

Perspectives from the stakeholder: Students' views regarding learning analytics and data collection

Deborah West
Flinders University

Ann Luzeckyj
Centre for Innovation in Learning and Teaching, Flinders University

Bill Searle
Charles Darwin University

Danny Toohey
School of Engineering and Information Technology, Murdoch University

Jessica Vanderlelie
Student Success, La Trobe University

Kevin R. Bell
Florida State University, Panama City, United States of America

This article reports on a study exploring student perspectives on the collection and use of student data for learning analytics. With data collected via a mixed methods approach from 2,051 students across six Australian universities, it provides critical insights from students as a key stakeholder group. Findings indicate that while students are generally comfortable with the use of data to support their learning, they do have concerns particularly in relation to the use of demographic data, location data and data collected from wireless networks, social media and mobile applications. Two key themes emerged related to the need for transparency to support informed consent and personal-professional boundary being critical. This supports findings from other research, which reflects the need for a nuanced approach when providing information to students about the data we collect, including what we are collecting, why and how this is being used.

Implications for practice or policy:

- When implementing the use of dashboards, institutions should ideally include opportunities for students to opt in and out, rather than being set so that students have agency over their data and learning.
- When undertaking work in relation to learning analytics, staff need to ensure the focus of their work relates to student learning rather than academic research.
- When institutions and academic staff collect and use student data (regardless of the purpose for doing so), all aspects of these processes need to be transparent to students.

Keywords: student perspectives, learning analytics, stakeholders, data collection, student data

Introduction

This article reports on research designed to explore students' attitudes to, and experience with, learning analytics (LA) in higher education. It builds on a previous study (West, Luzeckyj, et al., 2018; West, Tasir, et al., 2018) that explored teaching academics' views regarding key metrics that support learning and teaching in higher education. While academic insights provide critical value to help shape LA developments, they are only one of the major stakeholders. Student understandings of LA are not well explored, leaving a gap in our knowledge as there is little published work in this area within the sector in Australia or internationally. Additionally, the review of student perspectives on the merits and concerns related to LA is a crucial first step to underpin participatory design and co-creation of LA in higher education.

A brief background to the study foregrounds the paper, positioning it within a broader context of changes in the sector while relating it to the Innovative Research Universities (IRU) research upon which it builds. A literature review highlighting limited publications exploring student views undertaken to date then follows. An overview of the methodology and key findings is presented to support the discussion. The lessons provided by this study form the conclusion.

Background

The IRU is “a coalition of seven comprehensive universities committed to inclusive excellence in teaching, learning and research in Australia” (IRU, n.d.). As part of this mission, the IRU supports research into key issues related to contemporary higher education, including in the field of LA. A joint IRU and Malaysian Research University Network study completed in 2017 explored academic teachers’ views of LA, including the key ways that the teachers thought this approach would support the learning and teaching process (West, Luzecky, et al., 2018; West, Tasir, et al., 2018). The study identified a range of reports and data points, including student-facing ones. Overall, academics had strong opinions on what students need to do to improve outcomes.

The use of technology to support learning and teaching provides a viable digital footprint that lends itself to more effective LA. However, with such data come concern and responsibilities related to its appropriate use, data security and informed consent. A recent Jisc comparative study between Australia and New Zealand and the United Kingdom found increasing usage of digital technology in universities, with Australia remaining advanced in this space (Newman et al., 2019). The study also found that while the use of digital technology is increasing, students are concerned about the safety and privacy of their data in the digital world. With the connections between educational data, LA and the ethics and legalities of data use, the views of students are critical. With this important insight and the rate of progress in each institution on LA, the IRU commissioned the study to explore students’ views and attitudes on the use of LA to support their learning and to identify what types of data and reports they felt would be most useful.

Literature review

LA offers powerful opportunities to support the student experience, improve retention and enhance the quality of learning and teaching in higher education. Despite students being primary stakeholders for LA, there has been a distinct lack of student perspectives in LA decision-making across the sector and limited studies that empower students to contribute to the discussion (i.e., Bodily & Verbert, 2017; de Quincey et al., 2019; Prieto-Alvarez et al., 2018; Schumacher & Ifenthaler, 2018; Teasley & Whitmer, n.d.; Wook et al., 2016). However, most of these studies are small with limited student numbers. The notable exception is a mixed methods study undertaken by Arnold and Sclater (2017), who compared the views of students in relation to LA applications and privacy across two institutions in the United Kingdom and the United States of America. Other work has been either from the perspective of academics or it has been theoretical.

Australian researchers have reported that students lack awareness of what constitutes LA and have a limited understanding of the range of data currently collected by universities, who has access to it and the purposes for which the data is used (Fisher et al., 2014). Despite this lack of awareness, the limited work which directly seeks students’ views suggests that when provided with additional information, students generally express positive attitudes towards LA and can see the benefits and potential impact on their learning and experience (Arnold & Sclater, 2017; Brooker et al., 2017; Ifenthaler & Schumacher, 2016; Roberts, Howell, et al., 2016; Schumacher & Ifenthaler, 2018). Other positive aspects referenced in the literature from the student perspective include opportunities to increase the personalisation of the learning experience and the potential for improvements in student engagement, motivation and connection to academic content. Students have also indicated the potential for LA to support academics to understand them better and enable teachers to provide support or extension activities to students as required. Furthermore, they consider opportunities to connect with other students and improve the curriculum as important positives of LA (Brooker et al., 2017; Ifenthaler & Schumacher, 2016; Roberts, Howell, et al., 2016; Schumacher & Ifenthaler, 2018).

Additionally, students see value in the capacity of LA to enable them to compare their progress to that of others (Atif et al., 2015; Corrin & de Barba, 2014; Reimers & Neovesky, 2015). However, evidence also

suggests that students are concerned by the culture of competition over collaboration that would result from comparison data and highlighted apprehension regarding the potential impacts on student confidence and emotional well-being (Chen et al., 2018; Roberts, Howell, et al., 2016). Other negative perspectives expressed by students included the possibility that access to progress data may have the consequence of disengaging learners who have fallen behind and may also extend to those who are not achieving to a high standard despite significant effort (Jivet et al., 2017; Roberts, Chang, et al., 2016). Frequently, students identified a concern that LA would remove their anonymity and freedom to negotiate the learning environment as they choose (Beattie et al., 2014; Roberts, Howell et al., 2016) and ultimately impact upon their privacy (Hoel & Chen, 2016).

A deeper exploration of student privacy concerns suggests that despite being reasonably comfortable with sharing information related to their learning and engagement inside university systems, students are reluctant to share personal information, online behaviour, physical location and social media engagement outside of the learning environment (Ifenthaler & Schumacher, 2016). Importantly, students want the opportunity to opt in and out of sharing information and accessing dashboards, as noted by Roberts, Howell, et al. (2016). These conservative views on data sharing will no doubt limit the predictive and personalisation capabilities of analytics systems (Ifenthaler & Schumacher, 2016; Ifenthaler & Widanapathirana, 2014). However, several authors have hypothesised that by developing dashboards and raising student awareness, students will see the benefits of data sources to their learning and therefore be more willing to disclose a richer array of data sources (Ifenthaler & Schumacher, 2016; Scholes, 2016; Slade & Prinsloo, 2013).

A theme that is gaining traction and often discussed in the theoretical literature, alongside privacy, which leads directly to the issue of informed consent, is the strong desire from students for increased transparency in relation to the range of data collected and its use by universities (Slade & Tait, 2019). Currently, universities collect a range of data about students, their learning and engagement within the online and on-campus environment. In most cases, consent for the collection of this information is provided at enrolment, where the range of data being collected and the purposes for which it may be used are not necessarily made clear to students (Lawson et al., 2016; Slade & Tait, 2019). Lawson et al. (2016) have advocated for institutions to be clear about the purposes for which consent is provided and how data will be used. Pardo and Siemens (2014) have suggested that in addition to the principles of autonomy, beneficence, non-maleficence and distributive justice (Prinsloo & Slade, 2014; Slade & Prinsloo, 2013), it is important for institutions to consider the values of fostering trust, transparency and accountability.

To achieve these values, a focus on transparency must be built and applied at every stage of LA decision-making, data collection and use. Working with students provides opportunities for institutions to develop and supplement strong policies (Greller & Drachler, 2012) for LA decision-making (Slade & Prinsloo, 2013) with the creation of analytics approaches that recognise students' concerns and provide maximum opportunities for students to deepen their understanding. In addition, it is important to note that longitudinal changes in student attitudes remain unexplored, and variations between disciplines or demographics are not fully understood (Marzouk et al., 2016; Roberts, Howell, et al., 2016). If we are to work with students and take their views on board, it is vital that we explore and acknowledge student attitudes and understanding of the meaning of LA, the type of data collected and their concerns about its use (Prieto-Alvarez et al., 2018). Dollinger and Lodge (2018) also emphasised this point in their paper about issues and barriers and the use of co-creation strategies to address them. Parkes et al. (2020) also highlighted the importance of involving students from the outset (collecting and sharing data together) so that education becomes a democratic process. Our study adds to the work of those indicated here by further determining student understanding and concerns regarding LA so these views may be taken on board by institutions, thus allowing LA outcomes to meaningfully inform both policy and practice (for all stakeholders).

Methodology

The study discussed here was designed to explore students' perceptions in relation to LA and its key aims were twofold: firstly, to identify from the student perspective the data institutions can collect, leverage and provide to learners to support their university learning journey; and secondly, to identify the concerns learners have about their data and its use. Human research ethics approval was sought and received in late 2017 from one primary institution, and subsequent approvals provided by all participating institutions in line with the National Health and Medical Research Council (2019) guidelines.

The research methodology was based on the approach trialled when exploring the IRU and MRUN academic staff understanding of LA (West, Luzeckyj, et al., 2018). It was a three-stage process involving conducting initial focus groups with students, developing and distributing a survey which was based on discussions in the first focus groups and conducting follow-up focus groups to delve more deeply into survey results. The survey across the institutional membership of each group comprised mainly Likert responses, except where questions were linked to the respondents' demographics. In addition, open-ended questions inviting respondents to comment on their Likert responses were included. Survey responses were then scrutinised to identify questions requiring further exploration during the focus groups. This second set of focus groups was arranged with staff at the IRU institutions (in Australia) and explored areas where survey respondents indicated concerns or where there was broad agreement or disagreement across responses. The focus groups enabled us to delve more deeply into areas of interest or complexity identified in the survey results. Given this approach had been successful in our research with academic staff (West, Luzeckyj et al., 2018) and has been used and has been an accepted approach across sociology and other social sciences since the early 1990s (O'Brien, 1993; Wolff et al., 1993), we replicated it with students from six of the seven IRU institutions.

The questions for the survey were developed based on a review of the literature regarding student perspectives in LA (West et al., 2020) and work undertaken in and findings from the previous academic-focused study (West, Luzeckyj, et al., 2018; West, Tasir, et al., 2018) with further validation via the initial focus groups with students. Given most published work in this domain is theoretical in nature and very few studies had taken place with students when we initially embarked on our study in 2017 (with those that had taken place being almost entirely focus group-based and/or limited in scope), no survey instrument which existed at the time was deemed appropriate to repurpose. The initial set of focus groups was run with students in 2017 to help identify and shape the key questions for the survey. Fourteen students, who responded to a call for participants from four universities, participated in the first set of focus groups. The focus group participants were provided with a short introduction to LA and how data is collected in online systems. Participants were then asked about the type of data they thought institutions collected, what information the university can provide that could support them across their learning journey and the concerns they may have about the data collected and its use. These focus groups identified a high degree of similarity in the understanding and misconceptions surrounding LA, and from the areas identified and literature review, we developed the range of options provided in each survey question. However, to ensure clarity, the survey included brief information to contextualise the question.

The first set of focus groups was run as a two-part exercise which utilised both individual identification of responses via the use of sticky-note responses and discussion enabling ongoing analysis within the group as it progressed. In the first exercise, students were asked questions about the information they would find helpful to improving their experience at university. A series of prompt questions designed to aid discussion and help group responses then followed.

The second exercise followed a similar methodology. Students were asked to complete a specific task related to the data they believed universities collect about them and were provided with a series of questions designed to aid discussion and help with clustering responses and concurrent analysis within the group. The focus group discussion was recorded, transcribed and analysed using thematic analysis – a qualitative research method that explores data and enables themes to emerge (Fereday & Muir-Cochrane, 2006). Braun and Clarke (2006) defined themes as important elements in the data and suggested themes demonstrate where “some level of patterned response or meaning within the data set” occurs (p. 82).

As discussed above, the results from analysis of the initial focus group and the literature review were used to develop the survey. The survey questions were designed to identify both understanding and misunderstanding, and as such, it is important that we based our questions on the themes and concerns identified through our initial focus groups. We were very interested in any misconceptions students may have had so that the findings could provide guidance to improve institutional communication. The survey consisted of:

- questions related to respondents' demographic details to enable analysis across different student cohorts
- 23 questions about what data students think universities collect about them

- 23 questions about their level of comfort (on a 5-point Likert scale, with *very comfortable* scoring 5, and *very uncomfortable* scoring 1) concerning the use of the above data points to help support their learning
- 14 questions related to how useful they thought types of data and reports would be to their learning experience recorded using a 3-point Likert scale (indicating *positive*, *neutral*, *negative*)
- 10 questions related to how students felt about their data being used or handled in particular ways with responses being on 5-point Likert scale (*very concerned* scoring 5, and *not at all concerned* scoring 1)
- three questions about how they would like their information displayed consisting of “yes”, “no”, “not sure” responses
- one question related to frequency of notification regarding policies and practice on student data collection and use
- one open-ended question sought any other comments in relation to the ethics of digital data collection and use in higher education.

The survey was hosted online using the Qualtrics software platform, with the invitation to participate distributed to all undergraduate and postgraduate on-shore coursework students in six of the seven IRU universities (approximately 158,000 students) via email in Semester 1 of 2018. There were 2,017 valid responses; and while we acknowledge this is a small number (1% of potential responses), we believe our data set is one of the largest available on student perceptions to date. The survey data were processed in SPSS version 25 and Microsoft Excel. The Findings section includes further detail regarding the application of analytical techniques to the various questions.

Follow-up focus groups were held after survey analysis (in late 2018), to clarify the reasons behind results related to respondents’ understanding of key terminology and their level of comfort with key data types. All focus groups followed a set running sheet and explored each of the data types– [demographics] [location data] [social media group] [wireless network device usage] [mobile app services] [employment services data] [data to support you] – using these questions:

- What is your understanding of the term [data type]?
- What is it about [data type] that you think may cause students discomfort?
- Would you be concerned if the university used your [data type] to:
 - Identify risk markers to assist with retention strategies
 - Create a profile to determine risk or advise particular services
 - Provide you with a dashboard that you could use to support your learning
 - Develop dashboards that academics/teaching staff could use to understand learning patterns.

As data types vary considerably, we wanted to provide examples and explore these in relation to students’ comfort levels and the reasons for those. The following list provides key examples from the discussion:

- Access to and use of the learning management system (LMS) Blackboard or Moodle)
- Use of and access to university apps
- Access to email
- Physical location on campus
- Information about downloads
- Access to career advice services offered by the university
- Access to assistance with curriculum vitae
- Location on campus
- Class attendance
- Library use
- Access to social media sites (e.g., Facebook) site set up by the lecturer or fellow students for a topic, unit or course
- Access to social media sites set up by the Student Union or Student Association.

The second area explored further in the focus groups was associated with students' levels of concern about the university accessing and using their data. The questions within this area were also designed to determine students' knowledge of and comfort with data collection. These questions were linked to safety and security of data; allowing third parties to access data; using data to support students, improve teaching or trigger support or contact.

The second set of focus groups were conducted at the end of 2018 at four IRU institutions and attracted 20 students, who again responded to a call for participants. Two of us individually read and coded the transcripts, then analysed them using thematic analysis.

Findings

Given the first round of focus groups informed the development of the survey, we present only the results of the survey and the second round of focus groups.

As shown in Table 1, gender was the only demographic variable not comparable in relative proportions to the student body at IRU member universities. The survey cohort included a notably higher proportion of female students than are in the general student population.

Table 1
Demographic distribution of respondents

		% of respondents (<i>N</i> = 2,017)	% IRU universities combined ^a (<i>N</i> = 158,000)
Gender	Male	28.7	40.3
	Female	70.4	59.7
	Other	1.0	-
Origin	Domestic	82.7	78.8
	International	17.3	21.2
Level of study	Undergraduate	76.0	77.6
	Postgraduate	24.0	22.4
Study load	Full-time	79.2	71.6
	Part-time	20.8	28.4

^a Based on the last available national data (Department of Education and Training, 2018)

The survey data was analysed to identify potential differences in the responses based on gender, study status (full- or part-time; undergraduate or postgraduate; point in degree), domestic or international and mode of study. There were no differences in response patterns linked to any of the demographic variables, and patterns of responses were stable irrespective of institution.

The first step in leveraging and providing data to learners from their point of view to support their university learning journey requires gaining an insight into student understanding of concepts and ideas related to the type of data universities collect. To gain these insights, survey participants were asked to indicate what data they thought the university collected about them from a list of 23 options (see Table 2). Table 2 presents the percentage of students who believed that the university collected each data type (where students were asked to indicate *yes*, *no* or *don't know*). The respondents were clearly aware that data was being collected about them, particularly where it was directly relevant to their learning experience, for example, types of data associated with access to the LMS and associated systems. Interestingly, 82.3% of the respondents thought that their wireless network usage was recorded. Conversely, fewer than half of the respondents thought data related to their mobile phone (36.5%), social media (48.9%) or university mobile app usage was being collected.

Table 2

Data respondents thought universities collected about them (N = 2,017)

Type of data collected	% of respondents
Submission of assignments	98.8%
Completion of quizzes	97.5%
Use of text matching/originality software (e.g., Turnitin or SafeAssign or Urkund)	95.0%
Grades from the subjects you have taken	95.0%
When you accessed the LMS	94.6%
Access to particular content in the LMS	92.6%
Activity on discussion boards	91.2%
Demographic information (e.g., age; gender; address)	90.1%
Academic background (previous study, credit applications)	90.1%
Looking at your grades for assignments and quizzes	88.4%
Participation in online lectures, tutorials or web conferencing	85.4%
How long you spend in the LMS	84.4%
Accessing feedback from assignments	84.1%
Access to library borrowing services	83.8%
Wireless network device usage (e.g., university Wi-Fi, eduroam)	82.3%
Access to lecture capture recordings	77.9%
Use of video and audio learning materials	75.4%
Use of academic skills services	74.8%
Access to library support workshops and training	73.9%
University mobile app usage	65.5%
Access to employment services	62.9%
University social media groups	48.9%
Location data from your mobile phone	36.5%

Note. Results are presented in descending order of frequency rather than in the order the questions were asked.

As discussed, the survey included questions to determine respondents' comfort level regarding the university's use of various data sources to help support their learning experience. A summary variable, comfort level, was calculated as the mean score of the responses. As indicated in Table 3, respondents reported feeling more comfortable with items directly associated with their learning experience. The items attracting the lowest degrees of comfort were those not directly related to learning: location data from mobile phone (2.27/5), data about university social media groups (2.82/5) and university mobile app usage (3.06/5).

Table 3
 Respondents' levels of comfort with data usage (N = 2,017)

Use of data	Comfort level (/5)
Submission of assignments	4.25
Completion of quizzes	4.18
When you accessed the LMS	4.10
Use of text matching/originality software (e.g., Turnitin or SafeAssign or Urkund)	4.08
Access to particular content in the LMS	3.99
Accessing feedback from assignments	3.98
Looking at your grades for assignments and quizzes	3.96
How long you spent in the LMS	3.95
Activity on discussion boards	3.91
Participation in online lectures, tutorials or web conferencing	3.88
Access to lecture capture recordings	3.86
Grades from the subjects you have taken	3.85
Use of video and audio learning materials	3.84
Access to library borrowing services	3.83
Access to library support workshops and training	3.80
Use of academic skills services	3.74
Academic background (previous study, credit applications)	3.68
Access to employment services	3.47
Demographic information (e.g., age; gender; address)	3.47
Wireless network device usage (e.g., university Wi-Fi, eduroam)	3.14
University mobile app usage	3.06
University social media groups	2.82
Location data from your mobile phone	2.27

Note. Results are presented in descending order of frequency rather than in the order the questions were asked.

In order to determine students' understanding of the benefits that can potentially arise from LA, the survey included questions asking respondents to indicate whether they thought various practices might be useful to their learning experience (see Table 4). Respondents considered practices primarily concerned with the provision of additional materials or services most positively and were less positive about the provision of information that compared them in some way with other students in the class.

Table 4
Degree of usefulness of specific practices (N = 2,017)

Usefulness of practice in relation to learning experience	% positive responses
You are given information about additional materials (readings, resources) you might like to access based on an assessment you have coming up.	96.3%
You can see your progression through subject material.	96.2%
You are given information about additional services at the university (e.g., academic writing support; library) you might like to access based on an assessment you have coming up.	95.5%
You are given information about additional materials (readings; resources) you might like to access based on any grade received on an assignment or quiz.	94.3%
You are given information about additional services at the university (e.g., academic writing support; library) that you might like to access based on any grade received on an assignment or quiz.	93.5%
You are given information about additional materials (readings, resources) you might like to access based on a LOW grade received on an assignment or quiz.	91%
You are given information about additional services at the university (e.g., academic writing support; library) that you might like to access based on a low grade received on an assignment or quiz.	90.5%
You are given a projection of your likely final grade.	86.2%
You are given information that suggests that you will need to change your study behaviours in order to achieve a passing grade.	83.9%
You are given information that suggests that you will need to change your study behaviours in order to achieve a higher grade.	81.6%
You can see how much you are accessing the LMS.	80.5%
You can see your grades compared to others in class.	71.9%
How your access to the LMS compares to others in your class	61.1%
Number of times you accessed the LMS compared to others in class.	61%

Note. Results are presented in descending order of frequency rather than in the order the questions were asked.

Survey participants were asked to indicate how concerned they felt about several statements related to their data (see Table 5). Almost all respondents (90.3%) indicated concern about third parties receiving their data. Much lower levels of concern (fewer than 50% of the respondents indicated concern) were reflected where options involved provision of support or linked to improvements in learning and teaching, as can be seen in Table 5.

Table 5
Level of concern with data management and use (N = 2,017)

	% concerned responses
Third parties receiving your data	90.3%
Your data being kept safe and secure within the university	68.6%
Your data being used to trigger support services to contact you	63%
Your data being used to trigger academic staff to contact you	62.7%
Your data being used by the university for research	54.2%
Your data being used to provide support to you?	49.3%
Your data being used by academics for research into learning and teaching	49%
Your data being used by academics to improve their teaching	44%
Your data being used by the university to improve services	42.2%
Your data being used by the university to improve learning materials	41%

Note. Results are presented in descending order of frequency rather than in the order the questions were asked.

Questions regarding the display of LA information were also included (see Table 6). Most respondents (77.2%) indicated they did not want a compulsory dashboard to display their information. Rather, they preferred to either opt out of the dashboard (62.1% agree) or turn it on and off. Overall, 79.3% of the participants were clear they wanted a choice.

Table 6
Data display preferences (N = 2,017)

How would you like to have your information displayed?	Yes	No
In a compulsory dashboard	22.8%	77.2%
In a dashboard with the ability to opt out	62.1%	37.9%
In a dashboard you turn on or off	79.3%	20.7%

As discussed in the Literature review section, the literature increasingly suggests students want greater transparency regarding their provision of informed consent (Lawson et al., 2016; Pardo & Siemens, 2014; Slade & Tait, 2019). To explore students' attitudes further, the survey included questions about when students would prefer notification regarding the university's data policies and procedures. As can be seen in Table 7, more than 60% of the respondents preferred to be notified either annually or at the commencement of each semester.

Table 7
Preferred timing of notification of university data policies and procedures (N = 2,017)

Timing of notification	% of responses
When you first enrol at the university	23.05%
At the beginning of each academic year	31.23%
At the beginning of each semester	32.37%
When you enrol in each subject	13.34%

Students clearly wished to be asked for their consent to access data more often than on enrolment.

In order to contextualise students' concerns, it was important to gain greater insight into student understanding of those terms that attracted a lower comfort level. As outlined in the Methodology section, the focus groups provided this opportunity. Students were asked to discuss demographic data, location data, use of social media, wireless network device usage, use of mobile app services, employment services data and their concerns about them.

Participants' comments reflected a good understanding of the term *demographic data*. The following response is indicative of the types of answers students gave: "composition of people, about their age, nationality, distribution; about the lifestyle" (Focus Group 3). Respondents were concerned they may be placed into categories they did not necessarily identify with or that stereotyping, profiling or segregating may occur using their demographic data. The following statement from a Focus Group 1 participant reflects these sentiments:

My response to anyone asking me for information is something like, "Why do you want to know?", so in that case I think it's likely that students are asking you not to value their gender in terms of how good they do at study or what kind of resources you think they need. So, if they don't want their emphasis to be on gender, they're asking you not to focus on it either.

The response above suggests gender was not relevant to student learning and the respondent did not see why it is collected in relation to LA. They are also clearly positioning themselves as wanting to know why we are asking the questions we ask. Several participants held a similar position, requiring confirmation of the relevance of the data institutions were collecting. The kinds of data they questioned included students' location and their use of social media, wireless network device and mobile apps.

Participants were interested in ascertaining why universities might want to know where they were located, what they were doing with the location data and how it could be useful to institutions. The following Focus Group 2 participant cogently articulated these ideas when they said, "I would question why you would want that data and why does it matter and why is that being tracked? And you just start to feel a bit like [you are] being watched a bit." Although less clearly conveyed, other students reflected similar concerns.

In the survey, participants indicated they were more comfortable about the institution "tracking" them when they were using the LMS, if they were on campus or if it meant the data was used to improve services. However, the focus group participants clearly wanted to know how data was used and why, as outlined by this participant:

If we spend most of the time in the library are the university going to do more services for us in the library or for example student learning centre to enhance the service or to get more students involved with the services, is that the purpose? (Focus Group 1)

A Focus Group 4 participant tried to unpack why the university might want to use location data but reflected concern about how data might be used, saying:

It would probably help with the whole online learning thing, to see how many students are staying home. As well as that it's like they already have your home address, so I don't think I'd worry about that aspect. But it'd be going to friends' houses and stuff where it'd be like probably not that in-depth. So, I'm not sure how in-depth this would be. But just home and uni., I wouldn't really have second thoughts about it.

Another participant from the same group (Focus Group 4) indicated they felt the university accessing location data was "creepy", while another, from Focus Group 2, said it was "like big brother is watching me. So, it starts to feel like that". These quotes from the focus groups all indicate similar student concerns regarding the data institutions collect.

While some students were comfortable with teaching staff using social media for in-class purposes, other participants were more adamant about wanting social media to remain private. However, it is not clear whether social media or social media data was under discussion. The following response reflects these points:

My view of social media is, again, for being social. So, in the context of social well-being, I don't think that the lecturers or the facilitators would need to see what I share with my close group of friends or the circle. (Focus Group 3)

Although the response was not clearly articulated, this participant appeared to draw a line between personal and professional boundaries – another theme which prevailed across responses to many questions. When asked about wireless network device usage and the use of mobile apps, participants also drew a line between what was collected and whether it related to their studies (professional) or personal life. Participants also discussed wanting to be able to opt in or out when data was collected. The suggestion that students should be able to do this as well as have a say in the type of data that is shared was another common response across numerous data types.

Although there were no questions explicitly linked to articulating the benefits of LA included in the focus groups, responses clearly indicated a need for benefits to outweigh concerns. This point was specifically raised regarding the collection and use of location data and data related to the use of university employment or careers services. When discussing location data, participants felt the benefits they might receive as a result of the university keeping track of places they visited might help improve services. In the case of employment services, participants wanted the services provided to add value, as reflected by the following selected comment:

I wouldn't see it as a problem because it can be a value-added service in providing information to the students and bridging that gap between the university and the workforce. So, if university staff know what the students are more interested into or what job is most in at the moment, what is being discussed or most searched for, it can also help to improve the services in campus, such as providing career counsellor for that particular field. (Focus Group 4)

But as reflected by another Focus Group 4 participant, their comfort level depended on what data was accessed and how personal it was. The second participant said, "I think if it had something to do with Centrelink, I think people could get discomfort about that sort of thing".

The focus group participants reflected quite discerning and considered understandings of the terms being used and the data collected. They were clear about their comfort levels regarding university access to data and how it was accessed. They desired to opt in and out as they wished; were concerned about the security

of their data and the relevance of data collected to their study; and were not comfortable being monitored, tracked or profiled.

The second set of focus group questions were designed to gain greater insight into participants' specific concerns regarding universities' use of their data. The questions were intended to provide more in-depth insights into participant concerns about data being used to support them, for example, the development of staff-facing and student-facing dashboards as well as students being contacted by the institution. When discussing support triggers, we asked participants who they felt might be most appropriate to contact them (their teachers, professional staff or students in roles such as mentors).

Focus group responses to both sets of questions were similar. Participants required clarity around the intentions and purposes behind the collection of their data. They suggested the terms used were vague, as indicated by this Focus Group 1 response, "I think the phrase 'support you' is similar to say in your options here things like improve services and provide support, those are exceptionally vague phrases. I won't submit to something if I don't know what it is." The response is representative of what others also said.

When asked their thoughts about a dashboard available to academics, most students thought it a good idea and believed staff already access information about them, as stated by a Focus Group 4 participant, "Isn't there already something like that in [the LMS], in the student support section? So, there was like a little student activity section where it has a graph and it has if you've attended a tutorial or something?"

However, when it came to triggering support for students, participants' concerns resurfaced, this time regarding the kind of information under discussion and the type of support being triggered. The following response from Focus Group 2 was similar to what others stated:

The thing is, when you're talking about support, I think it is imperative for a student to know what support might [be being referred to] ... Is it my mental health, is that what the data is being collected for, is it my physical health, is it my learning environment, is it more social circle?

Participants were far more comfortable about being contacted about their learning than other issues (such as health or well-being) and preferred the contact to come from academic staff they knew. They were less happy about receiving generic emails or having peers contact them.

Questions regarding concerns about the use of data to develop dashboards that they might use or that staff might use to understand learning patterns were also included. Participants were generally favourable of the idea of seeing a student-facing dashboard, as indicated by the following statement:

And I think that's really important, like if you're trying to support yourself that's probably the best thing ever, that's a really good idea, the dashboard thing. Because some people like they got an HD [high distinction] and when they see okay, I'm lacking and I need a HD, that provides a theoretical data, which I could improve on and it could help me get a HD. Why not, why not use it? (Focus Group 2)

However, as indicated by the following comment from Focus Group 1, some students were less interested in knowing how they are faring, "I don't really want to know the average of my GPA". Another participant in the same group had a different response, stating:

Partially, I want to know what the average GPA is in my degree. Because that will give me an idea of if I'm above average or on average or below average. But that's it. An average across the degree.

These comments, which were selected because they best conveyed the spirit of the focus group responses, clearly support the student survey responses regarding wanting options to turn access to data on dashboards on or off.

Discussion

The study aim discussed here considers student perspectives regarding how universities access and use their data. The literature review revealed areas of interest identified in other studies, which our findings predominantly confirmed. The main difference between these earlier studies and this study is the size and breadth of the data set. The literature reported studies which were mostly smaller and/or located within one institution; in comparison, our study collected data across six institutions, resulting in 2,017 responses.

Several key themes emerged across our results including concerns about transparency, personal relevance and personal-professional identity. Transparency was inherently connected to informed consent, a critical issue influencing students' comfort levels. The literature has discussed students' lack of awareness regarding what constitutes LA, the data universities collect about them, who has access to it and how it is used, suggesting a lack of transparency by institutions when collecting, analysing and using data (Bodily & Verbert, 2017; Dollinger & Lodge, 2018; Schumacher & Ifenthaler, 2018). Our findings supported this connection with transparency additionally linked to informed consent. The participants in our study indicated comfort with data being collected about them, provided their permission was sought at least once per year and they were given clear information regarding who would use their data and how it would be used.

However, their comfort level depended on what specific data was used, for what purpose and by whom. In essence, students wanted these details before they provided any consent and were more comfortable when they could see the relevance, valued the purpose of the data analysis and understood who specifically would be using it.

Another theme that emerged related to personal relevance of applications of LA; for example, when data was used to support something, or linked to a service they valued and/or their personal values. Some literature has indicated students were comfortable with the use of dashboards that supported them to track progress and/or allowed them to be compared with others in their cohort (Atif et al., 2015; Corrin & de Barba, 2014; Reimers & Neovesky, 2015). However, other research has indicated students are concerned these resources might encourage competition and lead to students experiencing a lack of confidence and/or reduced emotional well-being (Chen et al., 2018; Roberts, Howell, et al., 2016). Participants in this study had mixed views about this issue although results did not appear to relate to any particular demographic, suggesting it may be a personal perspective or preference or connected to the students' own value base. The mixed response also suggests the need for system designs that enable students to choose whether they would like to view these dashboards.

The idea of relevance also connects to personal-professional boundaries. Participants in this study regarded professional boundaries in relation to their role as a student and our business as educators. Therefore, the data collected and utilised needed a clear link to their education studies rather than their personal lives. The respondents were hesitant about any data collection they perceived was not related to learning or teaching matters, a point also reflected in the literature (Brooker et al., 2017; Ifenthaler & Schumacher, 2016; Roberts, Howell, et al., 2016; Schumacher & Ifenthaler, 2018). Students raised concerns about the collection of data related to their demographics, mobile data, use of social media, mobile apps or other data they regarded as linked to their personal lives or the applications they accessed for personal use (or non-study purposes). Some students were more comfortable about the collection of this data if they were specifically informed about its intended use, especially if it might help them in their learning, to improve their engagement with the institution or to improve teaching. Although, as in similar concerns raised in theoretical papers (Lawson et al., 2016; Prinsloo & Slade, 2014; Slade & Prinsloo, 2013), many remained concerned about the potential use of data to categorise them or infringe on their privacy or autonomy.

However, the use of various devices, apps and social media increasingly blurs personal and professional boundaries. Many universities strongly encourage a policy of bring your own device; so, those devices gain use for both professional and personal applications and data collected from them would indiscriminately include both areas. Similarly, the use of social media apps such as Facebook for learning and teaching blurs boundaries regarding data collection. Further, participants regarded the university's collection and use of data from personal devices to gain insights regarding numbers at a public event (e.g., orientation week) as more acceptable than data regarding use of public transport, which students perceived as more personal.

These practices which many considered crossed a personal line, and some saw as “creepy”, may be connected to “tracking”.

With these points in mind, institutions must remain cautious and adhere to core business regarding education and data collection from systems and devices, ensuring these sit appropriately within the institution (e.g., eduroam, computers, the LMS). These boundaries pose some critical challenges as the sector tries to leverage the benefits of data to better support personalised learning and use more nuanced approaches to at-risk students. Additionally, access to student data may be limited and although universities believe we can achieve better modelling and potentially a better outcome for students with the additional data, a line needs to be drawn. As a sector and as academics, we need to consider whose purpose we are trying to achieve (the student’s or the university’s) and do the potential benefits outweigh the risks and the intrusiveness of using the data?

The study has several limitations. Despite the size of the sample, it is small in relation to the population (1%). It is, however, largely representative in terms of the demographic characteristics of the population, and analysis undertaken did not find any significant differences in responses between samples from the six institutions or the demographic characteristics. Additionally, the number of students participating in the focus groups was small due to timing in relation to the academic year. A replication of this study in other institutions would be beneficial to validate the findings.

Conclusions

This study provides important insights into the perceptions of students about their understanding of the data institutions might collect, leverage and provide to learners to support their university learning journey; and secondly, to identify the concerns learners have about their data and its use. These insights should be used to inform how universities proceed in supporting students and guide explorations regarding appropriate and ethical ways forward. Therefore, universities should give considerable thought to the data used, how it is used and under what circumstances or with what parameters.

Both the literature and this study provide evidence that as a sector, we need to take a far more nuanced approach to the use of student data, the information provided to them and the control they have over these elements. Transparency is critical. Clear communication with students on a regular basis about the data being collected but, also, why it is being collected and how it is going to be used are essential, as they are central to ensuring students are given the information they need to provide informed consent. Universities must consider the personal-professional line from the student perspective, ensuring that it is respected as universities progress the LA journey. It is imperative that LA practitioners and institutional policy makers work alongside and with students, rather than use them and the data collected from and about them as objects of research. Aligned with gathering students’ perspectives and respecting those perspectives is a requirement to continue engaging with our students via a range of avenues to explore the benefits and challenges of LA while ensuring an ethical and acceptable way forward.

Acknowledgements

The authors would like to acknowledge the support of the Innovative Research Universities.

References

- Arnold, K. E., & Sclater, N. (2017). Student perceptions of their privacy in learning analytics applications. In *Proceedings of the Seventh International Learning Analytics Knowledge Conference* (pp. 66–69). Association for Computing Machinery. <https://doi.org/10.1145/3027385.3027392>
- Atif, A., Bilgin, A., & Richards, D. (2015). Student preferences and attitudes to the use of early alerts. In T. Bandyopadhyay, D. Beyene, & S. Negash (Eds.), *Proceedings of AMCIS 2015: Twenty-first Americas Conference on Information Systems* (pp. 3410–3423). Association for Information Systems. <http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1297&context=amcis2015>

- Beattie, S., Woodley, C., and Souter, K. (2014). Creepy analytics and learner data rights. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology. Proceedings of the ASCILITE Conference* (pp. 421–425). <https://www.ascilite.org/conferences/dunedin2014/files/concisepapers/69-Beattie.pdf>
- Bodily, R., & Verbert, K. (2017). Review of research on student-facing learning analytics dashboards and educational recommender systems. *IEEE Transactions on Learning Technologies*, 10(4), 405–418. <https://doi.org/10.1109/tlt.2017.2740172>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brooker, A., Corrin, L., Fisher, J. & Mirriahi, N. (2017). Defining “data” in conversations with students about the ethical use of learning analytics. In H. Partridge, K., Davis, & J. Thomas (Eds.), *Me, Us, It! Proceedings of the ASCILITE Conference* (pp. 27–31). <https://2017conference.ascilite.org/wp-content/uploads/2017/11/Concise-BROOKER.pdf>
- Chen, B., Chang, Y. H., Ouyang, F., & Zhou, W. (2018). Fostering student engagement in online discussion through social learning analytics. *The Internet and Higher Education*, 37, 21–30. <https://doi.org/10.1016/j.iheduc.2017.12.002>
- Corrin, L. & de Barba, P. (2014). Exploring students’ interpretation of feedback delivered through learning analytics dashboards. In B. Hegarty, J. McDonald, & S.-K. Loke (Eds.), *Rhetoric and Reality: Critical perspectives on educational technology: Proceedings of the ASCILITE* (pp. 629–633). <https://www.ascilite.org/conferences/dunedin2014/files/concisepapers/223-Corrin.pdf>
- de Quincey, E., Briggs, C., Kyriacou, T., & Waller, R. (2019). Student centred design of a learning analytics system. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 353–362). Association for Computing Machinery. <https://doi.org/10.1145/3303772.3303793>
- Department of Education and Training. (2018). *Selected higher education statistics 2017 full year data – Section 2: All students*. <https://docs.education.gov.au/node/51321>
- Dollinger, M., & Lodge, J. M. (2018). Co-creation strategies for learning analytics. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 97–101). Association for Computing Machinery. <https://doi.org/10.1145/3170358.3170372>
- Fereday, J., & Muir-Cochrane, E. C. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods*, 5(1), 80–92. <https://doi.org/10.1177/160940690600500107>
- Fisher, J., Valenzuela, F.-R., & Whale, S. (2014). *Learning analytics: A bottom-up approach to enhancing and evaluating students’ online learning*. Office for Learning and Teaching. https://ltr.edu.au/resources/SD12_2567_Fisher_Report_2014.pdf
- Greller, W., & Drechsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society*, 15(3), 42–57. <https://www.jstor.org/stable/jeductechsoci.15.3.42>
- 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(1), 139–158. <https://doi.org/10.18608/jla.2016.31.9>
- Ifenthaler, D., & Schumacher, C. (2016). Student perceptions of privacy principles for learning analytics. *Educational Technology Research and Development*, 64(5), 923–938. <https://doi.org/10.1007/s11423-016-9477-y>
- Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning*, 19(12), 221–240. <https://doi.org/10.1007/s10758-014-9226-4>
- Innovative Research Universities. (n.d.). *Purpose*. <https://www.iru.edu.au/about/purpose/>
- Jivet, I., Scheffel, M., Drachsler, H. & Specht, M. (2017). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *European Conference on Technology Enhanced Learning* (Vol. 10474, pp. 82–96). Springer. https://doi.org/10.1007/978-3-319-66610-5_7
- Lawson, C., Beer, C., Rossi, D., Moore, T., & Fleming, J. (2016). Identification of ‘at risk’ students using learning analytics: The ethical dilemmas of intervention strategies in a higher education institution. *Educational Technology Research and Development*, 64(5), 957–968. <https://doi.org/10.1007/s11423-016-9459-0>

- Marzouk, Z., Rakovic, M., Liaqat, A., J., Vytasek, J., Samadi, D., Stewart-Alonso, J., Ram, I., Woloshen, S., Winne, P.H., and Nesbit, J.C (2016). What if learning analytics were based on learning science? *Australasian Journal of Educational Technology*, 32(6), 1–18. <https://doi.org/10.14742/ajet.3058>
- National Health and Medical Research Council. (2019). *Ethics and integrity*. <https://www.nhmrc.gov.au/research-policy/ethics-and-integrity>
- Newman, T., Beetham, H., & Knight, S. (2019) *Digital experience insights survey 2018: Findings from Australia and New Zealand university students*. Jisc. http://repository.jisc.ac.uk/6967/1/Digital_experience_insights_survey_2018.pdf
- O'Brien, K. (1993). Improving survey questionnaires through focus groups. In D. L. Morgan (Eds.), *Successful focus groups: Advancing the state of the art* (pp. 105–117). Sage. <https://doi.org/10.4135/9781483349008.n7>
- Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438–450. <https://doi.org/10.1111/bjet.12152>
- Parkes, S., Benkowitz, A., Bardy, H., Myler, K., & Peters, J. (2020). Being more human: rooting learning analytics through resistance and reconnection with the values of higher education. *Higher Education Research & Development*, 39(1), 113–126. <https://doi.org/10.1080/07294360.2019.1677569>
- Prieto-Alvarez, C. G., Martinez-Maldonado, R., & Anderson, T. (2018). Co-designing learning analytics tools with learners. In J. M. Lodge, J. C. Horvath, & L. Corrin (Eds.), *Learning analytics in the classroom: Translating learning analytics research for teachers* (pp. 93–110). Taylor & Francis Group. <https://doi.org/10.4324/9781351113038-7>
- Prinsloo, P. & Slade, S. (2014). Educational triage in open distance learning: Walking a moral tightrope. *The International Review of Research in Open and Distance Education*, 15(4) 307–331. <https://doi.org/10.19173/irrodl.v15i4.1881>
- Reimers, G., & Neovesky, A. (2015). Student focused dashboards: An analysis of current student dashboards and what students really want. In M. Helfert, M. T. Restivo, S. Zvacek, & J. Uhomoihi (Eds.), *Proceedings of the 7th International Conference on Computer Supported Education* (Vol. 1, pp. 399–404). Science and Technology Publications. <https://doi.org/10.5220/0005475103990404>
- Roberts, L. D., Chang, V., & Gibson, D. (2016). Ethical considerations in adopting a university- and system-wide approach to data and learning analytics. In B. Kei Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 89–108). Springer. https://doi.org/10.1007/978-3-319-06520-5_7
- 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*, 9(1959), 1–11. <https://doi.org/10.3389/fpsyg.2016.01959>
- Scholes, V. (2016). The ethics of using learning analytics to categorise students on risk. *Educational Technology Research and Development*, 64(5), 939–955. <https://doi.org/10.1007/s11423-016-9458-1>
- Schumacher, C., & Ifenthaler, D. (2018). Features students really expect from learning analytics. *Computers in Human Behavior*, 78, 397–407. <https://doi.org/10.1016/j.chb.2017.06.030>
- Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1509–1528. <https://doi.org/10.1177/0002764213479366>
- Slade, S., & Tait, A. (2019). *Global guidelines: Ethics in learning analytics*. International Council for Open and Distance Education. <https://www.icde.org/s/Global-guidelines-for-Ethics-in-Learning-Analytics-Web-ready-March-2019.pdf>
- Teasley, S., & Whitmer, J. (n.d.). *The impact of student-facing LMS dashboards*. Blackboard. <https://www.blackboard.com/en-uk/resources/impact-student-facing-lms-dashboards>
- West, D., Luzeckyj, A., Searle, B., Toohey, D., & Price, R. (2018). *The use of learning analytics to support improvements in teaching practice*. Innovative Research Universities. <https://www.iru.edu.au/wp-content/uploads/2018/04/MRUN-Learning-Analytics-report.pdf>
- West, D., Luzeckyj, A., Toohey, D., Vanderlelie, J. J., Bell, K. R., & Searle, B. K. (2020). Do academics and university administrators really know better? The ethics of positioning student perspectives in learning analytics. *Australian Journal of Educational Technology*, 36(2), 60–70. <https://doi.org/10.14742/ajet.4653>
- West, D., Tasir, Z., Luzeckyj, A., Si Na, K., Toohey, D., Abdullah, Z., Searle, B., Jumaat, N. F., & Price, R. (2018). Learning analytics experience among academics in Australia and Malaysia: A comparison. *Australasian Journal of Educational Technology*, 34(3), 122–139. <https://doi.org/10.14742/ajet.3836>
- Wolff, B., Knodel, J., & Sittitrai, W. (1993). Focus groups and surveys as complementary research methods: A case example. In D. L. Morgan (Ed.), *Successful focus groups: Advancing the state of the art* (pp. 118–136). SAGE. <https://doi.org/10.4135/9781483349008.n8>

Wook, M., Yusof, Z. M., & Nazri, M. Z. A. (2016). Educational data mining acceptance among undergraduate students. In *Educational and Information Technologies*, 22(3), 1195–1216. <https://doi.org/10.1007/s10639-016-9485-x>

Corresponding author: Deborah West, Deborah.West@flinders.edu.au

Copyright: Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence (CC BY-NC-ND 4.0). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.

Please cite as: West, D., Luzeckyj, A., Searle, B., Toohey, D., Vanderlelie, J., & Bell, K. R. (2020). Perspectives from the stakeholder: Students' views regarding learning analytics and data collection. *Australasian Journal of Educational Technology*, 36(6), 72-88. <https://doi.org/10.14742/ajet.5957>