_F2	2.59	.485	.935	.000	-.090	-.136	-.095	.361**	1

** Correlation is significant at the 0.01 level (2-tailed); * Correlation is significant at the 0.05 level (2-tailed)

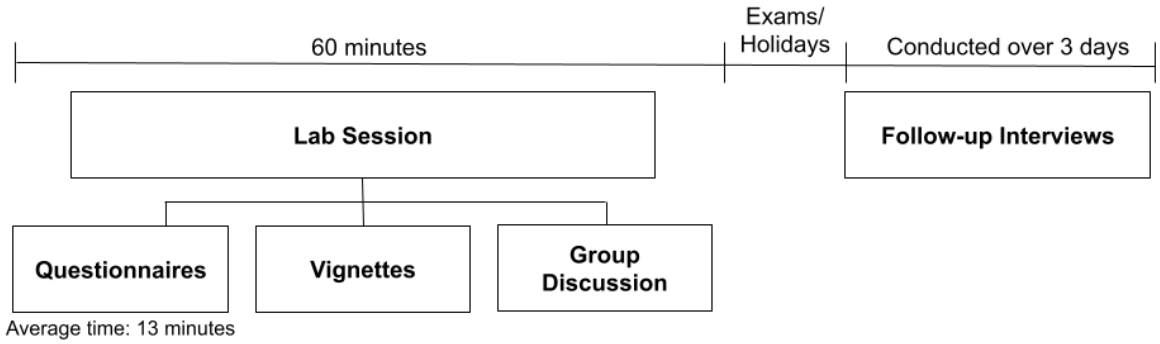

Figure 1: The design of the study

The authors declare that there is no conflict of interest.