Investigating Students' Experiences with Collaboration Analytics for Remote Group Meetings

Qi Zhou[1], Wannapon Suraworachet[1], Stanislav Pozdniakov[2], Roberto Martinez-Maldonado[2], Tom Bartindale[2], Peter Chen[2], Dan Richardson[2], Mutlu Cukurova[1]

[1] University College London, London, UK
[2] Monash University, Melbourne, VIC, Australia
qtnvqz3@ucl.ac.uk

Abstract. Remote meetings have become the norm for most students learning synchronously at a distance during the ongoing coronavirus pandemic. This has motivated the use of artificial intelligence in education (AIED) solutions to support the teaching and learning practice in these settings. However, the use of such solutions requires new research particularly with regards to the human factors that ultimately shape the future design and implementations. In this paper, we build on the emerging literature on human-centred AIED and explore students' experiences after interacting with a tool that monitors their collaboration in remote meetings (i.e., using Zoom) during 10 weeks. Using the social translucence framework, we probed into the feedback provided by twenty students regarding the design and implementation requirements of the system after their exposure to the tool in their course. The results revealed valuable insights in terms of visibility (what should be made visible to students via the system), awareness (how can this information increase students' understanding of collaboration performance), and accountability (to what extent students take responsibility of changing their behaviours based on the system's feedback); as well as the ethical and privacy aspects related to the use of collaboration analytics tools in remote meetings. This study provides key suggestions for the future design and implementations of AIED systems for remote meetings in educational settings.

Keywords: human-centred AI, remote meetings, collaboration analytics, ethics

1 Introduction and Background

There is an increasing amount of research that shows the positive impact of using Artificial Intelligence (AI) applications to support students' academic performance [1, 2], their affective engagement [3–5], and metacognitive development [6–8]. In the design of effective AI in Education (AIED) tools, most available research highlights the significance of robust technical approaches and the use of learning sciences principles [9, 10]. However, a range of other human factors related to AIED tools are often neglected, including students' preferences, why and how the tools will be used [11], the social contexts in which the tools will be used, and ethical [12] and societal implications related to fairness, accountability and transparency [13]. Understanding how human factors (i.e. the characteristics of students, educators, other relevant stakeholder and the environment) can shape the use of AIED tools is key for their successful adoption and the field's wider impact on Education. The value of research in human factors in the design and implementation of AI, in general, has now been established and is addressed in specific tracks of influential conferences including the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI) [14] and the Association for the

Advancement of Artificial Intelligence (AAAI) Conference on AI [15]. Yet, there is limited previous work addressing concerns with regards to the human factors of AIED.

Aiming to address such a gap, in a series of studies, Holstein et al., [16–19] investigated the iterative co-design of augmented reality glasses for an intelligent tutoring system (ITS) with K-12 teachers and students. The studies provided valuable insights into teachers' experiences and challenges in using an ITS in their classroom settings [18]. For instance, although teachers often preferred the automation of certain tasks to ease their teaching workload, over-automation of tasks in teaching environments was considered as a threat to their flexibility to choose and implement their own pedagogical goals. Similarly, Van Leeuwen and Rummel [20] documented the teachers' experiences after using three different AIED interfaces (aimed at mirroring, alerting and advising) and identified significant differences in the way teachers can use each of them [21]. Dillenbourg et. al., also investigated teachers' experiences while orchestrating ITSs in collaborative learning contexts [22] and co-designed a series of multimodal analytics prototypes with educators [23]. Just a few studies have focused on the potential role that students may play in the design of a data-intensive educational tool. For instance, Prieto-Alvarez et al. [24] encouraged students to co-create a learner-data journey based on their particular needs and Chen and Zhu [25] investigated students' experience with a visualisation tool that analysed their engagement and interactions with others through social network analysis. Similarly, Chaleer [26] studied students' experience and perceived awareness and usefulness with an ambient group awareness tool. However, the tool was evaluated in a single class, so the students' exposure to it was very limited.

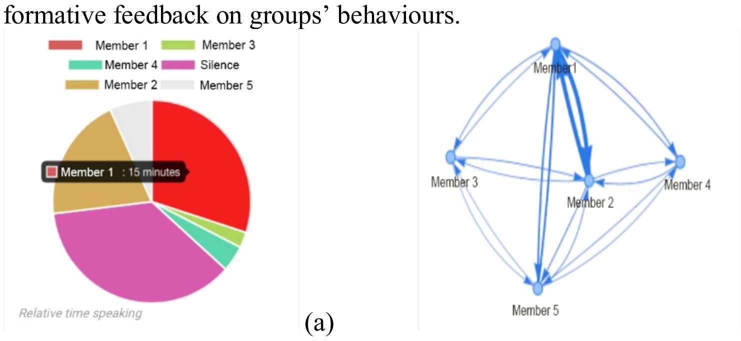
These studies have provided significant contributions to our understanding of teachers and students' experiences with AIED tools in real-world contexts, which then can be used to shape the design and implementation of AIED tools. However, prior work has focused on limited types of AIED tools (i.e., ITSs), limited instructional approaches and goals (i.e., monitoring student activities in classrooms), and mainly focused on the experiences of teachers rather than those of students. In this paper, we build on the emerging literature exploring students' experience of AIED implementations in real-world contexts. We contribute to this literature through the analysis of students' experiences with an AIED tool that monitors their collaboration in remote meetings (using Zoom) as part of a ten-week postgraduate course. The contribution of the paper is two-folded. First, the themes that emerged from the analysis of students' experiences can contribute to and shape the design features of similar systems and their further automation with AI. Second, since it focuses on a novel context for AIED systems -collaboration analytics in synchronous remote meetings using Zoom-, the findings of this study have significant implications for future pedagogical interventions. Remote meetings have become the norm for students studying synchronously at a distance during the coronavirus pandemic, which highlights the timeliness of these contributions.

1.1 Collaboration Analytics and AIED in Remote Meetings

The study presented in this paper was conducted in the context of the use of a collaboration analytics tool. The term Collaboration Analytics refers to AI and Analytics solutions aimed at scrutinising interaction group data to extract insights for supporting sense-making processes and the development of effective collaboration skills [27]. There are plenty of research studies in the literature that are explicitly or implicitly categorized under this umbrella. Some significant examples include but are not limited to AI assistants for scheduling group meetings [28], personal assistants for providing help in collaborative problem-solving [29], real-time gaze feedback with metacognitive supports from a pedagogical agent for dyads [30], utterance analytics of chats and discussion forums to support students' awareness in their involvement [31], feedback provision to groups of students based on their interaction patterns [32], external help-seeking support in collaboration contexts for students [33], and tools to provide summary information of student groups based on certain indicators to support teachers' class monitoring and control [21]. Most available studies describe the design of collaboration analytics in asynchronous online (e.g., [34]) or classroom settings (e.g., [35]). Whilst the virtual meetings have become crucial for remote education due to the need for synchronous collaboration, more work is needed to understand how AI innovations can support reflection and students' learning in such settings. For instance, Cornide-Reyes et al. [36] recently developed the NAIRA system, a real-time multimodal learning analytics tool that inspects students' level of participation within the remote meetings through an influence graph, a speech time distribution, and a silence bar. However, the study did not investigate the students' real-world experiences with the tool in detail.

2 The Context of the Study

The study was conducted in the context of a post-graduate course (covering the design and use of educational technology) that lasted ten weeks. A total of forty-four students completed the course. Students were divided into ten groups, ensuring each group was interdisciplinary (education, design, and technology graduate members) and mixed in terms of gender. Group sizes ranged from three to five. At the beginning of the course, each group was asked to identify an educational challenge. Then, they had to carry out an educational technology design case to solve the challenge and submit a design case solution in Week 10. Analytics generated from online group meetings were used to provide formative feedback on groups' behaviours.



(a) (b)

Fig. 1. (a) A pie chart represents the total speech time per student including the relative fraction of time a group has been silent. Each portion represents the relative speaking time of each student. (b) Turn-taking network represents conversational flows between students.

Groups used Zoom during their regular classes to conduct their planning and design meetings. The ZoomSense system's "sensor" appeared as a participant in the Zoom meetings, recorded the verbal utterances of each student in Zoom, and stored them in a cloud database. The actual content of the meetings was not recorded. Verbal utterances data were then used to model two constructs i) students' total speech time, ii) students turn-taking behaviours. Figure 1(a) presents the total speech time visualisation for group 8 in Week 3. This chart also includes the total silence time and relative speaking time of each student (i.e. the most verbally active group member spoke for 15 minutes). Figure 1(b) shows the turn-taking behaviours of students. This was presented as a network/sociogram, where the direction of the edges depicts the conversational flows from one student to another during the discussions. The thickness of the edges represents the mutuality of the conversation. After every remote meeting, these two visualisations and a written report were sent to each group separately via email. The report served to provide written feedback (a sample of email feedback) to students indicting how they could improve group interactions. In the later versions of the tool, the written feedback was also automatically provided via the Zoom chat to scaffold students' collaboration in *real-time*. In this study, the feedback was sent by teaching assistants every week after group meetings.

3 Methodology

In this paper, we addressed three research questions. i What are the specific needs of students' that can impact the *design features* of collaboration analytics in remote meetings? ii. What are the specific needs of students' that can impact the *educational implementation* of collaboration analytics in remote meetings? iii. What are the *ethical and privacy concerns* of students with regards to being monitored during remote meetings?

To address the research questions, we theoretically framed the student probes according to the components of the Social Translucence (ST) framework: Visibility, Awareness, and Accountability [37]. This framework was proposed to help investigate users' design needs for the particular purpose of computer-mediated, online group activities [23]. Based on ST, a total of twelve open-ended interview questions were used in retrospective semi-structured interviews at the end of the module. Interview questions covering the Visibility dimension (4 questions) focused on the significant aspects of students' online synchronous meetings and what features of their collaboration should be made visible to them. Awareness dimension's questions (3 questions) aimed at exploring to what extent the information provided by the analytics create a well-informed understanding of students' own and others' performance. Accountability questions focused on understanding to what extent the feedback provided by the analytics can help students take responsibility for improving their performance (2 questions). In addition to the ST framework, we added 3 questions to particularly explore privacy and ethics concerns of students with regards to the use of AIED tools in remote meetings.

In total twenty students (four male and sixteen female -representative of the cohorts' gender ratio) volunteered to participate in the interviews. At least one student from all ten groups was included in the sample. None of the participants had any experience of using collaboration analytics or similar AIED tools in the past. The study has received full ethics approval from the host institute of the lead author. All participants were clearly informed and signed consent forms accordingly.

The data analysis was conducted using Braun and Clarke's six phases of thematic analysis [38]. First, the data was transcribed verbatim. Initial thematic codes were generated by two independent researchers individually. After that, themes from two researchers were compared, discussed, and revised to make sure that emerging themes covered all the collected data and that they are auditable. This process led to an agreed final coding scheme. After this process was completed, the final coding scheme was applied to all transcriptions from scratch to ensure consistency.

4 Results

The thematic coding analysis described in the previous section led to the emergence of ten themes from the transcription data. The themes were then categorised into four dimensions: visibility (4.1), awareness (4.2), accountability (4.3), and the ethics (4.4).

4.1 Visibility

Comprehensibility of collaboration analytics. Thirteen participants responded positively with regards to the easiness to comprehend information and straightforward interpretation of the visualisations shown in Figure 1. For example, P11 reacted positively as follows: "*This is the first time that I have seen such a straightforward way to show the interactions during our collaborative learning.*" On the contrary, five participants partially agreed on this (P4,10,12,14,20), one firmly replied 'no' (P17) and one reported uncertainty to answer the question (P6). Overall, they pointed out that the definition of effective contribution was not clear to them and the analytics only covered partial contributions in speech time and turn-taking.

Accuracy of the analytics information. Fifteen participants reported that the graphs are accurate and "similar to their feelings" (P8,15). P5 elaborated: "*I think it clearly shows the volume of contribution. So those who are talking the most, [what] it is showing is quite accurate in terms of calculating who was the person that was talking the most and ... [with whom he was having] conversations with.*" However, four participants (P3,6,11,19) reported differences between the analytics presented and their actual experiences. Notably, P3 and P6 thought their participation was higher than depicted, while P19 argued s/he contributed to the discussion less. There was also a report from P11 that there was always a higher amount of silence presented in the analytics than they experienced as a group.

Lack of quality evaluations and partially represented contribution. However, all participants expressed concerns over the lack of quality evaluations of student contributions. Seven participants specifically raised concerns that their contributions were only quantitatively represented through speech time and turn-taking but it did not show the quality of their contributions which could be "*total rubbish*" (P1), "*off-topic*" (P12) or "*not useful*" (P16). Therefore, higher speech time did not always mean more actual contribution (P5,11,12,13,15,17,18,20). On the contrary, lower speech time could also represent a key contribution to the further progress of their work (P5,6,13,17). Generally, participants argued that the contributions in a group task are more about the quality of the content than its quantity (P1,4,5,6,10,15). Similarly, the turn-taking lines shown in the collaboration analytics, which show conversational flows between group members, were argued to provide potentially misleading information as explained by P14, as follows: "*sometimes someone spoke after me but what he said was not related to what I have said. I think he diverted the topic and I could not reply to him.*"

At the same time, six participants raised concerns over the limitation of unimodal data collection since the information represented with the analytics was only captured from the students' Zoom meetings. Students might be "recorded" as silent in the collaboration analytics, but they might have been focusing on completing their co-design tasks on another collaboration tool beyond what is captured by the system. Furthermore, participants also mentioned various group activities that were crucial to their group work but were excluded from the analytics including their chats via instant messaging platforms such as WhatsApp (P19), additional meetings of sub-groups or group as a whole (P16) that took place out of the module, the final presentation preparations (19) and other forms of preparation before the discussion (P12). To illustrate: "*During the meeting, we might express these points [prepared ideas] with a few sentences in a short time but we might have spent a significant amount of time and energy on preparing them. The speech time cannot represent these pre-meeting preparations.*" – P12.

4.2 Awareness

The value of seeing one's own performance. Participants mutually agreed upon the value of the tool to make them aware of their performance (19 participants), yet their reasons varied. Some reported, thanks to analytics, they ensured a high level of participation (P10) or maintained continuous participation in their meetings (P14). Importantly, the tool appeared to prompt students to reflect on their performance. As P13 reported, "*I asked myself, why was I the person who spoke the least?*" On the other hand, P11, who was a regular high contributor, reported that "*sometimes I would ask myself: Did I speak such a lot?*" In general, collaboration analytics were considered as external objective measures that can help students be less "biased" from their own experience when evaluating their performance in the group activities. As P5 pointed out: "*Obviously about the whole thing about eyewitness testimony, it can be distorted by events that happen post the experience. So, what the graph does, it really helps you to have a clear data point to say, Okay, this is what happened in the group.*"

The value of seeing others' performance. Not only the tool was considered as an enabler for students to reflect on their performance, but it was also considered as an enabler to reflect on others' performance. The majority of the participants (17) acknowledged that collaboration analytics can make them aware of their group members' contributions, and determine who is struggling or need help. P20 explained this as follows: "*[the analytics] can help you know others' contribution better or help you find their problem. We had a new member. He rarely participated in the group work in the last few weeks and he muted himself during the meeting.*" This potential was also recognised by P1, 4 and 9. Surprisingly, such awareness of a struggling member was not that evident without the weekly reports sent to students, as P1 pointed out: "*I didn't know that one of our group members didn't spend a lot of time speaking. I mean, it took him about seven weeks before he told us 'I struggle with your accents'.*"

4.3 Accountability

Collaboration analytics to foster group discussions. The collaboration analytics were considered as a medium for triggering discussions by almost half of the participants (9). While some groups reported having a specific discussion about the analytics occasionally (P3,4,5,9,13,14,19), some reported that constant discussions were going on in their weekly meetings about the previous weeks' feedback (P2,7). For example, P5 explained that "*It did work because one week our meeting started when we were discussing the graphs. The persons who were showing to be contributing less, were talking about why they felt they were doing that. And one highlighted an issue where somebody felt that they didn't understand the material enough to contribute that week.*"

Self-regulation and socially shared regulation of behaviours. At the individual level, nineteen participants tried to regulate their behaviours and adapt their level of interaction according to the collaboration analytics (i.e., if they had a high level of participation and dominated the discussion in one meeting, they tended to speak less in consecutive meetings). This was indicated by P1, as follows: "*...[after seeing analytics* *on their group behaviours] I shut up. I didn't talk for about half an hour.*" Similar incidents were reported by P15. In contrast, if they had a low level of participation, they tried to speak more. As P13 described "*once I was detected to have less speech time, I would speak more in the next time. I would try my best to catch up with my teammates and have more interactions with them.*" Some students also reflected on how their activity or lack of preparation outside of the meeting reflected their levels of interaction during the meeting. For instance, seven participants (P7,8,13,16,18,19,20) attributed their low level of interaction to lack of preparation for the meeting and hence, tried to prepare more in future meetings. To illustrate this, P19 explained that she could not contribute much if she did not finish the weekly readings. As a result, she aimed to finish the weekly readings, check the weekly tasks, and prepare contributions for the group discussions in advance.

Regulation of student behaviours appeared to occur also at a social level. Twelve participants reported various strategies they used to regulate their behaviours based on their understanding of others' needs. For instance, they were encouraging the less active speakers to speak more (P2,9,11,14,18); helping others diagnose their problems (P4); providing a further explanation and inviting struggling members to contribute (P19); and developing group strategies such as assigning a weekly host for the group discussions (P12). Some participants were also able to make informed strategic changes as P5 argued: "*for myself and another person in the group, we could see that we were talking back and forth quite a lot. So, one week, we made a pact to not keep responding to each other's points yet to open up the floor for others in the group to respond to questions.*" However, whether regulated behaviours were beneficial for learning or not was not clear. For instance, P4 reported that the analytics directed her towards responding to people, not about discussing the contents: "*I was very much concerned with making sure I had good airtime and decent thick lines between the various people. And so, it became more about a response, less thinking about what that person said.*"

Gaming the system. 'Gaming the system' refers to a situation where students attempt to accomplish a task within the system by not truthfully working on the tasks as intended but rather taking advantage from the gap within the system [39]. There were four reports of 'gaming the system' (P3,12,14,17). P14 acknowledged that for the least active speakers to have more interaction, s/he performed the following action: "*[another member] discussed something not related to our tasks but easy for [the least active speaker] to talk in the meeting.*" The same approach was followed in the group of P3, as she described: "*because we wanted to give space [to members spotted as less active] so that it would be more equal, we would end up letting someone talk about completely random subjects, just that they had enough time.*"

Swinging back to "normal", the tentative nature of the changes. Notably, the changes to the group discussions dynamics informed by the tool were not long-lasting. Seven participants reported swinging back their "normal" after a short while, whereas six participants noticed the tentative nature of the changes of other members' behaviours. Multiple reasons for the short-term nature of the changes were provided: including the lack of control during the heat of the discussions (P10,11), the restriction on their speech-time giving unspontaneous flows of conversation (P3,6), the lack of summative evaluations of their collaboration (P2,7,11,12,1316). Overall, one-third of participants argued for the value of integrating the tool and the assessment motives. As P7 elaborated: "*I am a behaviourist sort of thing. I feel like I don't really contribute much* *because I don't really focus there because I know this will not affect my final mark. Where if I was thinking maybe that is a 5% or 2% of our final marks will be affected. I think people would contribute more.*"

4.4 Privacy and ethics concerns

Half of the participants reported that they did not have any concerns and claimed they ignored the fact of being monitored in their group meetings, with P5 explaining: "*I'd say we completely forgot the sensors were there, aside from them just appearing in the panel, and we were like presenting our screens anyway.*" P3 reasoned that this comfort in being monitored might be due to the course's subject area. As she explained: "*We came on this course to learn about educational technology. So, in that sense to do that, it wasn't shocking, you know? Not like if I'd come, maybe on a different course, maybe then I'd find it really weird.*" P12 also reported no concerns due to her interest in AIED. Moreover, two participants (P15,18) argued their comfort was due to the formative use of the tool, as it was not for summative assessment: "*If they [the analytics] were only there for feedback but not assessment, I think that's alright to be monitored*" (P15). By contrast, one-quarter of participants said their concerns were rather fleeting and the other quarter added that they were significantly concerned. Four participants (P1,2,3,15) asked to confirm whether the tool recorded their voices as P2 described their group concerns that "*there is one thing that we always discuss about… are you [the lecturers] listening to everything that we are talking about?... some information even though it's supposed to be private, it is not really private.*" Additionally, five participants revealed uncomfortable feelings upon being monitored, such as feeling "*uncomfortable*" (P3), "*strange*" (P4), "*super-concerned*" (P6), "*nervous*" (P15), and "*being spied on*" (P4,15). Interestingly, these concerns were particularly observed from students with low contributions. As P6 stated: "*It was really, really challenging. So, knowing that something is monitoring how much time I speak, I had the pressure to do it and it went out of hand. The second week, I was under pressure. I think I spoke like two minutes or so.*" P3 reported that her group was more spontaneous when not being monitored: "*We had some sessions outside of the bots. And yeah, then we did not worry about that[being observed] anymore. Whoever needed to say something said it. If we wanted to have a chat, we had a chat….Personally, I was a bit different and I felt we were more spontaneous.*" This aligned with reports from P6,7,9,15,20 that they would have acted more openly if they were not being observed.

On the contrary, P4,5,7,8,14,17 asserted that the being monitored helped them to act productively as their group was "supervised" indirectly through the tool. P4 explained that "*this small thing that sits in your head is echoed publicly, in some way is representative of who you are, and your teachers are seeing this, and you don't want to look bad to your professors.*" P17 reported that: "*To be honest, I have stayed here [the university] for three years. I had my undergraduate here, acted as an invisible man. I don't have confidence so I rarely express my opinions in the class. Since this year we had the [tool], I forced myself to express more about my opinions.*"

5 Discussion and Conclusion

The results presented above have significant implications for the design and implementation of AI tools for collaboration in educational remote meetings. With regards to our first research question on the design implications, results show that the collaboration analytics in remote meetings have the potential to make students aware of their own as well as their group members' collaborative behaviours. However, students argued that the tool only represented a small part of their actual contribution and so they did not always perceive the tool as significant for their success in the course. The main critiques were the lack of content analysis and unimodal nature of the tool. Due to these design drawbacks, students struggled to make connections between what the tool represents and what really 'mattered' for their learning. It was argued that content analysis of the discussion that would provide proxies on the quality of the contributions by group members -in addition to the quantity of contribution- is essential for the uptake of the tool. Therefore, we suggest that future designs of similar AIED tools should consider involving the content analysis and multiple modalities in their collaboration analytics. For instance, detection of off-topic discussions and introduction of data from writing analytics from chats as a second modality can increase the value of collaboration analytics in remote meetings. Similarly, perhaps at a more practical level, future iterations that involve data analysis from multiple platforms (i.e. collaborative docs, chats, presentation platforms) can lead to more holistic representations of student contributions in remote meeting settings. In turn, such representations are more likely to lead to a stronger relationship between students' awareness of their performance and to what extent they change their behaviours accordingly [23, 40, 41].

Results also indicated that the reflections driven through awareness can lead students to change their behaviours in remote meetings. As discussed in self-regulated learning (SRL) literature [42, 43], by providing means to students to support evaluation not only of the overall progress of the group but rather to make an accurate attribution of personal contribution to the group progress (reflection phase), students can plan their future learning and correct their expectations (forethought phase) [42]. Therefore, the awareness provided by the tool has the potential to improve students' learning in remote meetings. However, such changes in student behaviours were argued to be temporary and many students returned to their "normal" behaviours in remote meeting interactions. This is aligned with research investigating the effects of digital tools on behaviour change persistency in general [44]. Multiple reasons were presented by students for the observed phenomenon of "regressing to business as usual". This phenomenon is partly related to the incomplete representation of students' contributions which we have discussed above. Moreover, students reported that this "back to normal" may be caused by the lack of intervention. Since the tool did not provide guidance or suggestions to the students during the meeting, it is challenging for students to make a change on time. Therefore, the future design of collaboration analytics tools should not only focus on providing visualisations but should also include real-time automated feedback on what actionable steps they can take to improve their collaboration behaviours. On the other hand, the guidance may also be structured into the implementation of the collaboration analytics tool which is explored in the second research question.

Our second question investigated the suggestions for educational implementation of AIED tools with collaboration analytics in remote meetings. Firstly, students would benefit from instructions that would scaffold them on what sort of actions they could potentially take based on their reflections of the collaboration analytics. As some students noted, although they realised that they needed to change certain behaviours, they did not know exactly how to do this. This may be due to the feedback sent regarding students' participation which did not have strong elements on how students' can regulate their actions. Therefore, they struggled to adapt and change their behaviours accordingly [45]. Future implementations should involve clear instructions on what further actions can be taken to address the tool's suggestions. Secondly, the learning context in which the tool was implemented significantly affected to what extent students engaged with it. For instance, in this study, the analytics were not considered as part of the summative assessment, so some students were not motivated to take long-lasting actions based on them. This leads to the suggestion that teachers and AIED designers should carefully align the collaboration analytics and the learning design including assessment [46]. Thirdly, better instructions on what kind of analytics outcomes are expected for different group tasks were deemed as important. Some students regulated their behaviours to equalise the contribution in their group discussions, others purposely made no effort in this regard as they considered some of the group meetings as peer learning opportunities rather than collaboration. They wanted to learn from the students who have more experiences and knowledge. This may indicate that students have varied definitions of collaboration for different group tasks. Therefore, an alignment of group tasks' learning design, its collaboration analytics, and their consequent visualisations should ideally be shared with students in advance. As discussed in the literature, there are distinctions between collaborative learning, cooperative learning and peer learning [47] which may require students to present different behaviours [48].

Regarding our third research question, we explored students' privacy concerns about being monitored by the collaboration analytics tools. Most students did not report negative emotions towards being monitored and some reported motivational value in being observed. One possible reason may be that the analytics were not part of the summative assessment. It was also argued that students were behaving more comfortably as they knew the system could not record the content of their discussions. This highlights the importance of informing students about what the AIED tool can and cannot do and how it will be implemented. Yet, this also leads to a significant dilemma. On the one hand, students asked for more detailed investigations of their collaborative behaviours (i.e., content analysis) and argued that the tool would make them more accountable if the analytics involved summative assessments. On the other hand, students argued that they would have more significant privacy concerns had this has been the case.

5.1 Limitations and Future Research

Since the participants were postgraduate students and the course was in educational technology, it is challenging to generalise the results. Similar studies in diverse contexts are called for drawing a better picture of student experiences. Moreover, although there were indications about the value of the tool to help students regulate their behaviours, future work is needed to delineate to what extent the tool supports self-regulation (SRL) ("regulate oneself"), co-regulation ("supporting each other") or socially shared regulation (SSRL) ("regulating together") [43]. Based on the findings, a future version of the system may include the generation of fully automated real-time prompts, to be sent to students via the Zoom chat, to scaffold students' collaboration based on the discussion dynamics, including SRL (e.g., ask the student who demonstrated no verbal activity in the last 5 minutes to verbally summarise the current state of discussion) and SSRL (e.g., advice to the most active students to involve less active students). However, further co-design evaluations of prompts are needed before any potential AI-driven automation to understand what exact behaviours need to be prompted, when exactly, and how.

References

1. Li, H., Gobert, J., Dickler, R.: Evaluating the Transfer of Scaffolded Inquiry: What Sticks and Does It Last? In: Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., and Luckin, R. (eds.) Artificial Intelligence in Education. pp. 163–168. Springer International Publishing, Cham (2019)
2. VanLehn, K., Banerjee, C., Milner, F., Wetzel, J.: Teaching Algebraic Model Construction: A Tutoring System, Lessons Learned and an Evaluation. Int. J. Artif. Intell. Educ. 30, 459–480 (2020). https://doi.org/10.1007/s40593-020-00205-3
3. Baker, R.S.J. d., D'Mello, S.K., Rodrigo, Ma.M.T., Graesser, A.C.: Better to be frustrated than bored: The incidence, persistence, and impact of learners' cognitive–affective states during interactions with three different computer-based learning environments. Int. J. Hum.-Comput. Stud. 68, 223–241 (2010). https://doi.org/10.1016/j.ijhcs.2009.12.003
4. D'Mello, S., Graesser, A., Picard, R.: Toward an Affect-Sensitive AutoTutor. Intell. Syst. IEEE. 22, 53–61 (2007). https://doi.org/10.1109/MIS.2007.79
5. Du Boulay, B.: Intelligence tutoring systems that adapt to learners motivation. Presented at the October 1 (2018)
6. Azevedo, R., Cromley, J.G., Seibert, D.: Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? Contemp. Educ. Psychol. 29, 344–370 (2004). https://doi.org/10.1016/j.cedpsych.2003.09.002
7. Laru, J., Järvelä, S.: Integrated Use of Multiple Social Software Tools and Face-to-Face Activities to Support Self-Regulated Learning: A Case Study in a Higher Education Context. In: Wong, L.-H., Milrad, M., and Specht, M. (eds.) Seamless Learning in the Age of Mobile Connectivity. pp. 471–484. Springer Singapore, Singapore (2015)
8. Winne, P.H.: Enhancing Self-Regulated Learning for Information Problem Solving with Ambient Big Data Gathered by nStudy. In: Adesope, O.O. and Rud, A.G. (eds.) Contemporary Technologies in Education: Maximizing Student Engagement, Motivation, and Learning. pp. 145–162. Springer International Publishing, Cham (2019)
9. Luckin, R., Cukurova, M.: Designing educational technologies in the age of AI: A learning sciences-driven approach. Br. J. Educ. Technol. 50, 2824–2838 (2019). https://doi.org/10.1111/bjet.12861
10. Rosé, C.P., McLaughlin, E.A., Liu, R., Koedinger, K.R.: Explanatory learner models: Why machine learning (alone) is not the answer. Br. J. Educ. Technol. 50, 2943–2958 (2019). https://doi.org/10.1111/bjet.12858
11. Buckingham Shum, S., Ferguson, R., Martinez-Maldonado, R.: Human-Centred Learning Analytics. J. Learn. Anal. 6, (2019). https://doi.org/10.18608/jla.2019.62.1
12. Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O., Rodrigo, M., Cukurova, M., Bittencourt, I., Koedinger, K.: Ethics of AI in Education: Towards a Community-wide Framework. Int. J. Artif. Intell. Educ. (2021)

13. Sjödén, B.: When Lying, Hiding and Deceiving Promotes Learning - A Case for Augmented Intelligence with Augmented Ethics. In: Bittencourt, I.I., Cukurova, M., Muldner, K., Luckin, R., and Millán, E. (eds.) Artificial Intelligence in Education. pp. 291–295. Springer International Publishing, Cham (2020)
14. Kuniavsky, M., Churchill, E., Steenson, M.W. eds: Designing the user experience of machine learning systems. In: AAAI Spring Symposium Proceedings (Technical Report SS-17-04). The AAAI Press, Palo Alto, CA, U.S.A. (2017)
15. Amershi, S., Weld, D., Vorvoreanu, M., Fourney, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P.N., Inkpen, K., Teevan, J., Kikin-Gil, R., Horvitz, E.: Guidelines for Human-AI Interaction. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–13. ACM, Glasgow Scotland Uk (2019)
16. Holstein, K., McLaren, B.M., Aleven, V.: Student Learning Benefits of a Mixed-Reality Teacher Awareness Tool in AI-Enhanced Classrooms. In: Penstein Rosé, C., Martínez-Maldonado, R., Hoppe, H.U., Luckin, R., Mavrikis, M., Porayska-Pomsta, K., McLaren, B., and du Boulay, B. (eds.) Artificial Intelligence in Education. pp. 154–168. Springer International Publishing, Cham (2018)
17. Holstein, K., McLaren, B.M., Aleven, V.: Informing the Design of Teacher Awareness Tools Through Causal Alignment Analysis. In: 13th International Conference of the Learning Sciences (ICLS 2018) (2018)
18. Holstein, K., McLaren, B.M., Aleven, V.: Designing for Complementarity: Teacher and Student Needs for Orchestration Support in AI-Enhanced Classrooms. In: Isotani, S., Millán, E., Ogan, A., Hastings, P., McLaren, B., and Luckin, R. (eds.) Artificial Intelligence in Education. pp. 157–171. Springer International Publishing, Cham (2019)
19. Holstein, K., McLaren, B.M., Aleven, V.: Co-Designing a Real-Time Classroom Orchestration Tool to Support Teacher–AI Complementarity. J. Learn. Anal. 6, 27–52 (2019). https://doi.org/10.18608/jla.2019.62.3
20. van Leeuwen, A., Rummel, N.: Comparing teachers' use of mirroring and advising dashboards. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. pp. 26–34. ACM, Frankfurt Germany (2020)
21. van Leeuwen, A., Bos, N., van Ravenswaaij, H., van Oostenrijk, J.: The role of temporal patterns in students' behavior for predicting course performance: A comparison of two blended learning courses: Comparison of two blended learning courses. Br. J. Educ. Technol. 50, 921–933 (2019). https://doi.org/10.1111/bjet.12616
22. Pierre Dillenbourg, Luis P. Prieto, Jennifer K. Olsen: Classroom Orchestration. In: International Handbook of the Learning Sciences. Routledge (2018)
23. Echeverria, V., Martinez-Maldonado, R., Buckingham Shum, S.: Towards Collaboration Translucence: Giving Meaning to Multimodal Group Data. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19. pp. 1–16. ACM Press, Glasgow, Scotland Uk (2019)
24. Prieto-Alvarez, C.G., Martinez-Maldonado, R., Shum, S.B.: Mapping learner-data journeys: evolution of a visual co-design tool. In: Proceedings of the 30th

Australian Conference on Computer-Human Interaction. pp. 205–214. ACM, Melbourne Australia (2018)

25. Chen, B., Zhu, H.: Towards Value-Sensitive Learning Analytics Design. Proc. 9th Int. Conf. Learn. Anal. Knowl. 343–352 (2019). https://doi.org/10.1145/3303772.3303798
26. Charleer, S., Klerkx, J., Duval, E.: Towards balanced discussions in the classroom using ambient information visualisations. Int J Technol. Enhanc. Learn. 9, 27 (2017)
27. Martinez-Maldonado, R., Kay, J., Shum, S.B., Yacef, K.: Collocated Collaboration Analytics: Principles and Dilemmas for Mining Multimodal Interaction Data. Human–Computer Interact. 34, 1–50 (2019). https://doi.org/10.1080/07370024.2017.1338956
28. Cranshaw, J., Elwany, E., Newman, T., Kocielnik, R., Yu, B., Soni, S., Teevan, J., Monroy-Hernández, A.: Calendar.help: Designing a Workflow-Based Scheduling Agent with Humans in the Loop. Proc. 2017 CHI Conf. Hum. Factors Comput. Syst. 2382–2393 (2017). https://doi.org/10.1145/3025453.3025780
29. Winkler, R., Bittner, E., Neuweiler, M.L., Söllner, M.: Hey Alexa, Please Help Us Solve This Problem! How Interactions with Smart Personal Assistants Improve Group Performance. Presented at the Fortieth International Conference on Information Systems , Munich (2019)
30. Hayashi, Y.: Gaze awareness and metacognitive suggestions by a pedagogical conversational agent: an experimental investigation on interventions to support collaborative learning process and performance. Int. J. Comput.-Support. Collab. Learn. 15, 469–498 (2020). https://doi.org/10.1007/s11412-020-09333-3
31. Trausan-Matu, S., Dascalu, M., Rebedea, T.: PolyCAFe—automatic support for the polyphonic analysis of CSCL chats. Int. J. Comput.-Support. Collab. Learn. 9, 127–156 (2014). https://doi.org/10.1007/s11412-014-9190-y
32. Kim, T., McFee, E., Olguin, D.O., Waber, B., Pentland, A. “Sandy”: Sociometric badges: Using sensor technology to capture new forms of collaboration: sensor technology and collaboration in teams. J. Organ. Behav. 33, 412–427 (2012). https://doi.org/10.1002/job.1776
33. Alavi, H.S., Dillenbourg, P.: An Ambient Awareness Tool for Supporting Supervised Collaborative Problem Solving. IEEE Trans. Learn. Technol. 5, 264–274 (2012). https://doi.org/10.1109/TLT.2012.7
34. Kent, C., Cukurova, M.: Investigating Collaboration as a Process with Theory-driven Learning Analytics. J. Learn. Anal. 7, (2020). https://doi.org/10.18608/jla.2020.71.5
35. Janssen, J., Erkens, G., Kirschner, P.A.: Group awareness tools: It’s what you do with it that matters. Comput. Hum. Behav. 27, 1046–1058 (2011). https://doi.org/10.1016/j.chb.2010.06.002
36. Cornide-Reyes, H., Riquelme, F., Monsalves, D., Noel, R., Cechinel, C., Villarroel, R., Ponce, F., Munoz, R.: A Multimodal Real-Time Feedback Platform Based on Spoken Interactions for Remote Active Learning Support. Sensors. 20, 6337 (2020). https://doi.org/10.3390/s20216337

37. Erickson, T., Kellogg, W.A.: Social translucence: an approach to designing systems that support social processes. ACM Trans. Comput.-Hum. Interact. 7, 59–83 (2000). https://doi.org/10.1145/344949.345004
38. Braun, V., Clarke, V.: Using thematic analysis in psychology. Qual. Res. Psychol. 3, 77–101 (2006). https://doi.org/10.1191/1478088706qp063oa
39. Baker, R., Walonoski, J.A., Heffernan, N., Roll, I., Corbett, A., Koedinger, K.: Why Students Engage in "Gaming the System" Behavior in Interactive Learning Environments. J. Interact. Learn. Res. 19, 185–224 (2008)
40. Prestigiacomo, R., Hadgraft, R., Hunter, J., Locker, L., Knight, S., van den Hoven, E., Martinez-Maldonado, R.: Learning-centred translucence: an approach to understand how teachers talk about classroom data. In: Proceedings of the Tenth International Conference on Learning Analytics & Knowledge. pp. 100–105. ACM, Frankfurt Germany (2020)
41. Szostek, A.M., Karapanos, E., Eggen, B., Holenderski, M.: Understanding the implications of social translucence for systems supporting communication at work. In: Proceedings of the ACM 2008 conference on Computer supported cooperative work - CSCW '08. p. 649. ACM Press, San Diego, CA, USA (2008)
42. Panadero, E.: A Review of Self-regulated Learning: Six Models and Four Directions for Research. Front. Psychol. 8, (2017). https://doi.org/10.3389/fpsyg.2017.00422
43. Järvelä, S., Järvenoja, H., Malmberg, J.: Capturing the dynamic and cyclical nature of regulation: Methodological Progress in understanding socially shared regulation in learning. Int. J. Comput.-Support. Collab. Learn. 14, 425–441 (2019). https://doi.org/10.1007/s11412-019-09313-2
44. Yanovsky, S., Hoernle, N., Lev, O., Gal, K.: One size does not fit all: A study of badge behavior in stack overflow. J. Assoc. Inf. Sci. Technol. n/a, (2020). https://doi.org/10.1002/asi.24409
45. Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., Kirschner, P.A.: Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. Comput. Hum. Behav. 107, 105512 (2020). https://doi.org/10.1016/j.chb.2018.05.004
46. Lockyer, L., Heathcote, E., Dawson, S.: Informing Pedagogical Action: Aligning Learning Analytics With Learning Design. Am. Behav. Sci. 57, 1439–1459 (2013). https://doi.org/10.1177/0002764213479367
47. Damon, W., Phelps, E.: Critical distinctions among three approaches to peer education. Int. J. Educ. Res. 13, 9–19 (1989). https://doi.org/10.1016/0883-0355(89)90013-X
48. Cukurova, M., Luckin, R., Millán, E., Mavrikis, M.: The NISPI framework: Analysing collaborative problem-solving from students' physical interactions. Comput. Educ. 116, 93–109 (2018). https://doi.org/10.1016/j.compedu.2017.08.007