Empowering online teachers through predictive learning analytics

Authors:
Herodotou, Christothea*; Hlosta, Martin; Boroowa, Avinash; Rienties, Bart; Zdrahal, Zdenek; Chrysoula Mangafa

*Corresponding author: Christothea.herodotou@open.ac.uk

Affiliation: The Open University UK

Dr Christothea Herodotou is an Associate Professor at the Institute of Educational Technology, Open University UK. She coordinates the evaluation of Early Alert Indicators, a university-wide predictive analytics project for improving student performance and retention.

Dr Martin Hlosta is the technical lead and data scientist of OU Analyse, at the Knowledge Media Lab, Open University. He has developed the predictive models and he leads the design of OU Analyse, and its integration with Open University systems.

Avinash Boroowa is the Project Lead of the Early Alert Indicators Project. He leads a multi-disciplinary team of senior academics, data scientists, data analysts and developers, providing thought leadership in the deployment of predictive analytics to support students.

Prof Bart Rienties is Professor of Learning Analytics at the Institute of Educational Technology, Open University. He is the programme director of learning design and learning analytics and lead of the OU data wrangling activities.

Prof Zdenek Zdrahal is Professor of Knowledge Engineering at the Knowledge Media Lab, Open university. He conceived the OU Analyse (2014-present) concept, which builds on his JISC project RETAIN (2011-2012) and the MSR-OU project with the MS Research Cambridge (2012-2013).

Dr Chrysoula Mangafa is a Lecturer at the Institute of Educational Technology, Open University interested in the use of technologies to support inclusive education. She has supported the OUA evaluation with teachers.

ABSTRACT
This study presents an advanced predictive learning analytics system, OU Analyse (OUA), and evidence from its evaluation with online teachers at a distance learning university. OUA is a predictive system that uses machine-learning methods for the early identification of students at risk of not submitting (or failing) their next assignment. Teachers have access, via interactive dashboards, to weekly predictions of risk of failing for each of their students. In this study, we examined how the degree of OUA usage by 559 teachers, of which 189 were given access to OUA, related to student learning outcomes of more than 14K students in 15 undergraduate
Online teachers and predictive analytics

courses. Teachers who made ‘average’ use of OUA were found to benefit their students the most; after controlling for performance differences, these students were found to have significantly better performance than their peers in the previous year’s presentation during which the same teachers made no use of predictive learning analytics. Implications for the teaching practice, in particular the empowerment of teachers in online and distance learning settings are discussed.

1. INTRODUCTION

Online teachers are professionals involved in the provision of online and distance learning who have a different set of roles and responsibilities compared to face-to-face teachers. Online teaching is often distributed amongst a group of teachers, rather than managed by a single teacher (Toetenel & Rienties, 2016). Teaching involves responsibilities such as designing the process and learning activities of online courses, facilitating online student discussions, creating audio-visual materials (e.g., videos, interactive exercises), setting up the course on an online platform, and ensuring rights clearance for learning material (Papathoma, 2018). Furthermore, an important role for online teachers is to coach students through individualised tutoring and feedback, and facilitate interactions with the materials and other learners (Chen & Lee, 2011; Haya, Daems, Malzahn, Castellanos, & Hoppe, 2015; J. M. Lin, Wang, & Lin, 2012; McKenney & Mor, 2015).

Amongst the roles particularly emphasised in the literature is the provision of pedagogical and personal support to students (Lin et al., 2012; Muñoz Carril, González Sanmamed, & Hernández Sellés, 2013; Rienties, Herodotou, Olney, Schencks, & Boroowa, 2018; Starčič & Vukan, 2019). Recent research amongst 72K students engaged in 74 modules indicated that 69% of weekly engagement by students were primarily determined by how online teachers designed online courses, and how they supported their students (Nguyen, Rienties, Toetenel, Ferguson, & Whitelock, 2017). Academic knowledge and success can be structured through teacher’s online presence that facilitates critical discussion and dialogue (Dockter, 2016). This role is particularly challenging given the lack of face-to-face interactions that can inform about difficulties students may face (Crawley, Fewell, & Sugar, 2009), as well as the large number of students a teacher may have to monitor and support.

A wide body of literature has identified that Predictive Learning Analytics (PLA) can help teachers, and institutions in general, to overcome these challenges, by identifying which groups of students might need extra support to reach desired learning outcomes (Conijn, Snijders, Kleingeld, & Matzat, 2017; Haya et al., 2015; C. Herodotou et al., 2017; Kovanović
et al., 2015; Scheffé et al., 2017). For example, Scheffé et al. (2017) found that visualising relative performance to 172 students using PLA supported learning processes over time. Yet, Conijn et al. (2017) found that PLA results varied substantially and concluded that Virtual Learning Environments (VLEs) engagement and clicking data are of limited value for effective prediction of learning outcomes.

While an emerging body of learning analytics literature has highlighted mixed results about the effectiveness of PLA for actionable interventions and proactive support (Ferguson & Clow, 2017; Fincham, Gasevic, Jovanovic, & Pardo, 2018; Viberg, Hatakka, Bälter, & Mavroudi, 2018), it is yet unanswered whether PLA can effectively empower teachers to intervene on time, on a large scale basis, and across a range of disciplinary contexts. As evidenced in several recent reviews about the uptake of learning analytics and PLA in particular (Ferguson et al., 2016; Ferguson & Clow, 2017; Viberg et al., 2018), although institutions and teachers are interested in learning analytics, their actual uptake in most institutions is rather limited. Also, there is limited evidence beyond relatively small case-studies of PLA usage with teachers (van Leeuwen, Janssen, Erkens, & Brekelmans, 2014), experimental lab studies (Worsley & Blikstein, 2015), and MOOC contexts (Kizilcec, Pérez-Sanagustín, & Maldonado, 2017) that PLA can have sustained positive effects on learners and learning outcomes (Fincham et al., 2018).

In this study, we aim to provide evidence from a large-scale implementation at a distance learning university about how and when PLA can used by teachers to effectively support student learning. In this study, 189 out of 559 online teachers were given access to OU Analyse (OUA), a PLA system, as a means to gain insights about the performance of 14,128 students and proactively support them. In our quasi-experimental design, we were particularly interested to explore how the teachers’ engagement with the PLA has a positive (or negative) impact on student performance in comparison to the previous year when these teachers did not have access to PLA data. This study aims to answer the following Research Questions (RQs):

RQ1: How does teachers’ usage of OUA relate to student learning outcomes?

RQ2: How does student performance compare to peers in previous years’ course presentation during which teachers had no access to OUA?

In the next section we discuss the role, responsibilities and challenges online teachers face and how insights from learning analytics could be used to support the teaching practice. We also present existing insights about how learning analytics and PLA can support student learning outcomes, when used by teachers.
2 THEORETICAL BACKGROUND

2.1 The role and challenges of online teachers

Online teachers are presented with a range of technology and pedagogy-related challenges that necessitate the transformation of the teaching practice when moving to instruction in online settings (Lin et al., 2012; Redmond, 2011). Amongst the challenges online teachers reported is the achievement of online presence (teaching, social, and cognitive), defined as the need for interactivity and encouraging learners to regulate their learning (Kilgour et al., 2018). For example, interviews with 23 postgraduate students at a Canadian distance education university (Edwards et al., 2011) indicated that exemplary or effective online educators are those who (a) challenge learners to perform beyond their current capacity, have high expectations for them, (e.g., through strong cognitive presence and materials that question and challenge learners), (b) affirm personal growth, (e.g., through strong online presence by recognizing the students’ potential problems and taking actions to support them), (c) influence learners through subject matter expertise and strong online presence (e.g., through, feedback and email communication).

The provision of effective pedagogical support may be particularly demanding and challenging in online settings, where teachers are often interacting with a large number of students (Toetenel & Rienties, 2016), and they rely on ‘clues’ rather than direct interaction with students (as in face-to-face teaching) to identify and solve difficulties students may face (Chen & Lee, 2011). In particular, online teachers make use of multiple data sources to inform their understanding of whether and how students engage with online courses (Fincham et al., 2018; Bart Rienties, Herodotou, Olney, Schencks, & Boroowa, 2018). Accordingly teachers at the Open University UK (OU) provide support using available student data (demographics, VLE engagement etc.), student contact history made available via so-called Student Support Teams (SSTs) who get in touch with students, as well as a system that monitors (non) submission of assessments. In addition to that, teachers were found to develop their own approaches of monitoring students and their interactions with them, such as the creation of excel sheets with information about students and their contact history (Walker et al., 2019). Yet, this is not the case for all online teachers. For example, Herodotou et al. (2017) showed that student support and monitoring approaches such as emailing, calling, and texting varied across teachers, with some teachers being very proactive and persistent in monitoring and supporting students, while others viewing SSTs as the agents in charge of supporting students.
In this study, we argue that it is the teacher responsibility to monitor student behaviour online in a proactive manner, identify, and provide support to students who may face some difficulties, while at the same time further challenging students who perform well in order to promote personal growth and development. In the next sections, we discuss how machine learning techniques and learning analytics tools can empower teachers in realizing this demanding, yet critical teaching responsibility.

2.2 Learning analytics and online teachers
Mixed findings are reported in studies assessing the use of PLA data and visualisations with teachers (C. Herodotou et al., 2017; van Leeuwen et al., 2014; van Leeuwen, Janssen, Erkens, & Brekelmans, 2015). For example, in a study with 28 teachers van Leeuwen et al. (2014) identified that teachers who received learning analytics visualisations of collaboration activities were better able to identify participation problems. Also, they were found to intervene more often in “problematic” groups as opposed to a control group of teachers who did not receive learning analytics visualisations. Positive outcomes are also reported by McKenney and Mor (2015) reporting that the teachers’ professional development was enhanced by engaging with analytics software.

Through a mixed-method study at a distance learning institution, Herodotou et al., (2017) found mixed effects on student performance when 240 teachers in 10 modules were given access to PLA, yet usage analysis showed that teachers made only limited use of PLA in their practice that could explain these mixed effects. Five follow-up interviews revealed that teachers had positive views about using PLA in teaching as they recognised their usefulness for complementing the teaching practice and being 'on top of things'. Also, Herodotou, Rienties, Borooa, Zdrahal, & Martin (n.d.) in a multi-methods study with 59 teachers, and more than 1,300 students identified that teachers overall engagement with predictive data is the second most significant factor explaining student performance, following previous best performance.

Similarly, Dazo, Stepanek, Chauhan, and Dorn (2017) analysed usage data of 14 teachers and identified that frequency of learning analytics visits decreased between semesters. A follow-up focus group with six teachers pointed out that teachers faced challenges in interpreting learning analytics data, and this led some of them to shift to other methods of monitoring students’ progress, such as reading their posts. Overall, there is an emerging body of evidence showing that PLA can be effective in some cases, yet not in others, raising the need for more, robust, longitudinal research beyond a single context or discipline. In order to provide
convincing evidence to involve stakeholders, including teachers, in this study we describe a large-scale study with more than 14K students whose teachers were granted access to OUA.

3 METHODOLOGY

3.1 OU Analyse (OUA)

OU Analyse (OUA) is a predictive system used to identify learners at risk of failing their studies (Fig. 1), which has been used at a large scale at the Open University (OU) since 2013. OUA predicts on a weekly basis whether (or not) a given student will submit their next teacher-marked assignment. It uses a traffic light system to pinpoint in red students at risk of not submitting, in amber those with a moderate probability of failing or barely passing, and in green those who are likely to succeed.

Figure 1: Section from the OUA dashboard with predictions for individual students about submitting their next assignment.

The validity and accuracy of OUA has been widely tested amongst 45K students and across 40+ courses (Hlosta, Zdrahal, & Zendulka, 2017) and has been relatively widely cited in the learning analytics community. Predictive models take into account students' demographic data and course activity in VLE for which predictions are calculated. The OUA dashboard visualises predictive information about who is at risk to submit the next assignment for individual students, as well as VLE engagement and assignment submission rates at the course level.
3.2 Participating courses
Fifteen courses (N=15) with 14,128 undergraduate students and 593 teachers, presented in the academic year 2017/18 joined the study from a range of disciplines (9 Science; 4 Technology; 1 Health and social care; 1 Law). Teachers teaching in more than one courses were excluded from the analysis (n=34) along with their students, resulting in a total of 559 teachers. It is noted that OU teachers work on a part-time basis, and they normally also hold other full-time jobs and/or other responsibilities. In addition, their employment contracts did not foresee participation in research activities, and no contractual obligations were included in terms of PLA engagement. These issues should be considered when interpreting our results. Of course we are mindful that active OUA engagement by online teachers does not automatically translate to being a “good” or a “bad” teacher, as there could be as well other effective online teaching strategies teachers may deploy. Furthermore, active engagement by teachers is no guarantee of students’ success. Access to OUA was given to teachers who volunteered to use the system (n=189), of which a majority (65.6%) made use of the system at least once. Degree of PLA usage was calculated in percentages considering for varied course lengths i.e., number of weeks a course ran. We used nonparametric statistics for comparing groups as after inspection of the skewness and kurtosis measures, histograms, normal Q-Q plots and box-plots and the Kolmogorov-Smirnov test (p<.001) the sample data were not approximately normally distributed. A non-parametric Levene's test verified equality of variances in groups (p>.05) (Nordstokke & Zumbo, 2010).

4 RESULTS
Weekly usage of OUA ranged between 3% and 84% of the course length. Figure 2 shows the percentage of teachers (tutors) who actually accessed OUA in relation to those who were originally given access to the system. Teachers in Science courses were more engaged with OUA than other teachers, and throughout the course presentation from the start to the end. Technology teachers mostly visited OUA at the beginning and middle of their courses while Health care and Law teachers are shown to access OUA at specific points during the course, that coincide with the weeks an assignment was due. Although we again stress that teacher engagement in OUA is not necessarily an indicator of being a great teacher, at least it provides an initial proxy of engagement.
To answer RQ1, we first aggregated individual teacher's usage data with respective students and grouped teachers into three groups: (a) Group “Low”: No or limited usage of OUA (0-8% of OUA use) (n=461), (b) Group “Average”: Average usage of OUA (10-39% of weeks of course length) (n=74) and (c) Group “High”: High usage of OUA (41-84%) (n=24). Figures 3, 4, and 5 show OUA usage by teachers per group. In the “Low” usage group (Figure 3), a larger percentage of teachers accessed OUA, yet in specific weeks only. In the “Average” and “High” groups (Figure 4 and 5), less teachers accessed the system yet participation was more systematic and spread throughout the course presentation.

Figure 3: OUA usage (%) by the “low” group per week
A Kruskal-Wallis H test showed a statistically significant difference in student overall performance between the three groups ($\chi^2(2) = 28.24, p<.001$). Post hoc pairwise comparisons using the Mann-Whitney test indicated that student performance was significantly greater for Group High (Mdn = 57) than for Group Average (Mdn = 50) ($U = 649355.5, p = .013$), and for Group Low (Mdn = 46.7) ($U = 2866806.5, p<.001$). Also, Group Average (Mdn = 46.7) was better performing than Group Low (Mdn = 50) (Adjusted $p$ value due to multiple comparisons $p=.005/3=.016$). These findings suggest that the degree of OUA usage related positively to student overall course performance, with higher usage significantly related to better learning outcomes.

Chi-square analysis with pass rates as dependent variable showed significant differences between groups ($\chi^2=17.19, p<.001$, df=2). Using post hoc cellwise residual
Online teachers and predictive analytics

analysis, we identified significant differences between Group Low and Group High (adjusted \( p \) level=.008). A greater percentage of students in Group High (55.9%; \( p=.0003 \)) passed their courses as opposed to Group Low (48.1%, \( p=.0006 \)), suggesting that teachers with high OUA usage had more students passing their courses compared to teachers who did not use or made very limited use of OUA.

Answering RQ2, we examined whether better student performance outcomes in Group “Average” and Group “High” may be explained by teachers who were performing “better” in general (over and above PLA use). Teachers (n=54) in these two groups who taught the same courses the previous two academic years, yet they did not use or have access to OUA in the past were included in the analysis. Average student performance in current presentations was compared to previous years’ presentations (in which teachers did not use OUA). Wilcoxon Signed Ranks test for Group “High” (n=12) showed no statistically significant differences between current and past course presentations (\( Z=-.549, \ p=.583 \)) (Mdn_current=61, Mdn_past=60.3). On the contrary, statistically significant differences were observed for Group “Average” (n=42) (\( Z=-2.44, \ p=.015 \)’ adjusted \( p \) value \( p=.05/2=.025 \)) (Mdn_current=59.8, Mdn_past=56.2), suggesting that teachers who made average use of OUA had better student outcomes in the current year compared to previous years in which they did not use OUA. In addition, we controlled for variation in previous students’ performance that may explain better learning outcomes in the current year. Wilcoxon Signed Rank test showed no statistically significant differences (Group Average: Mdn_past= 66.5, Mdn_current=68.9, \( p=.635 \), NS; Group High: Mdn_past= 73.73, Mdn_current= 74.33, \( p=.286 \), NS), suggesting that observed differences are not related to students' differences in previous performance. These findings suggest that insights from OUA were particularly useful to teachers who made average use of the system as these teachers were shown to have better student performance outcomes during the course presentation they had access to OUA. These differences were not related to different student performance characteristics.

5 DISCUSSION

In this study, we examined whether the level of OUA usage relate to student performance (RQ1) and whether differences are related to teacher and student performance characteristics (RQ2). In terms of RQ1, this study suggests that frequent access to OUA predictions by online teachers can enhance student performance and increase pass rates (C. Herodotou et al., n.d.). “Average” OUA usage (as low as 10% of the weeks a course runs) may have a positive impact on student outcomes. When used in a systematic way, PLA systems such as OUA can be
powerful tools for teachers to identify students at risk at an early stage, and provide just-in-time support, enhancing student performance. What is yet not known is how teachers intervened (e.g., medium and type of conversation) to ensure that students at risk will submit their next assignment. The limited evidence available showed a variation of strategies (e.g., email, phone, message, calling SST) used across different teachers, yet with no evidence as to whether and which had a positive effect on students’ engagement (Toetenel & Rienties, 2016).

Answering RQ2, online teachers in the high OUA frequency group were shown to have the same average student performance across presentations, indicating that OUA usage is less likely to have made an impact on their practice. In contrast, teachers in the “average” OUA frequency group had significantly better average student performance in the current course presentations in which they used OUA relative to the previous presentation without OUA. These findings may be explained by teachers' existing approaches of monitoring students' engagement with the course material online. As shown in other studies (C. Herodotou et al., 2017), PLA helped some teachers in monitoring their students' behaviour online; rather than looking for information about students in different places (forums, VLEs, etc), OUA provided them with all the information needed in a single space. Teachers in the “high” frequency group may be teachers who already proactively check on their students (even when they did not have access to PLA), and thus maintain consistently high student performance. On the contrary, teachers in the “average” usage group for whom OUA usage was shown to have a positive effect on students' performance, may be those who are less likely to have devised a specific approach for monitoring student behaviour. OUA may have empowered them to proactively support their students by giving them insights about their performance in a systematic manner during the entire course presentation. In terms of the low usage group, these findings suggest that OUA insights could help this group of teachers to better monitor their students and achieve better learning outcomes, especially if those teachers do not have alternative ways of monitoring students and proactively engage and support them.

6 CONCLUSIONS
This study described a large-scale case-study about how PLA could empower teachers to proactively support their students and promote better learning outcomes. It contributed to the lack of large-scale learning analytics implementations and reaffirmed the importance and value of teachers to improve student learning over and above artificial intelligence and other technology-enhanced advances (Lin & Spector, 2018). PLA is a human intelligence innovation that could have a positive impact on learning by addressing one of the major challenges online.
teachers are now facing, that of effectively supporting learners’ growth and development. In this quasi-experimental study, we showed how PLA can be used to identify and intervene with, in particular, students at risk of failing their studies. PLA could also be used by teachers to identify those students who are performing well and who may require additional challenge. PLA can empower teachers in a number of ways: (a) gather together and visualise distributed data about individual and cohorts of students, (b) provide systematic monitoring (e.g., weekly) of the engagement and performance of a large group of students, (c) identify and flag students who require immediate attention in order to succeed with their studies, and (d) identify and flag well-performing students who should be further challenged. PLA can be beneficial to both the teachers by complementing and enhancing existing teaching practices (C. Herodotou et al., 2017) as well as the students by facilitating better learning outcomes (C. Herodotou et al., n.d.)

Yet, as shown in this study the uptake of PLA by teachers remains relatively low, and this necessitates further examination. Factors that may explain this trend are the long-standing positions of teachers at the university, other work responsibilities, and teachers' contractual agreements that do not contain provisions for using PLA. Our previous work (Christothea Herodotou, Rientes, Verdin, & Boroowa, n.d.) suggests that PLA adoption in distance learning higher education can be facilitated through a number of actions related to both the teachers and the broader organisational context including: (a) the generation of evidence of PLA effectiveness - this study contributes towards this direction, (b) the identification of effective support interventions for students at risk, (c) the facilitation of communication across all involved stakeholders, (d) the use of PLA to inform decisions about who needs support, (e) the understanding and mitigation of teachers’ resistance, (f) the involvement of managers in the process of adoption and (g) the use of PLA as a means to enrich existing teaching practices.

It is in our future research agenda to explain the observed teachers’ usage patterns by interviewing teachers from the various OUA usage groups to understand the reasons why some teachers make frequent use of the system while others choose to check PLA rarely or never. We will also seek to understand whether and how broader teaching beliefs and intentions as well as digital literacy may relate to usage patterns and intervention strategies. At the moment, we are exploring whether email reminders from teachers’ managers can motivate frequent engagement with predictions and spark appropriate interventions.

REFERENCES
https://doi.org/10.1111/j.1467-8535.2009.01035.x


Dockter, J. (2016). The Problem of Teaching Presence in Transactional Theories of Distance Education. *Computers and Composition, 40*, 73–86. https://doi.org/10.1016/j.compcom.2016.03.009


Online teachers and predictive analytics


