Learning in the Panopticon: Examining the Potential Impacts of AI Monitoring on Students

Learning in the Panopticon

Examining the Potential Impacts of AI Monitoring on Students

Bingyi, Han*

School of Computing and Information Systems, The University of Melbourne, bingyih@student.unimelb.edu.au

George Buchanan

iSchool, University of Melbourne, george.buchanan@unimelb.edu.au

Dana McKay

School of Computing Technologies, RMIT University, dana.mckay@rmit.edu.au

In a panopticon, people are intrusively monitored across all areas of their lives. AI monitoring has been ever more widely adopted in education, with increasingly intrusive monitoring of students. These changes potentially create ethical harms, but current ethical discussions predominantly focus on legal and governance issues. The concerns of the majority of users—namely students—are neglected. Overlooking students' concerns further increases their vulnerability. We use a student-centred and speculative approach through the Story Completion Method (SCM) to explore how students would potentially respond to intrusive AI monitoring in a higher education setting. Our study included 71 participants who elaborated on the story stems we provided to them. Through a blending of thematic analysis coding and the techniques of developing grounded theory, we reveal that the common responses of students to extensive AI monitoring included impacts on personal psychology, changed behaviour, and cognition. There are likely major disruptions to personal autonomy, identity and educational relationships. If we are to avoid a future 'big brother' classroom, further investigations using HCI methods are critical to understanding how to protect students in AI-dominated learning.

CCS CONCEPTS • Human-centered computing → User studies; •Human-centered computing → Human computer interaction (HCI); Applied computing → Education → Computer assisted instruction

Additional Keywords and Phrases: AI Monitoring, Learning Analytics, Surveillance, Privacy, Story Completion Method, AI Ethics

ACM Reference Format:

First Author's Name, Initials, and Last Name, Second Author's Name, Initials, and Last Name, and Third Author's Name, Initials, and Last Name. 2018. The Title of the Paper: ACM Conference Proceedings Manuscript Submission Template.

* Place the footnote text for the author (if applicable) here.

1 INTRODUCTION

The educational landscape has been shaped by the increasing adoption of AI [1]. In the conventional education system, educators evaluate students' performance by observing and monitoring their in-class behaviours and activities [2], then reflect on learning suggestions and adjust teaching strategies [3]. The introduction of AI in Education (AIEd), especially the use of automated Learning Analytics, results in constant observation of students (by e.g. close real-time monitoring [4]) and the ability to collect learning artefacts and evaluate them at scale [5]. The extent of this monitoring would not be possible with human effort alone. The AIEd approach potentially facilitates more frequent formative assessment and predictions [6] or more insight into students' learning, for example, successful learning patterns [7]. Those benefits rely heavily on the extensive data that is pervasively collected and processed by AI during students' educational practices, which raises ethical concerns over the heightened surveillance and privacy issues [8, 9]. However, the ethical consequences from the lens of students are little understood [10].

A significant gap remains in the field: existing ethical discussions are mostly theoretical [11] and focus on governance [12]. While governance has been considered, the students affected by AIEd have been neither consulted nor considered: existing work is not student-centric [10]. There is also an absence of qualitative studies to understand the social-technical factors of AIEd's design [13] and empirical understanding of the ethical implications of AIEd's use [14]. An ethical discussion informed and led by the HCI community is missing. In pursuing the positive potential of emerging educational technology, the unintended (potentially negative) consequences are often overlooked [15]. HCI researchers are ideally placed to understand the potential negative effects of AIEd, particularly in vulnerable populations. HCI research can also advocate for a more fair and ethical approach. The HCI community can significantly contribute by providing a foundational understanding of the ethical problems by focusing on students' needs and concerns.

Within AIEd, AI monitoring technologies have been increasingly adopted in school sectors, from campus security to in-class behaviour management [8]. Techniques are becoming more detailed and intrusive, from facial recognition to check attendance [16] to affective recognition of emotions such as boredom, confusion, frustration, or delight, to in turn detecting students' (dis)engagement during learning [17]. AIEd has long been investigating facial expressions and affective analysis to detect signs of learning engagement, especially in the higher education [8]. Currently, there are no mature applications yet in the market but it remains a goal for both academia and industry [17, 18]. What is unclear is the degree to which the possible ethical risks of this approach are understood. Typical 'perpetual beta' approaches, where the negative consequences of technology are considered only post-implementation are unethical and ineffective in sensitive settings such as education. Hence, rather than ask about current existing experience, we will take a further step by asking about the "might-become" impacts in the context of "what if" (when future monitoring AI comes to students) by using the Story Completion Method in a new study.

The Story Completion Method (SCM) leverages speculative design [19, 20] and research fiction [21, 22]. One of the significant threads of speculative design is to explore the ethical and moral dilemmas of emerging technology [19]. As a speculative method, the SCM has been specifically used as a critical tool to question the status and development of the existing or future technology [23]. In Story Completion Tasks (SCT), participants are invited to create narratives, which present their socially constructed ideas [24, 25]. Those fictions help speak out about the potential implications and social consequences of adopting a technology [26].

We collected 71 stories created by our participants with two story stem openings that describe different scenarios of using AI monitoring in the classroom, to see how students would likely react and respond to the AI during learning. Specifically, our focus is the role of AI monitoring in tertiary education. We then analysed those stories qualitatively. We revealed several unintended consequences of adopting AI monitoring in education, such as negative emotional/psychological impacts, behavioural changes in response to the presence of such systems, and influenced cognitions. Understanding those implications for individual students and their attitudes is key for the further design of safer educational environments. We summarise the contributions of this paper as: 1) A critical investigation of monitoring AI's impact on students from a user-centred HCI perspective; 2) Revealing the potential interactions among the impacts of AI monitoring, such as how emotional responses might lead to students' behavioural changes and disrupted cognition; 3) An initial study of the consequential ethics of the deployment of AI in Education.

2 RELATED WORK

In this section, we will review some existing monitoring educational technologies both developed for the research use and the market, followed by previous discussions around the impact of those technologies, to better understand the gap of lacking students' perspectives on AI monitoring's adoption and its implications.

2.1 Monitoring Technologies used in Education

Several monitoring educational technologies have been recently developed both in the market and for research. For instance, Holstein [27] developed the Spatial Classroom Log Explorer (SPACLE), a prototype tool that produces a moment-by-moment analysis of students' and teachers' behaviours within a particular learning context. Their study with SPACLE suggests that students' awareness of being monitored by teachers will promote learning; fewer "off-task" or "gaming" behaviours were shown. After SPACLE, Lumilo was developed by the same research group [28]. Lumilo is a mixed-reality tool provided for teachers, which monitors and evaluates students' in-class performance by displaying the real-time indicators of students' metacognitive and behavioural learning process. It then offers customised suggestions of intervention for teachers. It demonstrated that this real-time analytics for teachers can support students' learning by integrating human and machine intelligence [28, 29]. Apart from these research-focused tools, there are several other AIEd applications in the market that some schools have implemented. One such popular tool is the data-driven behavioural management system used in classroom settings called ClassDojo [30], another is the virtual school software called Class by Intel [31]. Both primary affordances include tracking and categorising students' in-class behaviours and reporting to teachers, or providing immediate feedback to students [30]. These systems aim to support personalization, reflection, recommendation and intervention [32] in learning and teaching.

Some AIEd applications, specifically detect and quantify students' facial expressions and possible emotional status. These applications aim to increase teachers' awareness of students' emotional states to help teachers reflect on their teaching and promote positive learning outcomes for students. Emotion detection applications include the Intel Class [31], EMODA [33] and EMODASH [18]. Except for those emotional detection techniques which are mainly provided for teachers, it has also been raised that emotion analysis could be offered to students as a self-regulated learning tool. The findings of detection (such as students' behaviour, interaction, and affective status) can encourage self-reflection [34-36] and motivate positive changes [37, 38] during their learning process.

Those monitoring tools are often investigated for intended positive outcomes such as enhancing students' learning [28] and adjusting educators' teaching strategies [39] by providing personalization and intervention. However, unintended impacts caused by the AI monitoring do potentially exist. Negative concerns have been lightly discussed in previous research that report AIEd developments. For instance, privacy issues have been mentioned with the monitoring tools which detect cheating behaviours in online tests [40]. However, privacy might not be the only ethical impact. Another recent CSCW study investigated the positive support for learning and teaching by tracking students' real-time understanding and sharing for both students and instructors [41]. In their future work, they note how this will impact interactions or discomfort that need further investigation. The following subsection will present existing ethical discussions about monitoring AI.

2.2 Impacts of Monitoring Technology on Education

Some studies aim to explore the impacts of AI monitoring tools across multidisciplinary fields involving not only HCI but also education and pedagogy. For instance, pedagogical studies raised concerns about the datafication of education [42, 43], which leads to changed relationships and role reconstruction within the educational environment [44]. Another example is concerns about quantification of students' behaviours and in-class performance, which encourages obedience in the classroom practices [9] and decreases learning autonomy [45]. Arguments have also been raised around other widespread ethical risks from the technological perspective such as data management and invaded privacy [15, 46], algorithmic fairness and bias issues [11, 47, 48], and labelling and stereotyping learners [7, 49]. To this end, several frameworks aim at mitigating the ethical harm to students, such as Slade and Prinsloo's framework [50], Sclater's Code of Practice [51], and Southgate's The Education, Ethics and AI (EEAI) framework [52]. However, most of those discussions are theoretical [11] or from a governance perspective [12] with a top-down approach. They still lack actual users' perspectives and thus understanding of the socio-technical implications [14] of monitoring AIEds.

Several recent empirical studies from HCI such as Lu's [47], involve teachers in investigating the unintended effects resulting from the use of a classroom surveillance system, ClassDojo, which monitors and quantifies students' in-class performance and reports to teachers. Their interview study with teachers reveals the negative psychological impacts on children which come from the monitoring technologies' measuring, codifying, and simplifying the psycho-social factors existing in the educational environment. The same research group of Lu's [53] has also revealed that the monitoring AI has challenged teachers as they also feel surveilled by parents, students, and school administrators. Consequently, teachers had to be careful of using students' data, and negotiate care and control of those data. They have to balance the optimistic hopes that students can be better cared for (e.g., emotional issues), with ethical dilemmas such as intrusion. This series of unintended consequences might disrupt the pedagogical relationships in the conventional education landscape. However, this work is explicitly from the teachers' perspective, even though those studies aimed at revealing the impact on students. Understanding based on students' actual responses is still scant.

Some empirical research on AIEd involves students but the objectives are students' functional preferences for the AIEd tool design and implementation, rather than ethical impacts [54, 55].

There are some studies investigating ethical considerations from the students' perspective., but these are often focused on a specific known issue. One such issue is privacy, which is the main or even only targeted problem in some studies [15, 56]. Similarly, Tasi et al. [57] explore students' expectations of privacy, however, they mainly focus on the technical perspective of ethical issues such as transparency and information asymmetry. Hence, there remains a lack of students' perspectives in exploring the impact of monitoring AI.

We have summarised the literature on AIEd adoption. The spread of monitoring technologies is growing, and these technologies are increasingly intrusive. Some governance frameworks focusing on issues such as fairness have attempted to address the use of AIEd, but these have been highly theoretical. Some behavioural changes, disruptions to relationships, and negative consequences as a result of AIEd have already been detected in studies of teachers. Ultimately, though most of the impacts of the AIEd adoption, both positive and negative, are directly on students, and studies of students' experiences of monitoring technology have been highly focused on privacy, to the exclusion of other concerns. A lack of "evidence of impact" [58] especially from the perspective of learners increases the difficulty of understanding and preventing ethical harm, which is also seen as a critical barrier to buy-in for AIEd adoption [59]. To address this, discussions of AIEd must adopt consequential ethics, an understanding of the ethical issues that arise as the result of the adoption of these technologies. However, technical artifacts of AI are often perceived as isolated in the ethical discussion while the consideration of the broader contexts in which AI systems are embedded is lacking [60]. We cannot use a framework of consequential ethics until the consequences surface. Hence, we conducted a speculative study to explore the possible consequential ethical risks of adopting monitoring AI that might come to students.

3 RESEARCH METHOD

The purpose of this study is to explore the possible impacts of the prospective AI monitoring applications on students in tertiary education. We focus on facial recognition and emotion detection techniques which are most ambitiously envisioned by both academia and industry [17, 18]. We hope to bring the future prospective AIEd scenarios to the current table of ethical discussions, and thus prevent ethical harm to students before it happens. Hence, we adopted the Story Completion Method (SCM) to collect data by inviting participants to co-create stories with the given Story Completion Tasks (SCT), where those stories are informed by their social and cultural knowledge or experience. This research approach was approved by the Human Research Ethics Committee at the University of Melbourne.

One criticism of this method is that SCM cannot inform us of actual psychological states and should be only considered as stories [24]. So, we will use SCM to investigate commonly held assumptions from the perspective of social constructivism and perceive our results as a "might be" instead of a "will be". By conducting this speculative study under dramatic futuristic scenarios of the classroom with AI monitoring, we will reveal the socially constructed visions of its impacts on students. Recent HCI studies have effectively used SCM for exploring emerging issues in sensitive topics [61] of technology design such as VR pornography [25] and sex robots [62]. In the school context, student groups comprise a vast population with a complex variety, such as sexual identity, race, and a range of cognitive abilities. We thus perceive SCM as an inclusive approach to address potential sensitivities, as the ethical concerns among cultural or identity groups vary.

The SCTs we crafted (see Table 1) constrained the participants to the context of our research interests [63]. Specifically, we focused on the invasive adoption of AI monitoring techniques in university classrooms. We left other details deliberately ambiguous such as the explanation of mentioned AIEd artifacts, functions, or concepts. This deliberate vagueness is a common tactic in SCM to motivate participants to cast their own interpretation and imagination [64]. Unlike other former HCI studies exploring future technology design (e.g. [25, 62]) which used story openings with only one or two sentences, our SCTs offer more immersive scenarios. We provided more descriptions to help participants understand the story context. A similar example of providing immersion through SCT can be seen in a recent HCI research [61] which extends the SCM with interactive storytelling gamification. At the end of the SCT, we prompted our participants by asking them what would happen next and how the character would think, feel, and behave.

Table 1. Story Completion Tasks

Scenario 1: During this summer break, Charlie's university has implemented an intelligent AI System with an education system which will assist teaching. It is the first lecture of this semester. Charlie walks into the classroom. He heard his classmates whispering: There will be an In-Class Surveillance by using the Classroom Behaviour Analysis System: - When students walk into the classroom, the system records their attendance by Facial Recognition Technology. It will then be generated as engagement data stored in Classroom Attendance System. - During the class, the system will capture students' affective (emotional) states, such as when a student is either confused or confident, using an Emotion Recognition Technology. Then the analysis report will be created for teachers to understand students' emotional readiness and cognitive needs for their learning. - Each student's learning material and feedback will be generated via real-time analysis in class, to build a profile of the student's progress in each class. Across a term, this data will generate a personal "Education Profile". Now the class bell rings, Charlie feels …	**Scenario 2:** The lecturer said: "let's have a group discussion". Charlie then heard another whisper from his classmates: During the in-class groupwork practice, the system will also identify what each student is doing and saying, by using Facial Recognition and Voice Recognition. Eye Tracking Technology will be used to capture the amount of time that each student is focused on each learning resource. Then the analysis report will be created by the Eye Tracking for the teachers to understand the learning progress of the class and the efficacy of the chosen learning recourses. Now the classroom becomes noisy and full of conversations, Charlie …

We recruited 71 participants who completed both story completion tasks, along with a brief demographic survey that recorded gender, age, cultural background, and whether they had any teaching experience. Participants included current university students, and recent graduates. We recruited our participants from two cultural backgrounds: Chinese and Australian. We did this to understand whether there are cultural differences when perceiving ethical concepts or harm, because these two cultures demonstrably have varied cultural perceptions of, e.g., privacy [65]. We further split the Chinese group into students who only have only experience of studying domestically (within China), and students who have more than one year's experience studying abroad. We did this to understand whether lived experiences in different cultures can influence people's perceptions of ethical concepts. Finally, participants with teaching experiences were singled out in the lab, as we are curious whether having teaching experiences will differ their insights into learning with AI. In total, there are 25 participants from the Chinese cohort, 25 from the cohort of Chinese with international living experience, and 15 from the Australian cohort. 23 participants had teaching experience or are currently working as teachers. We translated the Chinese stories to English before data analysis. Each participant is labelled with their cultural background (CHN for Chinese, IECHN for Chinese participants who studied abroad, and AUS for Australian participants), a number, and the suffix T if they have teaching experience, so CHN1-T is Chinese participant 1, who has teaching experience.

We firstly conducted the Thematic Analysis (TA) based on the participants-created stories, where we identified the lower-level codes and higher-level themes. In addition, we realised complex relationships exist between various codes and themes: many of them interplay or interact with each other, within and among different themes. We thus found that TA did not provide explicit support for this analysis. We therefore decided to adopt part of Grounded Theory Methodology (GT). Grounded Theory provides an exploratory method that enables the identification of relationships among concepts [66]. Building upon qualitative data sets [67], Grounded Theory offers a structure of three major coding stages: Open Coding, Axial Coding, and Selective Coding in the analysis procedure, which helps with identifying the concepts, high-level phenomena, and interpretation. We translated the initial codes from TA to the open codes for GT. By following the GT analysis structure, we further identified the Causal Condition (e.g., a specific function - surveillance), the Phenomenon (e.g., the affected emotions of students), the Objective (e.g., to gain good performance), the Action and Interactive Strategy (e.g., behavioural change), and the Consequences (e.g., the affected learning experiences). Then we recognised the Process Relationship and Cause-Effect Relationship among those identified impacts due to the AI monitoring adoption. The analysis will be further demonstrated in the following sections.

4 FINDINGS: IMPACTS OF AI MONITORING IN EDUCATION

Through thematic analysis of the collected 71 stories, we identified the unintended social-technical and ethical impacts of AI monitoring on various elements in the educational system. Affected parts of the system include learners, their learning status (e.g., learning autonomy), pedagogical roles and relationships, learning environments and experiences. So far, there is not enough evidence to indicate significant cultural differences in participants' attitudes toward AI monitoring adoption, so this is not further discussed in this paper. This paper will mainly present the impacts on individual students due to their direct interaction with AI monitoring, including the affected personal psychology, changes in behaviour and disrupted cognition. It was clear that those impacts entangled with each other, for instance, the impact on students' emotions led to behavioural responses. So, by further conducting the analysis following the techniques of constructing Grounded Theory, we finally unfold the consequential relationships between those identified codes, as shown in Fig.1 (See Section 4.4).

Table 2: Codes developed based on participants' responses

<table>
<tr><th>Core Category
Axial Code</th><th colspan="4">Sub-Category
Open Code (Mentioned Frequency)</th></tr>
<tr><td colspan="5">IMPACTS ON INDIVIDUALS</td></tr>
<tr><td rowspan="2">Behaviour Change</td><td>Cheat the surveillance (25)</td><td>Pander to the surveillance mechanism (16)</td><td>Elude surveillance - privacy seeking (23)</td><td>Rebellion against the system (19)</td></tr>
<tr><td>Forced performance (15)</td><td>Decreased social interactions (7)</td><td>Unauthentic behaviours: "Not being his usual self" (7)</td><td>Have to be cautious of in-class behaviours (5)</td></tr>
<tr><td rowspan="2">Affected Personal Psychology</td><td colspan="3" rowspan="2">Personal Discomfort
- Anxiety results in being monitored (21)
- Need time to be away and off (21)
- Pressure results in being constantly evaluated (15)
- Fear of bad performance being recorded (6)</td><td>Sense of uncertainty (9)</td></tr>
<tr><td>Unconfident (9)</td></tr>
<tr><td>Disturbed Cognition</td><td colspan="2">Sense of Disturbance:
- On individual identity and autonomy (12)
- On social interactions (7)</td><td>Sense of Disturbance: Invaded Privacy (30)</td><td>Sense of Disturbance: Concentration on study (27)</td></tr>
<tr><td colspan="5">IMPACTS ON PEDAGOGICAL ELEMENTS</td></tr>
<tr><td rowspan="2">Affected Learning Environment and Learning Experience</td><td colspan="2">A scary learning environment
(Students generally live with fear) (22)</td><td>Rules and supervisions are too strict (18)</td><td>More efforts are needed for learning (6)</td></tr>
<tr><td>More engaged class (13)</td><td>Educational disruption (27)</td><td colspan="2">Losing enjoyment and fun of learning experiences (6)</td></tr>
</table>

4.1 Influence of AI Monitoring on Personal Feelings and Emotions

As we did our first inductive coding, we saw there is a consistent focus by most participants on the predominantly negative emotional effects of constant surveillance and evaluation. We will discuss these speculative emotional responses in detail in this section.

Personal Discomfort

The most frequently mentioned emotional response is the strong personal discomfort. For instance, being monitored and constantly evaluated make students feel “apprehensive” (AUS10-T), “stressed and under pressure” (AUS3-T), “nervous” (CHN22-T), “scared” (IECHN12), “fear” (AUS15-T), “irritable” (IECHN17), “overwhelmed” (CHN17), “anxious” (AUS14-T), “self-consciousness” (AUS10), “worried” (AUS13-T), “panic” (AUS2), “insecure for lack of privacy” (CHN2), “as he knows Big Brother is watching” (AUS3-T), or “cannot relax at all once stepping into the classroom with this surveillance” (IECHN11).

Lack Of Transparency Leads to Sense of Uncertainty and a Fear of Showing Poor Performance

Under the surveillance, the student “fears his poor performance being recorded” (CHN2) and worries if "the system “interpreted or labelled his behaviour as negative” (IECHN8-T). It might be due to the system’s lack of transparency as we did not explicitly give the participants a detailed description of the monitoring and evaluation mechanism. Consequently, this opaque system provides the context which leads to the student’s sense of uncertainty. This sense of uncertainty includes the uncertain boundary and scope of monitoring, and the uncertain influence of evaluation. These uncertainties further construct the unclear standard of how students should behave in this learning environment with a monitoring AI. This feeling of uncertainty raises stronger personal discomfort such as worries, which further bring suspicion towards both surveillance and evaluation:

> “The lecturers have told the students to just ignore it, but Charlie knows that he will not be able to do that. He feels uneasy knowing that his every move may be captured and calculated by the class behaviour analysis system. What exactly is it looking for? Should he be writing notes constantly? Should he be nodding along even when the teacher is not looking at him?” (AUS4-T)
>
> “He feels really confused and taken aback that he had no idea about the system before beginning the class. He starts to feel angry that the university wasn't transparent about how they were going to run the class. He worries that if he says the "wrong" thing it will go on his record and it will impact his grades. Charlie doesn't know what to do - should he ask the lecturer more about the system? Should he lodge a complaint? He worries that his expression is being recorded and that his anger and anxiety will show through. He wonders what the consequence of that could be.” (AUS14-T)

Uncertainty and a New Surveillance Panopticon

It is worth noting how other consequences chain from the oft-mentioned sense of uncertainty described above. Once non-transparent surveillance with unquantified rules and judgement mechanisms is implemented, a power distance exists between the individual student and the system. As a result, the sense of uncertainty leads some individuals to subconsciously assume particularly stringent requirements alongside a high standard for behaviour. Participants described assuming that “the monitoring is ubiquitous” (CHN6) and “eyes staring at me all the time” (IECHN12) where the surveillance catches every slack moment in class and “being distracted (or loaf on the study) is not allowed” (IECHN1). This psychological process was seen in many stories.

Consequentially, students considered the classroom as an intense and strict learning environment with “a scary atmosphere” (CHN7) where individuals “generally live with fear, the fear of poor performance being recorded” (IECHN23). For this reason, the students believe more efforts are needed for learning under this system where he/she “must be constantly switched on” (AUS6-T), or “to engage more actively” (AUS12), or “to act as a perfect self” (IECHN2), to ensure his/her performances meet the imagined required evaluation standard. Unsurprisingly, this constant performance leads to personal discomfort. This discomfort is described in a large number of stories, for example, the participant described the environment as “too stressful to be a good student” (IECHN1); or said, “it is tiring as it seems not allowed to be away and off with the surveillance… even tigers sometimes doze” (CHN9-T).

The following story clearly shows a student’s assumption of strict surveillance where being distracted is not allowed:

> “He feels uncomfortable that every word and deed of his own will be under surveillance…Thinking of this, Charlie suddenly realised that the scene of his wandering mind just now should be recorded by the surveillance. He then stopped thinking and continued to listen to the class.” (CHN3)

Two stories show the student has to always perform to avoid poor performance scores:

> “I would feel very nervous because I feel like all my actions are being monitored and recorded and my every move affects my grades. So I would try to behave in an unnatural way/ in a way that I think will give me high marks from the perspective of the machine…Now with the additional voice recognition and eye tracking, I would feel even more nervous since not paying 100% attention at all times will be tracked, and therefore can impact my grades.” (AUS5)
>
> “Charlie is worried… he can't let his guard down and must be constantly "switched on". If he is labelled as disinterested or if he has a bad day he may be called up and reprimanded. If he is performing poorly his teacher may assign more onto his workload and stress him out.” (AUS6-T)

Uncertainty Decreases Confidence During Learning

In addition to the consequence of the uncertainty demonstrated above, as this feeling of uncertain leads to the student assuming the strict rules with high standards, and assuming that the system compares all individuals. As a result, the student becomes less confident in his/her in-class performance, as described in many stories. Furthermore, this fear of poor performance being recorded seems to exacerbate confidence problems:

> “As the system recorded his performance, he already felt nervous and worried about seeming dumb in front of his classmates. He is worried if the system shows that he is terrible at this subject and if his peers notice this and treat him differently than others. His mind is racing. He can't concentrate on learning as he is worried about so many things, such as if this will affect his grades, stop him from graduating, or affect his ability to get a job.” (AUS17-T)
>
> “… He is also a little reluctant to participate in answering questions since he fears that if his responses are not high quality, he may look a bit stupid.” (AUS7-T)

Alongside losing self-confidence, a few respondents expressed their worries that their historical recorded poor performance will make teachers lose confidence and raise biased opinions about them:

“Charlie was worried that the poor evaluation scores would stop his teachers from believing him as a good student and not giving him the attention that he used to have.” (IECHN7)

4.2 Behaviour Change and Performativity

We have just discussed one major theme of emotional and psychological responses. Interestingly, we have noticed in many stories that, most of those emotional impacts have later consequences on how individuals react to AI surveillance and evaluation. In this section, we will discuss the second major theme of potential behaviour changes of the students. We will demonstrate examples of the consequential relationships: each is labelled in italics in this section, and we use a narrative approach to show their interplay and similarities.

Most frequently, the student’s *Fear of Poor Performance Being Recorded* associated with their wish to gain good performance scores leads to the student’s various interactive strategies. Holding this objective, mentioned actions include *Cheating/Pandering to the Surveillance Mechanism* with *Performativity* after figuring out how to survive and adapt to it. E.g., the student would be compelled to act: “a smarter student is likely to become an actor and pretend with perfect performances” (IECHN2). More specifically, “everyone deliberately performed as a good student, pretending they understand the new delivered knowledge” (CHN21-T). Other mentioned strategies include “developing an Emotional Self-Check Software… to avoid being identified as a negative emotion” (CHN18), or “create the High-Score-Dictionary” which summarised the “high-frequency words” that help with gaining scores (CHN19). One story also mentioned, “to buy high score notes from seniors as a cheat sheet… to pass the strict evaluation” (AUS2).

There are also *Privacy-Seeking Attempts* to escape from monitoring, such as by speaking another language (IECHN13) or “undetectable accents and dialects” (IECHN8) to elude voice recognition, or by hiding real emotional or facial expressions for getting away from the emotion/facial recognition. Participants also described seeking private spaces for study or practice so the student can “be free to make mistakes” (AUS5) as “those mistakes won’t be tracked and recorded” (IECHN17-T). A more extreme story mentioned that the student would “drop out of this school”, as believing that “learning within this environment is harmful to mental health.” (CHN2)

The practice of *Avoiding Mistakes* commonly led to students *Being Cautious of In-Class Expressions and Behaviours*. Feeling this “self-conscious” (AUS17-T) is often accompanied by discomforts such as “feel more anxious” (AUS17-T). Many stories also expressed a “forced performance” (CHN10-T), especially with more active and positive participation during class:

> “Charlie forced himself to participate in the discussion although he had no idea about the topic. He participated only because he’s being watched, and he wants to behave himself.” (IECHN25)
> “I have to be active in the group discussions but to make sure my expressions are correct.” (CHN8-T)
> Knowing he was being monitored, “he started to pretend to engage more in the class, such as asking questions, taking notes, and wearing a positive emotion on his face. Although this may somehow motivate him to be actively involved in the class, a 2-hour lecture seems quite stressful and leads to some pressure he might not be able to handle.” (IECHN18-T)

In contrast, some stories consider this forced performance as a positive consequence of *Enhancing the In-class Engagement* of students:

"Even Kevin, the naughty boy who is usually the most difficult to engage in class, is actively participating in the discussion and expressing his views…Charlie felt that the classroom behaviour analysis system seemed to play a positive role in promoting learning in the classroom." (CHN21-T)
"He should behave better than what he used to be... It is certainly a good thing that students can be made more engaged with the class and their learning materials." (IECHN14-T)

More often, these behaviour changes especially forced performance and being cautious of self-behaviours, are aware as *Unauthentic Performance*, which is unnaturally different compared to how the student is supposed to behave without the monitoring AI. It is often accompanied by or leads to *Personal Discomforts*, such as feeling self-conscious and panicking:

"Charlie feels quite self-conscious about what his facial expressions may reveal. He tries to keep a 'poker' face all through the lesson. He is also a little reluctant to participate in answering questions since he fears that if his responses are not high quality, he may look a bit stupid. His friend is aware that Charlie is not being his usual self…" (AUS7-T)
He felt scared because of the AI monitored him invasively, "he used to be eloquent and outgoing, now he became taciturn, which was quite wrong compared to usual. (IECHN12)
"It forces me to perform in a certain manner which I may not be comfortable with." (AUS16)

4.3 Disrupted Concentration, Privacy, and Identity

We have seen the speculative emotional and behavioural responses to monitoring AI adoption in the classroom. Beyond that, there are other disturbances to the internal cognitive world of students, such as the threats to concentration on the study, privacy, social relationships, and self-identity. These concerns were not just within one person, but also compounded with the emotional and behavioural changes. In this section, we will reveal those cognitive disruptions in detail.

4.3.1 Educational Disruption

Those earlier-mentioned forced pretending or acting behaviours, as well as being self-conscious, are frequently perceived as "an extra effort" (AUS12) responding to the student's fear of bad performance being recorded. These extra efforts in turn disrupted the student's learning. Many stories more directly pointed out that the existence of monitoring and constant evaluation itself "distracts the student's concentration on studying" (IECHN4) and especially when the student is aware of the monitoring leading to personal discomfort:

"Eventually the emotion tracking will just encourage more people to put effort into faking emotions and distract from focusing on work." (AUS6-T)
He knows that all his behaviours will be recorded and "he is feeling nervous… He struggles with the exercise because he is so distracted..." (AUS17-T)

Moreover, the sense of uncertainty distracts the student away as they have to spend time wondering about the unclear standard, and be cautious of their in-class behaviours:

He is stressed about the group discussion after knowing he will be watched and assessed constantly by the AI… "He is hardly thinking about the question the teacher asked, rather he is concerned about when he should enter the discussion. What if he accidentally cuts someone off? Will the machine interpret that as poor social behaviour? How can it interpret the natural flows and intricacies of conversation? What was the question again? Charlie's head feels like spinning and he tries to bring his thoughts back to the class..." (AUS4-T)

As a result, quite a few respondents believe their learning environment and experience are dramatically influenced by the existence of the surveillance and evaluation system:

> "He must behave carefully to meet the requirements of evaluation. Under this intense monitoring, group discussion loses its enjoyment. The learning experience is no longer relaxing and rewarding because every motion is captured for analysis. He thinks about the new system instead of the learning content… He feels no longer easy, creative to express his ideas as he used to be." (IECHN19-T)

4.3.2 Internalised Disturbances

Alongside learning-related practical influences, many stories also expressed a direct disturbance to individuals' sense of privacy, their social relations, and to minority groups' social identities as a result of surveillance.

Nearly half the participants expressed concerns about violated privacy, along with varying levels of personal discomfort such as feeling "a bit worrisome" (AUS15-T), oppressed, anxious, and "unprecedentedly uneasy" (CHN9-T). Even more, many colourful similes/metaphors describe their psychological response to their invaded privacy. For example: "I feel myself being naked both physically and mentally" with nowhere to hide the mental state and body (CHN9-T); "feel like a captive animal" in front of the surveillance for no privacy (CHN25); whilst the whole environment "sounds like a prison" (IECHN17-T) where students are "studying in a cage, not a university" (IECHN25); "everyone in this classroom is like a white mouse in a laboratory, being observed all the time with performed behavioural analysis" (CHN22-T). These privacy concerns predominantly come from the AI detection of in-class behaviours and emotions.

Concerns about other non-learning-related information about individuals being revealed have also been raised. Most of these worries are related to an individual's social relations, especially the emotional intimacy that happens in schools:

> "He is concerned that he spent some time looking at one of the new female students and thought he might want to get to know her better - did the AI system see that emotion?" (AUS13-T)

Consequently, many participants described their decreased willingness to have informal interactions with other students in the classroom and attempts to hide or avoid the social behaviours from the monitoring. This would in turn influence the social and pedagogical relationship among students:

> Knowing all students are observed, "I wouldn't feel comfortable with discussing with my classmates and sometimes joking around." (IECHN17-T)
> "He avoids eye contact with the friend who teased him before." (AUS7-T)
> "He doesn't interact with his fellow students anymore, in case the system sees this as time-wasting or that he is lacking confidence. He misses that interaction - that's where he learnt a lot." (AUS13-T)

Interestingly, two stories reveal disturbance for individuals from a minority social identity group, such as being homosexual, an identity that can be either public or closeted. For closeted students, stress comes from the fact that the system might capture this secret, and the possibility of being outed becomes a burden to the individual student:

> "Will the AI analysis conclude on his sexual orientation after recording the content he browsed on TikTok during the class?" (CHN15)
>
> "I don't like the feeling of being monitored. Maybe it is because of my fear that the teacher or other students could know my secret beyond my study through monitoring records…Then Charlie suddenly thought of his best friend John. Charlie has a special feeling to John… he is worried about his special feeling toward John will be monitored through his eyes and mood, which will all be captured and recorded. He didn't want to be treated differently by others because of his little secret..." (IECHN7)

It is unethical that people cannot keep things they wish to hide from their groups such as classmates and schools. As respondent CHN16 expressed, "Everyone needs a sense of safety that comes from holding their secrets." However, whether the AI monitor can walk away, or perform with silence and allow those things shrouded in secrecy is yet to be discussed.

4.4 Summary and Overview: Impacts of AI Monitoring and Their Interplay

We abstracted the most frequently mentioned consequences from the stories, including: the function of Surveillance and Evaluation directly impacts the Emotions (Section 4.1); many of those emotional responses further lead to Behavioural Change (Section 4.2), and Cognitive Disruption (Section 4.3). The Learning Environment and Experience and Educational Relationships as a pedagogical impact are also disturbed. There are other pedagogical impacts as further consequences, such as the Status of Learners, the Pedagogical Roles, which will not be discussed in this paper.

Participants' stories have revealed various levels of psychological discomfort due to constant monitoring and evaluation. While on the other hand, the system's lack of transparency leads to the student's sense of uncertainty about the rules, supervision, and behaviour standards. As a result, individuals assume that the learning environment is harsh, with more or extra learning efforts needed for learning. The sense of uncertainty also decreases students' confidence, which further causes behavioural change with frequent performativity. That is, the student panders to the surveillance by pretending or acting, to gain good performance scores. Also, the student fears showing poor/inappropriate performance as the academic scores will be affected. Consequently, they would force themselves to engage during the class, while being cautious of their in-class expressions and behaviours to avoid making mistakes. Generally, inauthentic behaviour or performance has been frequently expressed as a reaction to the monitoring, whilst these behavioural changes also cause personal discomfort. Those behavioural responses under the surveillance disrupt students' concentration on the study and the initiative for learning. Moreover, the concern of invaded privacy includes a disturbance of individual identity and their social relation which all increases the vulnerability of students.

It is clear to be seen that those impacts do not exist standalone, they rather interact with each other which complicates matters further. In another word, one impact is often the result or cause of other impacts. Further, those interactions happen across the individual and pedagogical dimensions, as shown in Fig.1.

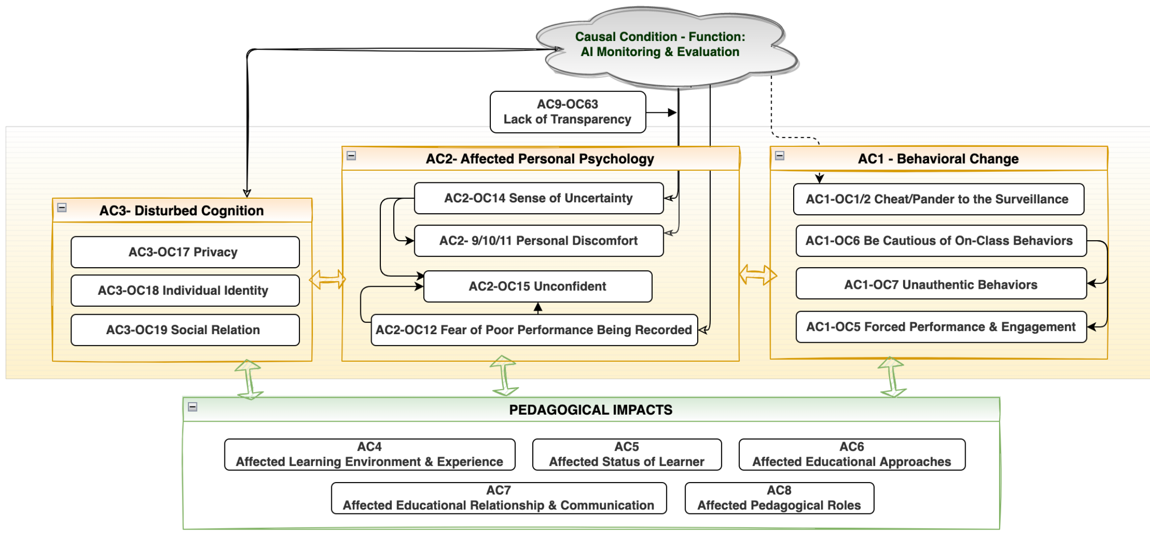

Fig.1. Impacts of AI Monitoring & Constant Evaluation on Individuals

While most emotional and behavioural responses were described by participants as negative, there are still few positive attitudes. For example, several respondents see forced engagement as a way of enhancing learning during class. We have demonstrated both sides of responses from students' perspectives where our focus in this paper is the classroom monitoring setting's direct impact on individuals. In the next section, we will draw on broader discussions around the mentioned concerns.

5 DISCUSSION

As pointed out earlier by Prinsloo [68], surveillance has increased students' vulnerability to privacy invasion and information collection, processing, and dissemination. These concerns have been reflected in our data: constant collection and evaluation of behavioural causes intrusion to privacy and individual identity, and disturbs learning. These concerns all potentially exacerbate the vulnerability of students, which consequently increases the difficulty of protecting students from ethical harm.

In addition to the increased vulnerability, the speculated experience of students touches all three components of the Breckler's [69] ABC Model of Attitude, i.e., Affect, Behaviour and Cognition, reflecting students attitude toward the adoption of AI in the classroom. We have seen the possible negative emotions as affective responses; the unauthentic behavioural changes and acts; and the cognitive beliefs of being invaded and disrupted. These concerns demonstrate the seriousness of AI monitoring, since it potentially affects the student's spirit (or emotion), body, and mind. In the following subsections, we will discuss in detail the implications of those three affected elements: emotional responses, behavioural change, and disrupted cognition. We will also discuss the utility of the Story Completion Method for exploring sensitive topics.

This paper is part of a wider study on how AIEd adoption might impact students' learning. This paper focuses on AI monitoring before we explore further functions such as customised recommendations and intervention. The impacts mentioned are from the perspective of individual students. There are aggregated pedagogical impacts such as the benefits and issues that come from the evaluation of the whole class, and also the support of learning processes. Those topics will be covered in another paper.

5.1 The Emotional Impacts of AI Monitoring

From our findings, recurring personal discomforts come directly from the constant surveillance and evaluation, such as anxiety, fear, pressure, insecurity, and lack of confidence. Similar concerns have also been raised by other earlier research, which shows that constant surveillance used in education will impact students' physical and psychological well-being, e.g. the chilling effect of students' fear of surreptitious monitoring, or amplified performance-related stress [70]. We should be alert to these negative reactions, as these discomforts might change how people receive the adoption of Learning Analytics and how they live with it. Further, most psychological impacts can directly lead to students' behavioural changes.

Students' fear of been seen to perform badly is an important social factor related to self-esteem and fear of failure. The feelings of discouragement and insecurity found in our data contravene requirements identified in earlier work which includes student requests to learn in a safe environment where they could make mistakes without the feeling of fear [46]. Importantly, this fear of acquiring a record of bad performance results from students' sense of uncertainty, which itself is due to a lack of transparency about the system and its operation. In turn, this further intensifies discomforts and leads to further behavioural change. The lack of transparency and the related ethical concerns do not only exist in our stories; they have also been reported in existing real-world studies. For example, Slade and Prinsloo [71] have long raised concerns that a lack of transparent data collection process of educational institutions on students' information potentially leads to ethical harm. The lack of precise transparency in our description of the putative monitoring system is commonplace in users' understanding of such systems in the real world (e.g., [72]).

In a different context, an early empirical study investigated the long-term effects of ubiquitous surveillance in the home [73]. The results revealed that the surveillance may not be potent enough to bring significant stress and negative mental health effects., but it was shown to be a cause of personal discomforts such as annoyance, concern, and anxiety. The reason for these discomforts was tentatively attributed to the violation of privacy functions of the home, though most participants acclimatised to the system within months. In contrast to the home, students are not typically involved when making decisions about monitoring in classroom environments. Students are thus a vulnerable group within the learning environment. We have seen the speculated negative impacts of surveillance on learners' affective states in our context of higher education. We need careful consideration of the psychological implications, of such monitoring as those impacts are entangled with changes in students' behaviour.

5.2 Further behavioural change due to affected personal psychology

There are theories from previous research describing people's behavioural change due to the existence of surveillance systems, such as panopticon theory, chilling effect, and performativity [30]. These have also been seen in our data. In this section, we proved that these effects also happen in different contexts, e.g., more intrusive AI techniques used in education. Furthermore, we revealed how these behavioural changes were raised from the affected personal psychology (emotions).

5.2.1 Surveillance, Panopticon, and Transparency

Foucault [74] has argued that the school functions as a social control mechanism. It governs students where the architecturally designed surveillance is deployed with a disciplinary and hierarchical power relation between students and the educational system [75]. In this way, the governmental mechanism of the education system reflects the panopticon. With the use of AI techniques in education, the information collection and data-based calculations of the surveillance of individuals will be more intrusive and overwhelming, which increases the vulnerability of students [68]. Criticisms from the pedagogical community have argued that the monitoring tools replicate and extend the Foucauldian Panoptic Surveillance on students' in-class behaviour (e.g., [9, 30]).

Our data reflects the panopticon in AIEd that students' sense of uncertainty about the surveillance and behaviour standards leads to their assuming the monitoring is ubiquitous. This is due to the untransparent monitoring mechanism and evaluation operations where an asymmetric information relationship exists between students and the surveillance system. Students' fear of the monitoring and learning environment makes them believe they should put more effort into learning to meet the imagined standard, such as assuming they "must be constantly switched on", It brings further negative psychological impacts such as anxiety and exhaustion.

To mitigate the students' personal discomfort that comes from panoptic surveillance, it is necessary to first take care of students' sense of uncertainty by making the monitoring mechanism and procedures entirely transparent. It must be clear to students what information will be collected, the detailed rules of the evaluation and calculation, whether their performance will be ranked, and the consequences of the evaluation. It has been suggested that providing transparency and avoiding ambiguity in the data collection and utilization process is helpful to building trust [76] and improving students' comfort in sharing data and providing consents [77]. In addition to data use issues, our findings suggest that making monitoring operations more transparent might decrease students' discomfort, mitigating the panopticon effect when being monitored in an educational environment.

5.2.2 Constant Evaluation and Performativity:

Manolev argued that a learning environment where the students' behavioural performance is constantly evaluated creates a culture of performativity as a behavioural control mechanism [30]. Students will constantly be subject to, and participate in calculating and organising their in-class performance, in response to evaluations and benchmarks [78]. This has been seen in our data where students tend to pander to the evaluation mechanism by performing themselves. Consequently, the learning experience under constant evaluation reshapes students, especially their learning initiatives and autonomy.

This culture of performativity leads to various behavioural changes in class. We will not dive deeply into these changes in this paper. However, when students deliberately present themselves in an artificial way, such as forcing themselves to participate and act perfectly, or cheating and pandering to the surveillance and evaluation, this gives rise to adverse behavioural impacts [51] and creates personal discomforts such as the sense of unnaturalness ("not being the usual self"), exhaustion, and anxiety. An ethical argument has been raised that students should be free of the feeling of embarrassment caused by less active participation [79]. However, as seen in our findings, the undesirable behavioural changes and emotional discomforts interact with or even exacerbate each other.

5.2.3 Chilling Effect:

Previous studies argued that monitoring systems have chilling effects on students [43, 80]. That is, surveillance promotes individuals' self-censorship, conformity, and inhibition, which renders them more cautious and careful of their engagement in candid discussion [81-83]. This is consistent with our data including complaints that the student must be cautious of in-class expressions and behaviours, with feelings of discouragement, lack of confidence and increased panic. These behavioural changes and psychological impacts also affect learners' status, such as self-determination and learning autonomy.

Our findings might address the gap in Holstein's experiment [28], which suggested that real-time teacher analytics enhances students' learning. Holstein's study [28] has shown that, when students are monitored by real-time analytics, they are more willing to receive hints, spend less time on unproductive persistence, make fewer mistakes on specific questions, and have more idle time in their learning process. The context of Holstein's study is with school children while our context is in tertiary education. His results echo our findings that students are cautious of their learning practice and avoid making mistakes, as they fear poor performance being recorded by the system. Our results can be seen as a supplement with psychological factors to his findings that students' awareness of being monitored reduces the frequency of maladaptive learning behaviours. However, on the basis of our data we remain concerned about whether all these influences on students are positive, even though Holstein's group perceived the improved learning outcomes as an advance.

We have seen how psychological impacts intertwined with behavioural changes, an interaction to which we must remain alert. These interactions refocus our attention on the rationale for designing and deploying AIEd systems. Participatory design techniques that include learners are necessary to understand negative consequences. We might also need to engage in specific focus on different contexts or lenses toward more vulnerable groups, such as younger students, or students such as LBGTQI students or students of colour for whom extra monitoring will create increased anxiety [43].

5.3 Privacy and the Disturbed Individual Identity

The invasion of privacy by AI monitoring is much discussed in the literature (e.g., [84]). When considering privacy in the educational context, it is necessary to consider the existing relationships, e.g., between student-teacher, student-school, or even younger student-parents. These relationships bring value tensions among different groups which add complexities when addressing privacy concerns, even in other contexts. For instance, early research has shown clashing preferences between parents and teens regarding what data should be included in audit logs on Home-Entryway Surveillance [85]. Teens coming and going will be monitored by their parents and their privacy needs to be negotiated with the parents. However, even though both parents and children agree that children's privacy should be respected, they perceive privacy concepts differently [86]. Those differences which lead to value tensions need to be considered.

Further, the unclear balance between the monitoring's essential caregiving and privacy invasion for child-parent relationship [87] is also reflected in student-school or student-teacher relationships. In addition, there are inherent power differentials between students and teachers, or between students and schools backed by authorities that exist in the educational context. These power differentials have been indicated as one of the features of intimate threats towards privacy invasion, which makes it complicated to ameliorate [88].

We are aware of the privacy-preserving computing technology developed for facial recognition or eye-tracking, which might mitigate privacy concerns by anonymising both data collection and use [89]. However, these techniques may not solve be able to solve the privacy issues seen in the context of AIEd. The ethical use of AIEd tools regarding privacy is in tension with the tool's utility and delivered benefits, especially when providing customised recommendations for individual learning. Those intended outcomes might not be realised without a linkage between individuals and their personal data [90]. Hence, this might be a new field for privacy preserving HCI and should consider students' experiences when trying to realise the intended benefits.

Except for the privacy invasion by the monitoring that has been popularly discussed [56, 91], students feel squeezed in the learning environment as surveillance causes educational disruption. It is similar to another context of surveillance at home that, the monitoring devices, e.g. cameras were perceived as life-disrupting [73].

More sensitively, the ethical concerns among cultural or identity groups vary where unique vulnerability needs to be considered (e.g. for students with disabilities [92]). Our data raises these issues with a few sexuality-related responses. We have mentioned the concern from a minor social identity group that the surveillance might be a burden to disturbing their individual identity and revealing their privacy. A recent article has shown that some school surveillance systems flag LGBTQ keywords from students, which may unintentionally out students' sexuality [93]. This surveillance might be more harmful to certain marginalised groups than others, regarding well-being and safety (abuse issues) for individual students and their families.

School is supposed to be a safe place for all students. We should have inclusive consideration to make AIEd more thoughtful and friendly to all individuals, regardless of race or sexuality. As pointed out earlier how the surveillance horrified LGBTQ students, individuals from different cultural backgrounds or with different identities might differ in the way they perceive the AIEd environment [58, 94]. These differences might also exist in other dimensions of Affective, Behavioural, and Cognitive responses of the ABC model of attitude, while these need careful consideration to prevent ethical harms coming to more vulnerable groups.

5.4 Contribution to the Story Completion Method

Like other recent HCI studies (e.g., [25, 95]) using Story Completion Method (SCM), our findings showed that SCM could probe sensitive topics by constructing fictional responses from nonpersonal storytelling with indirectness [63]. Participants can be more comfortable discussing sensitive topics using this third-person lens than asking them to tell "your" perspective. For instance, our respondents spoke out about sexuality-related concerns smoothly during made up narratives, which are informed by and reflect their experience and socio-cultural values [24].

In contrast to that earlier work in HCI, though, we have not exclusively used thematic analysis to code our stories. Instead, we have followed the techniques for developing Grounded Theory. This different analytical method allowed us to tease out the relationships among codes and themes in our complex dataset. By extending the analytical techniques used, we have demonstrated the potential of SCM to further reveal the complexity of the social-technical implications of merging technology's adoption.

6 FUTURE WORK

Our immediate future plan is to carry on the exploration of AIEd's ethical impacts on students including analysing our data with a pedagogical theme. This includes the behavioural change in detail, the affected learner status, learning environment and relationships, learning approaches and pedagogical roles.

The scenarios of this study focus on the context of higher education where the student groups narrowed to university students. Future studies might involve other age groups such as under-18 teenagers and younger children. We also did not consider stigmatised or marginalised groups within the student community. One potential avenue for future research may be considering the consequences or perceived threats to LGBTIQA+ students from deploying AI monitoring.

Extending the work into the future by repeating this work in the context of a real-world deployment, to validate the perceived likely threats as ones that emerge in actual real-world contexts is still key. As we noted above, some of the reported concerns of students have already been discerned in actual deployments. However, we have found new likely consequences that future developers should be aware of. However, this raises the concern that exposing students to stressful and threatening technologies may be unethical. How to forge a path towards understanding real-world consequences that minimise the risk of real-world harm is itself a vital concern to address as work in AIEd progresses.

7 CONCLUSION

The traditional learning landscape involving teacher and student is increasingly interrupted by a third party: AI. We now need to build a model of how learning works for students including this third "person", and to understand whether and how to safely integrate AI into education. We need to understand the voices and concerns of all the stakeholders, including teachers, school, governance, and especially learners and their family.

It is crucial to understand the possible impact of AIEd on students, to further ensure the ethics provide a workable ground and a useful baseline for all the components of the system [1]. This understanding might include their acceptance or preferences towards technology, how they would negotiate with themselves to balance the benefit and harm that technology brings, how their tolerance would change when facing different cases. We must also consider whether there are any cross-cultural differences, and what the cultural reasons are for these differences. If we can identify the constraints that bring this community together and create a new sustainable model, we will provide a foundation of AI in education and improve the confidence of educators and students to use any future systems.

REFERENCES

[1] Luckin, R., et al., Intelligence unleashed: An argument for AI in education. 2016.

[2] Romero, C. and S. Ventura, Data mining in education. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2013. **3**(1): p. 12-27.

[3] Margolis, A.A., Zone of Proximal Development, Scaffolding and Teaching Practice. Cultural-Historical Psychology, 2020. **16**(3): p. 15-26.

[4] Marcu, G. and A.N. Spiller. Collaborative Aspects of Collecting and Reflecting on Behavioral Data. in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020.

[5] Siemens, G., Learning analytics: The emergence of a discipline. American Behavioral Scientist, 2013. **57**(10): p. 1380-1400.

[6] Bulger, M., Personalized learning: The conversations we're not having. Data and Society, 2016. **22**(1): p. 1-29.

[7] Siemens, G. and P. Long, Penetrating the fog: Analytics in learning and education. EDUCAUSE review, 2011. **46**(5): p. 30.

[8] Andrejevic, M. and N. Selwyn, Facial recognition technology in schools: Critical questions and concerns. Learning, Media and Technology, 2020. **45**(2): p. 115-128.

[9] Williamson, B., Datafication of education: a critical approach to emerging analytics technologies and practices. Rethinking pedagogy for a digital age, 2019: p. 212-226.

[10] Mangaroska, K. and M. Giannakos, Learning analytics for learning design: A systematic literature review of analytics-driven design to enhance learning. IEEE Transactions on Learning Technologies, 2018. **12**(4): p. 516-534.

[11] Arnold, K.E. and N. Sclater. Student perceptions of their privacy in learning analytics applications. in Proceedings of the seventh international learning analytics & knowledge conference. 2017.

[12] Schiff, D., et al. What's next for ai ethics, policy, and governance? a global overview. in Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society. 2020.

[13] Wise, A., Y. Zhao, and S. Hausknecht, Learning analytics for online discussions: Embedded and extracted approaches. Journal of Learning Analytics, 2014. **1**(2): p. 48-71.

[14] Krach, S.K., M.P. McCreery, and H. Rimel, Examining teachers' behavioral management charts: A comparison of Class Dojo and paper-pencil methods. Contemporary School Psychology, 2017. **21**(3): p. 267-275.

[15] Kumar, P.C., et al. Privacy and security considerations for digital technology use in elementary schools. in Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019.

[16] Selwyn, N., Digital lessons? Public opinions on the use of digital technologies in Australian schools. Melbourne, Monash University. Retrieved May, 2019. **10**(2020): p. 2019-01.

[17] Dewan, M., M. Murshed, and F. Lin, Engagement detection in online learning: a review. Smart Learning Environments, 2019. **6**(1): p. 1-20.

[18] Ez-Zaouia, M., A. Tabard, and E. Lavoué, EMODASH: A dashboard supporting retrospective awareness of emotions in online learning. International Journal of Human-Computer Studies, 2020. **139**: p. 102411.

[19] Dunne, A. and F. Raby, Speculative everything: design, fiction, and social dreaming. 2013: MIT press.

[20] Wong, R.Y. and V. Khovanskaya, Speculative design in HCI: from corporate imaginations to critical orientations, in New Directions in Third Wave Human-Computer Interaction: Volume 2-Methodologies. 2018, Springer. p. 175-202.

[21] Blythe, M., et al. Imaginary design workbooks: Constructive criticism and practical provocation. in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018.

[22] Troiano, G.M., J. Tiab, and Y.-K. Lim. SCI-FI: shape-changing interfaces, future interactions. in Proceedings of the 9th Nordic Conference on Human-Computer Interaction. 2016.

[23] Lawson, S., et al. Problematising upstream technology through speculative design: the case of quantified cats and dogs. in Proceedings of the 33rd annual ACM conference on human factors in computing systems. 2015.

[24] Kitzinger, C. and D. Powell, Engendering infidelity: Essentialist and social constructionist readings of a story completion task. Feminism & Psychology, 1995. **5**(3): p. 345-372.

[25] Wood, M., G. Wood, and M. Balaam. " They're Just Tixel Pits, Man" Disputing the'Reality'of Virtual Reality Pornography through the Story Completion Method. in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017.

[26] Lindley, J., P. Coulton, and M. Sturdee, Implications for Adoption, in Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017, Association for Computing Machinery. p. 265–277.

[27] Holstein, K., B.M. McLaren, and V. Aleven. SPACLE: Investigating learning across virtual and physical spaces using spatial replays. in Proceedings of the Seventh International Learning Analytics & Knowledge Conference. 2017.

[28] Holstein, K., B.M. McLaren, and V. Aleven. Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms. 2018. Cham: Springer International Publishing.

[29] Holstein, K., B.M. McLaren, and V. Aleven, Informing the design of teacher awareness tools through causal alignment analysis. 2018, International Society of the Learning Sciences, Inc.[ISLS].

[30] Manolev, J., A. Sullivan, and R. Slee, The datafication of discipline: ClassDojo, surveillance and a performative classroom culture. Learning, Media and Technology, 2019. **44**(1): p. 36-51.

[31] Kaye, K., Intel calls its AI that detects student emotions a teaching tool. Others call it 'morally reprehensible.'. 2022: Protocol.

[32] Khalil, M. and M. Ebner. Learning analytics: principles and constraints. in EdMedia+ Innovate Learning. 2015. Association for the Advancement of Computing in Education (AACE).

[33] Ez-Zaouia, M. and E. Lavoué. EMODA: A tutor oriented multimodal and contextual emotional dashboard. in Proceedings of the seventh international learning analytics & knowledge conference. 2017.

[34] Lim, L., et al. Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference? in Proceedings of the 9th international conference on learning analytics & knowledge. 2019.

[35] Verbert, K., et al., Learning analytics dashboard applications. American Behavioral Scientist, 2013. **57**(10): p. 1500-1509.

[36] Noroozi, O., et al., Multimodal data indicators for capturing cognitive, motivational, and emotional learning processes: A systematic literature review. Education and Information Technologies, 2020. **25**(6): p. 5499-5547.

[37] Park, Y. and I.-H. Jo, Development of the learning analytics dashboard to support students' learning performance. Journal of Universal Computer Science, 2015. **21**(1): p. 110.

[38] Jivet, I., et al., From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. The Internet and Higher Education, 2020. **47**: p. 100758.

[39] Holstein, K., B.M. McLaren, and V. Aleven, Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms. 2019, Springer International Publishing. p. 157-171.

[40] Li, X., et al., Massive Open Online Proctor: Protecting the Credibility of MOOCs certificates, in Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing. 2015, Association for Computing Machinery: Vancouver, BC, Canada. p. 1129–1137.

[41] Sung, G., T. Feng, and B. Schneider, Learners Learn More and Instructors Track Better with Real-time Gaze Sharing. Proc. ACM Hum.-Comput. Interact., 2021. **5**(CSCW1): p. Article 134.

[42] Selwyn, N., Data entry: Towards the critical study of digital data and education. Learning, Media and Technology, 2015. **40**(1): p. 64-82.

[43] Kwet, M. and P. Prinsloo, The 'smart'classroom: a new frontier in the age of the smart university. Teaching in Higher Education, 2020. **25**(4): p. 510-526.

[44] Jarke, J. and A. Breiter, Editorial: the datafication of education. Learning, Media and Technology, 2019. **44**(1): p. 1-6.

[45] Lundie, D., Authority, autonomy and automation: The irreducibility of pedagogy to information transactions. Studies in philosophy and education, 2016. **35**(3): p. 279-291.

[46] Drachsler, H., et al. Ethical and privacy issues in the application of learning analytics. in Proceedings of the fifth international conference on learning analytics and knowledge. 2015.

[47] Lu, A.J., et al., Coding Bias in the Use of Behavior Management Technologies: Uncovering Socio-technical Consequences of Data-driven Surveillance in Classrooms, in Designing Interactive Systems Conference 2021. 2021, Association for Computing Machinery: Virtual Event, USA. p. 508–522.

[48] Durán, J.M. and K.R. Jongsma, Who is afraid of black box algorithms? On the epistemological and ethical basis of trust in medical AI. Journal of Medical Ethics, 2021. **47**(5): p. 329-335.

[49] Lawson, C., et al., Identification of 'at risk'students using learning analytics: the ethical dilemmas of intervention strategies in a higher education institution. Educational Technology Research and Development, 2016. **64**(5): p. 957-968.

[50] Prinsloo, P. and S. Slade, Student vulnerability, agency, and learning analytics: An exploration. Journal of Learning Analytics, 2016. **3**(1): p. 159-182.

[51] Sclater, N., Developing a Code of Practice for Learning Analytics. Journal of Learning Analytics, 2016. **3**(1): p. 16-42.

[52] Southgate, E., et al., Artificial intelligence and emerging technologies in schools. 2019.

[53] Lu, A.J., et al., Data Work in Education: Enacting and Negotiating Care and Control in Teachers' Use of Data-Driven Classroom Surveillance Technology. Proceedings of the ACM on Human-Computer Interaction, 2021. **5**(CSCW2): p. 1-26.

[54] Whitelock‐Wainwright, A., et al., The Student Expectations of Learning Analytics Questionnaire. Journal of Computer Assisted Learning, 2019. **35**(5): p. 633-666.

[55] Whitelock-Wainwright, A., et al., An exploratory latent class analysis of student expectations towards learning analytics services. The Internet and Higher Education, 2021: p. 100818.

[56] Jones, K.M., et al., "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, 2020. **71**(9): p. 1044-1059.

[57] Tsai, Y.-S., A. Whitelock-Wainwright, and D. Gašević, The privacy paradox and its implications for learning analytics, in Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. 2020, Association for Computing Machinery: Frankfurt, Germany. p. 230–239.

[58] Macfadyen, L.P., Institutional Implementation of Learning Analytics: Current State, Challenges, and Guiding Frameworks. by Charles Lang, Alyssa Friend Wise, Agathe Merceron, Dragan Gaševic, and George Siemens. 2nd ed. Vancouver, Canada: SOLAR, 2022.

[59] Tsai, Y.-S., et al., Learning analytics in European higher education—Trends and barriers. Computers & Education, 2020. **155**: p. 103933.

[60] Hagendorff, T., The ethics of AI ethics: An evaluation of guidelines. Minds and Machines, 2020. **30**(1): p. 99-120.

[61] Troiano, G.M., et al. "Are We Now Post-COVID?": Exploring Post-COVID Futures Through a Gamified Story Completion Method. in

Designing Interactive Systems Conference 2021. 2021.

[62] Troiano, G.M., M. Wood, and C. Harteveld. " And This, Kids, Is How I Met Your Mother": Consumerist, Mundane, and Uncanny Futures with Sex Robots. in Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 2020.

[63] Clarke, V., et al., Editorial introduction to the special issue: using story completion methods in qualitative research. 2019, Taylor & Francis. p. 1-20.

[64] Braun, V., et al., Qualitative story completion: Possibilities and potential pitfalls. Qualitative Research in Psychology, 2019. **16**(1): p. 136-155.

[65] Ma, Y., Relational privacy: Where the East and the West could meet. Proceedings of the Association for Information Science and Technology, 2019. **56**(1): p. 196-205.

[66] Walker, D. and F. Myrick, Grounded theory: An exploration of process and procedure. Qualitative health research, 2006. **16**(4): p. 547-559.

[67] Adams, A., P. Lunt, and P. Cairns, A qualititative approach to HCI research. 2008.

[68] Prinsloo, P. and S. Slade, Student Vulnerability, Agency and Learning Analytics: An Exploration. Journal of Learning Analytics, 2016. **3**(1): p. 159-182.

[69] Breckler, S.J., Empirical validation of affect, behavior, and cognition as distinct components of attitude. Journal of personality and social psychology, 1984. **47**(6): p. 1191.

[70] Reidenberg, J.R. and F. Schaub, Achieving big data privacy in education. Theory and Research in Education, 2018. **16**(3): p. 263-279.

[71] Slade, S. and P. Prinsloo, Learning analytics: Ethical issues and dilemmas. American Behavioral Scientist, 2013. **57**(10): p. 1510-1529.

[72] Roberts, L.D., et al., Student attitudes toward learning analytics in higher education:"The fitbit version of the learning world". Frontiers in psychology, 2016. **7**: p. 1959.

[73] Oulasvirta, A., et al., Long-term effects of ubiquitous surveillance in the home, in Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012, Association for Computing Machinery: Pittsburgh, Pennsylvania. p. 41–50.

[74] Foucault, M., Security, territory, population: lectures at the Collège de France, 1977-78. 2007: Springer.

[75] Foucaul, M., Discipline and Punish: the Birth of the Prison. New York: Vintage, 1979.

[76] Corrin, L., P.G. de Barba, and A. Brooker, Understanding the learner perspective to inform institutional learning analytics strategy and practice, in Handbook of Digital Higher Education. 2022, Edward Elgar Publishing. p. 248-259.

[77] Li, W., et al., Disparities in students' propensity to consent to learning analytics. International Journal of Artificial Intelligence in Education, 2021: p. 1-45.

[78] Ball, S.J., The teacher's soul and the terrors of performativity. Journal of education policy, 2003. **18**(2): p. 215-228.

[79] Chen, B. and H. Zhu, Towards Value-Sensitive Learning Analytics Design, in Proceedings of the 9th International Conference on Learning Analytics & Knowledge. 2019, Association for Computing Machinery: Tempe, AZ, USA. p. 343–352.

[80] Kwet, M., Operation Phakisa education: Why a secret? Mass surveillance, inequality, and race in South Africa's emerging national e-Education system. Mass Surveillance, Inequality, and Race in South Africa's Emerging National E-Education System (December 4, 2017), 2017.

[81] Solove, D.J., A taxonomy of privacy. U. Pa. L. Rev., 2005. **154**: p. 477.

[82] Penney, J., Internet surveillance, regulation, and chilling effects online: A comparative case study. Regulation, and Chilling Effects Online: A Comparative Case Study (May 27, 2017), 2017. **6**(2).

[83] Sklansky, D.A., Too much information: How not to think about privacy and the Fourth Amendment. Calif. L. Rev., 2014. **102**: p. 1069.

[84] Alkhatib, S., et al. Privacy and the Internet of Things (IoT) Monitoring Solutions for Older Adults: A. in Connecting the System to Enhance the Practitioner and Consumer Experience in Healthcare: Selected Papers from the 26th Australian National Health Informatics Conference (HIC 2018). 2018. IOS Press.

[85] Ur, B., J. Jung, and S. Schechter. Intruders versus intrusiveness: teens' and parents' perspectives on home-entryway surveillance. in Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014.

[86] Cranor, L.F., et al. {Parents'} and {Teens'} Perspectives on Privacy In a {Technology-Filled} World. in 10th Symposium On Usable Privacy and Security (SOUPS 2014). 2014.

[87] Ghosh, A.K., et al. A matter of control or safety? Examining parental use of technical monitoring apps on teens' mobile devices. in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018.

[88] Levy, K. and B. Schneier, Privacy threats in intimate relationships. Journal of Cybersecurity, 2020. **6**(1): p. tyaa006.

[89] Erkin, Z., et al., Privacy-Preserving Face Recognition. 2009, Springer Berlin Heidelberg. p. 235-253.

[90] Rubel, A. and K.M. Jones, Student privacy in learning analytics: An information ethics perspective. The information society, 2016. **32**(2): p. 143-159.

[91] Stahl, B.C. and D. Wright, Ethics and privacy in AI and big data: Implementing responsible research and innovation. IEEE Security & Privacy, 2018. **16**(3): p. 26-33.

[92] Baek, C. and S.J. Aguilar, Past, present, and future directions of learning analytics research for students with disabilities. Journal of Research on Technology in Education, 2022: p. 1-16.

[93] Carballo, A. Remote Learning Accidentally Introduced a New Danger for LGBTQ Students. 2022 [cited 2022 Jun 14, 2022]; Available from: https://slate.com/technology/2022/02/remote-learning-danger-lgbtq-students.html.

[94] Dawson, S., et al. Rethinking learning analytics adoption through complexity leadership theory. in Proceedings of the 8th international conference on learning analytics and knowledge. 2018.

[95] Kelly, R.M., et al. "It's About Missing Much More Than the People": How Students use Digital Technologies to Alleviate Homesickness. in Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. 2021.