Using Learning Analytics to Scale the Provision of Personalised Feedback

Abelardo Pardo, Jelena Jovanovic, Shane Dawson, Dragan Gašević, Negin Mirriahi

Abelardo Pardo is Associate Professor in the School of Electrical and Information Engineering at The University of Sydney. His areas of research are educational technology and learning analytics. Address for correspondence: Abelardo Pardo, Electrical Engineering J03, The University of Sydney 2006 NSW Australia. Email: Abelardo.pardo@sydney.edu.au.

Jelena Jovanovic is Associate Professor in the Department of Software Engineering at the University of Belgrade, Serbia. Her research interests include semantic technologies and learning analytics. Address for correspondence: Jove Ilica 154, 11000 Belgrade, Serbia. Email: jeljov@gmail.com.

Shane Dawson is the Director of the Teaching Innovation Unit and Professor of Learning Analytics at the University of South Australia. His research interests lie in learning analytics, social learning and organisational leadership. Address for correspondence: Shane Dawson, Teaching Innovation Unit, University of South Australia, 5001, SA, Australia. Email: shane.dawson@unisa.edu.au

Dragan Gašević is Professor and Chair in Learning Analytics and Informatics in the Moray House School of Education and School of Informatics at the University of Edinburgh. His research interests are in learning analytics, self-regulated and social learning, and educational policy. Address for correspondence: Dragan Gašević, Informatics Forum, 10 Crichton Street, Edinburgh, EH8 9AB, United Kingdom. Email: Dragan.Gasevic@ed.ac.uk

Negin Mirriahi is Senior Lecturer in the Teaching Innovation Unit at University of South Australia. Her research interests are in the area of technology adoption, blended and online learning, academic staff development, MOOCs, learning analytics to inform pedagogical practice, and video analytics to enhance learning.

Abstract

There is little debate regarding the importance of student feedback for improving the learning process. However, there remain significant workload barriers for instructors that impede their capacity to provide timely and meaningful feedback. The increasing role technology is playing in the education space may provide novel solutions to this impediment. As students interact with the various learning technologies in their course of study, they create digital traces that can be captured and analysed. These digital traces form the new kind of data that are frequently used in learning analytics to develop actionable recommendations that can support student learning. This paper explores the use of such analytics to address the challenges impeding the capacity of instructors to provide personalised feedback at scale. The case study reported in the present paper showed how the proposed approach had a positive impact on student perception of feedback quality and on academic achievement. The study was conducted with first year undergraduate engineering students enrolled in a computer systems course with a blended learning design across three consecutive years (\(N_{2013}=290\), \(N_{2014}=316\), and \(N_{2015}=415\)).

Keywords: Feedback, learning analytics, student engagement
Introduction

Student directed feedback has been identified as one of the most important factors influencing a student’s academic achievement (Hattie, 2008). Although the research literature offers numerous suggestions and principles concerning the effective provision of feedback (e.g. Evans, 2013; Gikandi, Morrow, & Davis, 2011), students are generally dissatisfied with the quality of the feedback they receive (P. Ferguson, 2011). At the same time instructors, an especially those responsible for large or highly diverse student cohorts, are under increasing pressure to dedicate resources and time for providing feedback to improve student outcomes, retention and satisfaction. In essence, as course enrolments increase there is a diminishing level of time per student that an instructor can devote in order to develop timely, personalised feedback.

A solution may lie in the use of Learning Analytics (LA). The increase in technology mediation of learning activities is producing data sets with unprecedented detail about how students interact in virtual learning spaces. Research areas such as Educational Data Mining (EDM) and LA explore how this wealth of data can be used to increase understanding, and overall quality of learning. The ultimate objective of collecting and analysing such data is to produce actionable knowledge connected with the learning environment that can be used to inform learning and teaching practice.

There are two major challenges to be addressed to use LA techniques to support feedback processes. The first lies in the development of informative lead indicators of student learning progression that can be used by instructors to create real-time feedback. The second is related to the scaling of such feedback to large and diverse student cohorts. This research study aims to demonstrate how these impediments to providing high quality feedback can be addressed. The paper explores how LA can be used to support instructors in providing meaningful and personalised feedback to large student cohorts in data-rich learning contexts and how to quantify the effect of this feedback in terms of student satisfaction with feedback and academic achievement.
Related Research

Feedback

The concept of feedback in education has been comprehensively studied. Early studies in the area of management theory provided a basic initial definition of feedback in terms of information about the gap between the actual and reference levels of a parameter in a system and identified the main factors that contribute to its effectiveness (Ramaprasad, 1983). Various studies explored the effect of factors such as frequency of feedback on student learning gains (Black & Wiliam, 1998). Several theoretical models have been proposed to explain how feedback is produced, received, and what effect it has on students (Boud & Molloy, 2013a; Butler & Winne, 1995; Evans, 2013; Kluger & DeNisi, 1996). These models have been used as a framework to inform the principles related to the provision of feedback (Chickering & Ehrmann, 1996; Nicol & Macfarlane-Dick, 2006). Despite the volume and rigour of the studies investigating feedback, the concept is still considered fragile (P. Ferguson, 2011) and inherently problematic (Higgins, Hartley, & Skelton, 2010). This situation is especially delicate in the context of contemporary higher education where students attend large classes that effectively diminish an individual’s opportunity to interact with teaching staff. As such, it is not overly surprising that several studies have shown that students’ dissatisfaction with the quality and quantity of feedback is a common concern raised across the majority of higher education institutions (Carless, 2006; Krause, Hartley, James, & McInnis, 2005).

Later studies pushed the definition of feedback away from the mere provision of information towards a dialogic framework (Hounsell, 2007; Nicol, 2010; Sadler, 2010; Yang & Carless, 2013) and Boud and Molloy (2013b) finally redefined it as a process in which learners obtain information that helps them appreciate the similarities and differences with some appropriate standards to improve their work. This new approach prompted the need for sustainable feedback strategies (Boud & Molloy, 2013a; O’Donovan, Rust, & Price, 2015) based on a dialogue between students and instructors or among students that emphasizes the connection with learning and promotes higher student engagement.

However, the recent trend in higher education towards massive student cohorts and the adoption of active learning techniques poses a serious hurdle to the scalability of dialog-based solutions. While these practices accommodate economies of scale they make the provision of feedback a complex and intensive undertaking. Teacher-student interaction is difficult to sustain in such environments and often typically degrades into a one-way; one to many communication process (Nicol, 2010). Although methods relying on peer feedback may reduce the required instructor time (Carless, Salter, Yang, & Lam, 2011), the role of the expert is still needed to maximise learning gains (Chang, 2011; Porte, Xeroulis, Reznick, & Dubrowski, 2007).

To address the challenges noted above, the study in this paper explored the effect of providing personalised feedback messages for large student cohorts at the levels of learning process and self-regulation as suggested by Hattie and Timperley (2007) and observing the principles of facilitating self-assessment, provide opportunities to close the gap, and clarify what good performance is as suggested by (Nicol & Macfarlane-Dick, 2006).
Learning analytics

The field of Learning Analytics (LA) emerged with the objective of using data to increase the insight in learning experience and better support students (Dawson, Gašević, Siemens, & Joksimovic, 2014). LA makes use of machine learning and predictive modelling techniques to analyse learners’ digital traces to understand and optimise learning processes.

LA research represents a broad array of methods that are used to derive support actions for students (R. Ferguson, 2012). One particular category includes methods for the so-called Early Warning Systems (Jayaprakash, Moody, Eitel, Regan, & Baron, 2014; Lonn, Krumm, Waddington, & Teasley, 2012). These systems make use of existing data sets collected by student information systems and learning management systems (LMSs). Such datasets are used for building machine learning models aimed at predicting when students will require special support actions to avoid dropping a course, alleviate existing difficulties, or eliminate misconceptions. The information derived from these models is either made available to instructors, or directly shown to the students in the form of different LA dashboards (Verbert et al., 2014) to support reflection (Krumm, Waddington, Teasley, & Lonn, 2014; Tanes, Arnold, King, & Remnet, 2011).

More recent research in LA investigates student reactions to presented data visualisations regarding their engagement and how such information is used to inform any change in their approach to study. For example, (Corrin & de Barba, 2015) found that students were unable to accurately interpret information provided in commonly used dashboards and thus, the effects on their learning were either non-existent and sometimes even negative. As pointed out by some researchers (Gašević, Dawson, & Siemens, 2015; Wise, 2014), additional focus needs to be placed on how to transform the use of data into a positive influence in learning scenarios. In other words, there is a gap between the outputs of a data-processing algorithm and interpretation for a positive impact on a student’s learning experience.

The approach adopted in this study proposes the use of previously collected detailed data about student behaviour to support instructors to create quasi-immediate personalised feedback messages for courses with large student cohorts.

Method

Approach

The paper presents a novel approach for the provision of feedback based on a set of engagement indicators that are derived from the students’ activities in the learning environment. The approach assumes that students engage in activities organised into cycles with an identical curriculum structure. The purpose of these cycles is to promote the use of the feedback received at the end of one cycle to improve the activities in the following cycle (Hounsell, McCune, Hounsell, & Litjens, 2008). In the study, each cycle has a duration of one week and contains a set of tasks presented to students through the LMS. The activities include interactive formative assessments (e.g., videos followed by formative multiple-choice questions). The LMS captures students’ digital traces as a collection of learning events that represents a detailed account of how students interacted with the resources.

In a traditional setting, the provision of feedback at the end of each cycle would require an instructor to review the engagement measures for each student and provide a set of comments according to the observed behaviour. In a fully automated
approach, an algorithm receives as its inputs a model of the domain of knowledge (the topics covered by the activities), a set of observed events associated with a particular student, and a set of rules. Based on this information the automated model selects an appropriate feedback item for the student (Graesser, Conley, & Olney, 2012). In these two (traditional and automated) scenarios, the feedback message would be personalised.

This paper suggests a combination of the previous two procedures. Specifically, the proposed model aims at enhancing the capacity of human instructors to provide personalised comments to groups of students depending on their engagement observed in the learning environment at the learning process and self-regulation levels (Hattie & Timperley, 2007). For each activity defined in the course design, instructors prepare in advance a set of feedback messages for different levels of interaction with learning resources. For example, if an activity contains a video, the instructors will provide feedback for students who barely glanced over the video, for those who watched a significant portion of the video, for those who watched the entire video, and for those who watched the video several times. The assumption behind this approach is that instructors can use the level of engagement with the activity to modulate the comment sent to the students so that it is much more personal and connected with the students’ behaviour. In this scenario, a cycle with $n$ activities and $m$ students would require instructors to write $k \times n$ comments in advance thus avoiding the dependency on the number of students in the cohort. The number $k$ is derived from the number of engagement categories instructors are interested in; in the above example of video use, four different categories were used (i.e., $k=4$). That is, the proposed approach is not depended on the number of students and thus, it is scalable.

Once these comments are created, an algorithm is executed at the end of each instructional cycle. For each student and for each activity in the cycle the algorithm selects the appropriate comment based on the level of the student’s participation in the activity. The comments selected for the activities are collated into a detailed feedback message that is then sent to the student either through a virtual learning environment, or a personalised email. The personalised suggestions can then be taken into account by the students for the next cycle (feed-forward).

**Hypotheses**

The research hypotheses explored in the study are:

- RH1: The provision of feedback through personalised messages based on the students’ engagement with the learning tasks increases the students’ satisfaction with feedback in a blended learning course.
- RH2: The provision of feedback through personalised messages based on student engagement with learning tasks increases academic achievement.

RH1 targets the commonly identified gap in the perception of useful feedback between instructors and students (Huxham, 2007). RH2, on the other hand, seeks to verify that the personalised messages have an impact on academic achievement as documented in other studies (Hattie, 2008; Kluger & DeNisi, 1996).

**Context**

The course under analysis is a 13 week first year computer engineering course at a large research intensive university in Australia. The intervention took place in weeks 2-5 of the 2015 course edition and the results were compared with the 2013 and 2014
editions of the same course for RH1, and with the 2014 edition for RH2. The design of the course involves 1-week long cycles whereby students engage with course activities including videos, formative multiple-choice questions, and summative exercises. All activities were available through the institutional LMS that captured all student interactions with the implemented learning activities and resources. The course assessment was comprised of a midterm examination accounting for 20% of the full course marks, session preparation accounting for 20%, a project accounting for 20%, and a final exam accounting for 40%.

The instructor created four feedback comments for the 37 activities in the four cycles (weeks 2-5) preceding the midterm examination for a total of 138 short (one or two sentence) text snippets. The algorithm to combine the comments related to all the activities for each individual student was executed at the end of each of the four weeks and the collated set of comments were included in a personalised email sent to each student at the end of each week. The diagram shown in Figure 1 depicts the workflow used for each week.

![Figure 1 Weekly workflow to compose the messages](image)

The effect of this approach was measured with respect to two aspects of the course in order to answer RH1 and 2: self-reported student satisfaction with the quality of feedback, and academic performance in the midterm exam. Student satisfaction was considered for the course offerings in 2013, 2014 and 2015. The academic performance was considered for the 2014 and 2015 editions. The exam material is not disclosed from year to year, and therefore was used identically in both editions. The exam from the 2013 edition was excluded due to a different exam structure. The personalised emails were deployed only in the 2015 edition.

**Participants**

The number of students enrolled in the 2013, 2014, and 2015 editions were respectively 291 (46 females, 245 males, 97.3% engineering students), 315 (57 females, 257 males, 1 other, 93.6% engineering students), and 414 (75 females, 339 males, 94.2% engineering students).

**Measures**

Three data sources were used in the study. The first one was derived from the interactions of the students with the resources available in the LMS. Three types of interactions were recorded: watching a video, completion of multiple-choice questions, and engagement with a sequence of summative exercises. These data were
used only during the 2015 edition to create the personalised feedback messages. The second data source was derived from the institutional student evaluation of teaching survey, which was made available to the students during the last three weeks of the semester in the 2013, 2014, and 2015 editions. Students answered to a set of questions that were based on a Likert-like scale with values “strongly disagree”, “disagree”, “neutral”, “agree”, and “strongly agree”. Question 6 of the survey was selected as relevant for the study since it addressed specifically how the students perceived the provision of feedback. During the 2013 and 2014 editions of the course, the question was “Feedback from my assessment and otherwise was useful in helping me learn”. In the 2015 edition, the institutional survey was redesigned and the question was reworded to “I have been guided by helpful feedback on my learning”. Given the institutional nature of the survey, the authors had no control over the deployment of this reworded version. The data obtained from the surveys was completely anonymous.

The third data source included the scores of the midterm examination. The exam took place in week 6 of the semester and consisted of 20 multiple-choice questions similar to those that students had available as part of their course activities.

**Procedure**

The intervention was implemented in weeks 2 to 5 of the 2015 edition of the course and consisted of sending each student a personalised email at the end of the week with detailed feedback about their engagement with the activities proposed for that week. Thus, each student received four emails before the midterm exam that took place in week 6.

For each week and each activity, the instructor prepared text messages for four categories of engagement with the following rules:

- Engagement with video activities. Two variables were considered: the number of times a student “played” a video, and the number of seconds the video was watched. These variables were discretised by dividing their values into quartiles.
- Engagement with multiple-choice questions. Different messages were written for each quartile of the distribution of the number of incorrect questions.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Message</th>
</tr>
</thead>
<tbody>
<tr>
<td># incorrect &gt; 22</td>
<td>“Make sure you practice again the exercises. Make sure you understand two concepts: memory operations, and memory size. See if you can go through the entire sequence without errors”</td>
</tr>
<tr>
<td># incorrect &lt;= 22 and # incorrect &gt; 11</td>
<td>“Good initial work. However, you should try again and make sure you fully understand how memory works. Choose those answers that you don’t understand why they are correct, and post them in the forum”</td>
</tr>
<tr>
<td># incorrect &lt;= 11 and #incorrect &gt; 6</td>
<td>“Good work with the exercises. You may want to review the answers again in a few days to make sure the concept of how memory works is fully understood”</td>
</tr>
<tr>
<td># incorrect &lt;= 6</td>
<td>“Excellent work with the exercises. You may want to keep the link handy to give it a final review before the exam”</td>
</tr>
</tbody>
</table>
Engagement with sequence of summative exercises. Different messages were written for each quartile of the distribution of the number of incorrect exercises.

The use of quartiles translated into a process with \( k = 4 \) as per the formula explained in the previous section, thus requiring instructors to write four comments per activity. Table 1 shows an example of the four engagement categories (conditions) resulting from the analysis applied to a sequence of summative exercises in 2015. In the given example, the number of incorrect exercises had a distribution with quartiles delimited at 6, 11, and 22 incorrect answers. The message shown in the first row targets students with the largest number of incorrect answers whereas the last row shows the message for those that submitted the least number of incorrect answers.

![Example of a personalised message.](image)

Table 1 shows an example of the four engagement categories (conditions) resulting from the analysis applied to a sequence of summative exercises in 2015. In the given example, the number of incorrect exercises had a distribution with quartiles delimited at 6, 11, and 22 incorrect answers. The message shown in the first row targets students with the largest number of incorrect answers whereas the last row shows the message for those that submitted the least number of incorrect answers.

![Example of a personalised message.](image)

Once all messages were created for each activity, an algorithm was used to assign the appropriate text to each student based on their engagement behaviour and performance in the formative quizzes. For each student, the engagement indicator is compared with the conditions (first column in Table 1) and the appropriate text is selected. The messages associated with all the activities were then collated into a personalised email. Figure 2 shows an example of a portion of the personalised email sent to a student.

### Results and Discussion

Table 2 shows the number of students that received a message, the number of activities considered for the message, and the number of unique emails sent (percentage over the number of students). The number of students decreased over time because those with no activity recorded during a week were removed from the list for the following week to avoid sending repetitive messages about their lack of engagement. The last column shows the percentage of unique emails over the number of students.

Table 2 Description of the email processing parameters and results for each week

<table>
<thead>
<tr>
<th>Week</th>
<th># Students</th>
<th># Activities</th>
<th># Diff Emails</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>444</td>
<td>7</td>
<td>262 (59%)</td>
</tr>
<tr>
<td>3</td>
<td>417</td>
<td>5</td>
<td>254 (61%)</td>
</tr>
<tr>
<td>4</td>
<td>408</td>
<td>5</td>
<td>226 (55%)</td>
</tr>
<tr>
<td>5</td>
<td>394</td>
<td>4</td>
<td>180 (46%)</td>
</tr>
</tbody>
</table>
Change in perception of feedback by the students

A one-way between-subjects ANOVA was conducted to compare the effects of the year on the level of student satisfaction with feedback reported by the 2013, 2014, and 2015 cohorts. There was a significant effect of the year on the reported level of satisfaction with feedback at the p < 0.05 level for the three years [F(2, 615) = 20.41, p =2.60e-9, eta-squared=0.062]. Post hoc comparisons using the Tukey HSD test indicated that the mean score for the level of satisfaction with feedback in 2015 (M = 3.82, SD=0.90) was significantly different than the values in 2014 (M=3.35, SD=1.03) and 2013 (M=3.25, SD=0.97) with p<0.05. The effect size (Cohen’s d) was 0.49 with respect to 2014, and 0.61 with respect to 2013. There was no statistically significant difference between the 2013 and 2014 editions. These results suggest that the treatment in the 2015 edition in which the personalised feedback was included had a large effect on how students perceived feedback.

Figure 3 shows the change in the mean of students’ perception of the usefulness of the received feedback over the 3 examined course editions as well as the confidence intervals derived from the ANOVA.

Impact on academic performance

Unlike in the previous section, the impact on academic performance was measured comparing the scores of the midterm exam for the 2014 and 2015 editions. The exam of the 2013 edition was not considered because it contained different questions. An independent-samples t-test was conducted to compare the midterm scores in the two editions of the course. There was a significant difference between the answers received in the 2014 edition (M=12.801, SD=4.794) and the 2015 edition (M=13.835, SD=4.893); t(684.499) = -2.859, p = 0.002, with an effect size (Cohen’s d) of 0.213. This result suggests that the provision of personalised feedback messages in the 2015 edition had a small to medium positive effect on the midterm score.
Although the size effect of the intervention was not as large on the midterm score as the change in perception of the quality of the feedback, the combination of these two results shows that the proposed approach had a significant and positive impact on the learning experience of the students.

The messages in this case study were sent only in the weeks preceding the midterm examination. The approach can be easily extended to be maintained throughout the semester. The messages included detailed suggestions about how to approach the tasks scheduled for each week. For this reason, the suggestions could be extended not only for the entire semester, but for a variety of courses (each course would have different suggestions for the different tasks).

Conclusions

The provision of effective feedback that help students to reflect on the similarities and differences between their work and the expected standard is increasingly challenging in the current conditions found in higher education institutions. Feedback is typically facilitated either through increasingly labour intensive manual procedures, or with fully automated tools. The current availability of data sets with detailed accounts of student interactions offers the possibility of exploring how technology can augment human intelligence to provide personalised feedback at scale for large student cohorts.

The study described in this paper highlights the potential of combining digital traces typically captured by technology mediation in a learning environment with instructor knowledge to provide frequent and personalised feedback messages for large student cohorts. The approach included a learning design that used cycles that repeat throughout the length of the experience, a set of messages created by the instructor for each task depending on four levels of engagement, and a mechanism to combine these messages into personal emails. The instructor produced 138 text snippets with advice on 37 tasks that were combined based on the observed indicators. If scaled, the approach could include the provision of a set of initial comments to the instructors to modify or tailor to the specifics of a course. Also, the approach could accommodate a fully automated approach by algorithmically identifying the text to be provided in a rule. The results in the study show that the messages had a positive effect both in student satisfaction with feedback, and academic performance in a midterm exam. These results point to feedback as an important factor to support student success and the need for new approaches to establish a connection between student data and how to consider it to provide high quality feedback.

The results also point to the important role of instructors when deploying support actions for students. A limitation of this study is that it did not include the point of view of the instructor. Future studies may explore aspects such as the time an instructor needs to dedicate to this approach, the opinion of experienced instructors, or the ideal message wording to maximise engagement (Wright, McKay, Herschok, Miller, & Tritz, 2014). The study also highlighted that detailed knowledge of the learning experience places instructors at an advantageous position to use technology to transform their expertise into highly situated, personalised student-directed feedback. Future research directions derived from this study point to the need for better methods for identifying individual differences in student engagement and their overall learning experience. Other factors such as learning strategies or study habits, if properly detected, could offer the possibility of interventions with potentially high effects.
Ethical Guidelines

The research adhered to ethical standards and guidelines as the nature of study demanded. The statistical analysis was performed using already existing non-identifiable data.

Potential Conflict of Interest

The authors have no conflicts of interest to declare in relation to this work.

References


