Learning analytics dashboard for improving the course passing rate in a randomized controlled experiment

Author(s)
Hellings, Jan

Publication date
2016

Document Version
Final published version

Published in
Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)

Practitioner Track Proceedings

6th International Learning Analytics & Knowledge Conference (LAK16)

University of Edinburgh, Edinburgh, UK
April 25–29, 2016

Edited by
Rebecca Ferguson, Mike Sharkey and Negin Mirriahi

SoLAR
SOCIETY for LEARNING ANALYTICS RESEARCH

A SoLAR Publication

Suggested citation:

Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). Society for Learning Analytics Research

The Society for Learning Analytics Research (SoLAR) is an inter-disciplinary network of leading international researchers who are exploring the role and impact of analytics on teaching, learning, training and development.

LAK16 was sponsored by Brightspace, Blackboard, HT2 learning, IBM and McGraw-Hill Education

CONTENTS

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

INTRODUCTION

Practitioners spearhead a significant portion of learning analytics, relying on implementation and experimentation rather than on traditional academic research. Both approaches help to improve the state of the art. The LAK conference has created a practitioner track for submissions, which first ran in 2015 as an alternative to the researcher track.

The primary goal of the practitioner track is to share thoughts and findings that stem from learning analytics project implementations. While both large and small implementations are considered, all practitioner track submissions are required to relate to initiatives that are designed for large-scale and/or long-term use (as opposed to research-focused initiatives). Other guidelines include:

- *Implementation track record* The project should have been used by an institution or have been deployed on a learning site. There are no hard guidelines about user numbers or how long the project has been running.
- *Learning/education* related Submissions have to describe work that addresses learning/academic analytics, either at an educational institution or in an area (such as corporate training, health care or informal learning) where the goal is to improve the learning environment or learning outcomes.
- *Institutional involvement* Neither submissions nor presentations have to include a named person from an academic institution. However, all submissions have to include information collected from people who have used the tool or initiative in a learning environment (such as faculty, students, administrators and trainees).
- *No sales pitches* While submissions from commercial suppliers are welcome; reviewers do not accept overt (or covert) sales pitches. Reviewers look for evidence that a presentation will take into account challenges faced, problems that have arisen, and/or user feedback that needs to be addressed.

Submissions are limited to 1,200 words, including an abstract, a summary of deployment with end users, and a full description. Most papers in the proceedings are therefore short, and often informal, although some authors chose to extend their papers once they had been accepted.

Papers accepted in 2016 fell into two categories.

- **Practitioner Presentations** Presentation sessions are designed to focus on deployment of a single learning analytics tool or initiative.
- **Technology Showcase** The Technology Showcase event enables practitioners to demonstrate new and emerging learning analytics technologies that they are piloting or deploying.

Both types of paper are included in these proceedings. The technology showcases are i by the word 'Showcase' at the start of their title.

Rebecca Ferguson, The Open University
Mike Sharkey, Blackboard
Negin Mirriahi, University of New South Wales

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

Learning analytics dashboard for improving the course passing rate in a randomized controlled experiment

Jan Hellings
Amsterdam University of Applied Science
j.f.hellings@hva.nl

The intention of this study was to increase the number of students passing a Java programming course by giving feedback on their online behaviour through a learning analytics dashboard. The treatment consisted of a randomized controlled experiment with 556 students. Of these, 276 students received for eight weeks an email with a link to their dashboard and 280 students were assigned to the control group. The treatment and control groups are comparable on the analysed characteristics. There was no significant difference in withdrawal between the control and treatment group. The results show that offering the dashboard had no significant effect on the percentage passing and the exam marks. The analysis of the use of Myprogramminglab showed that the treatment group practised 5% more than the control group. This result is significant at p= .1 . There was a significant difference between the online activity of students in the academic year 2015 and that in 2014. In 2014, online activity was much higher than in 2015.

INTRODUCTION

For computer science students at the Amsterdam University of Applied Science (HBO-ICT) in Amsterdam, 'Programming' is a first-year 10-week course of Java programming. The course is supported by a number of e-learning systems including the Myprogramminglab of Pearson[1] and the Moodle[2] environment. The classes were designed according to the 'Flip the Classroom' (Davies, Dean, & Ball, 2013) principle. At home the students prepared their lesson by taking online quizzes and exercises.

At the end of the course in week 10 the students are tasked with programming a small Java program and are then graded on the results of their efforts. The results of the programming course were insufficient, about 42%, instead of the desired 30% of the students n=1217 failed in 2013 and 2014. The activities of the students in e-learning environments tend to be predictable for results on the course (Hu, Lo, & Shih, 2014; Tempelaar, Rienties, & Giesbers, 2014). The analyses of the online behaviour in the population of 2014 n=684 showed that there was a correlation between online behaviour and the result of programming. Students who did all the online exercises scored an average of 6.8 (SD=2.7) and the ones who did not scored an average of 4.8 (SD=3.5) ($t(682)= -7.84$ $p < .000$) These results led to the development of an learning analytics dashboard displaying information about how well the students perform on their online task and the predicted result and chance of passing the course.

---

[1] http://www.pearsonmylabandmastering.com/northamerica/myprogramminglab/

[2] http://moodle.org

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

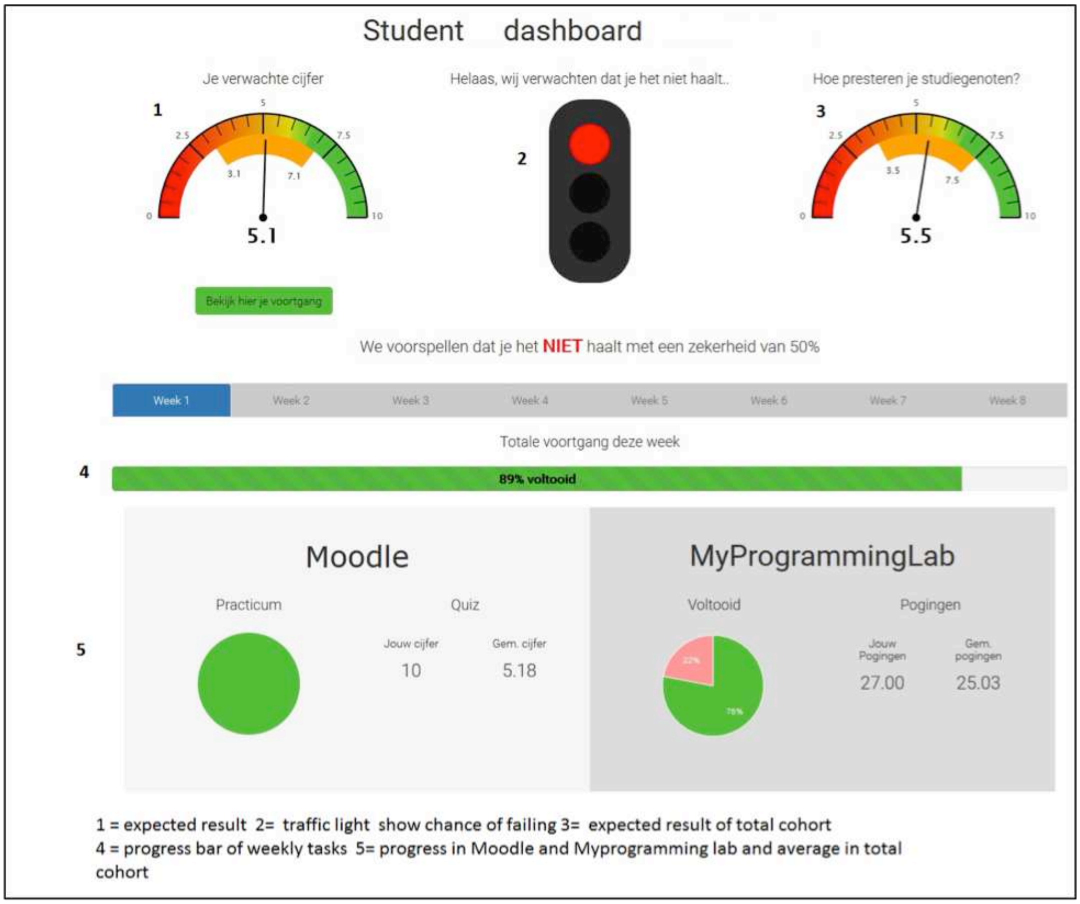

Figure 1. Student dashboard for Week 1

The purpose of the dashboard is to encourage the students to finish their online tasks. The effectiveness of a learning analytics dashboard is examined in the paper by (Lauría, Moody, Jayaprakash, Jonnalagadda, & Baron, 2013). This article gives the effect of the use of a so-called Early Warning system (EWS) on the study results. An EWS is used to identify high-risk students in a course (Hu et al., 2014) as early as possible with the aim of changing their learning behaviour. The students in courses with an EWS achieved on average a 6% higher final grade than the control group. An effect of the system is, that of the treatment group a larger group withdrew from the course: 25.6%, compared to 14.1% of the control group.

The Learning Analytics dashboard (Figure 1) was implemented on a weekly basis. The dashboard visualized the expected result and risk of failure for the student. For the expected result, linear regression models are used, and for the risk of failure decision tree algorithms (Decision Stump, Adaboost) (Hu et al., 2014) are used. These models were created through WEKA 3.6 . The failure risk models in week 8 correctly classified 88.8 % of the instances (607), 11.2% of the instances are incorrectly classified (77). The expected results are calculated with linear regression models of WEKA[3]. In week 8

[3] http://www.cs.waikato.ac.nz/ml/weka/

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

the model has a R^2= .41. The dashboard application was developed in PHP[4] with a MySQL[5] database and the prediction models were used in Kettle[6] to generate the predictions. A Perl script converted the data from Moodle and Myprogramminglab into Excel files. The Excel files are imported in SPSS and uploaded to the MySQL database of the dashboard application. The dashboard application generated the dashboards on the basis online behaviour data and the generated predictions (Figure 2).

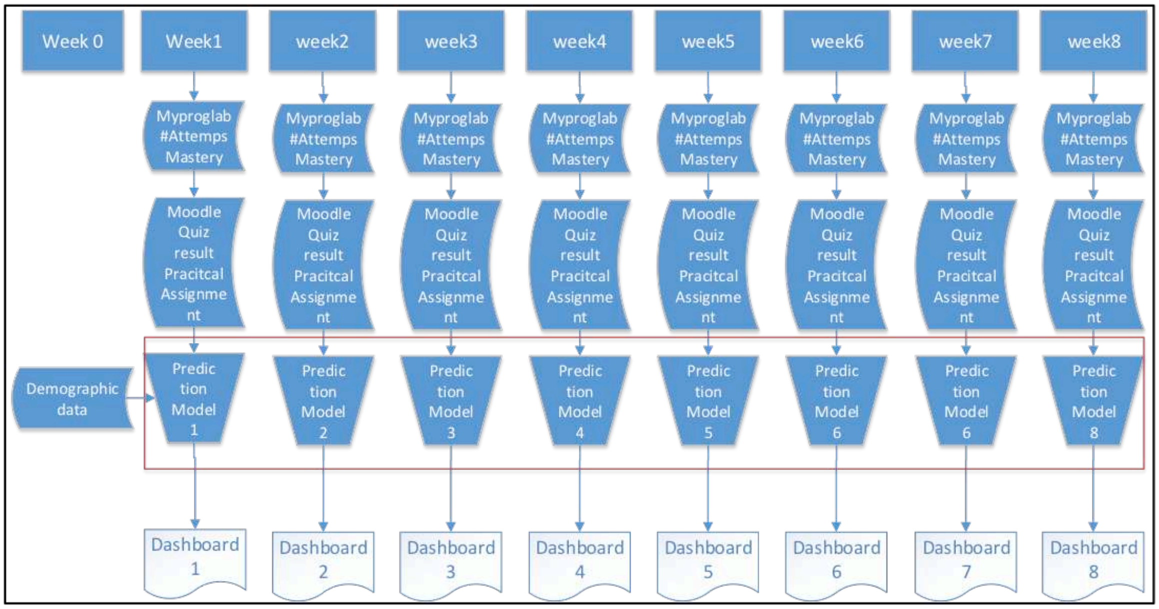

Figure 2. Visualization of dashboard generation

The treatment with the dashboard took place during the programming course in the first 10 weeks of the academic year 2015-2016. The exam was on 2 November 2015 and the retake on 10 December 2015. The course is to taught all computer science freshmen (n = 558) and given by 14 different teachers to 20 classes. The students had two lessons of two hours per week. This study gives insights into how a learning analytics dashboard has an effect on the student success rate in a programming course. Another contribution is that it shows if the dashboard has any effect on the online behaviour. This leads to the following research questions:

1. Will the learning analytics dashboard improve the success rate and the results of the students participating in the Java programming course?
2. Will the learning analytics dashboard increase the online activities of students?

---

[4] http://php.net/

[5] https://www.mysql.com/

[6] http://community.pentaho.com/projects/data-integration/

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

RESEARCH METHOD

The study is set up like a RCT and 556 students are involved in the experiment, with 276 in the treatment group and 280 in the control group (Figure 3). From previous cohorts (2013, 2014) it was known that the results of programming differ according to the specialization of the student. To achieve internal validity of the experimental design, the students were conditionally randomized per their specialization. The students from the treatment group received a mail with a link to their dashboard every week. In total the students received eight mails with a link to their dashboards. At the end of the course, the results of the exams were collected for both the treatment and control group.

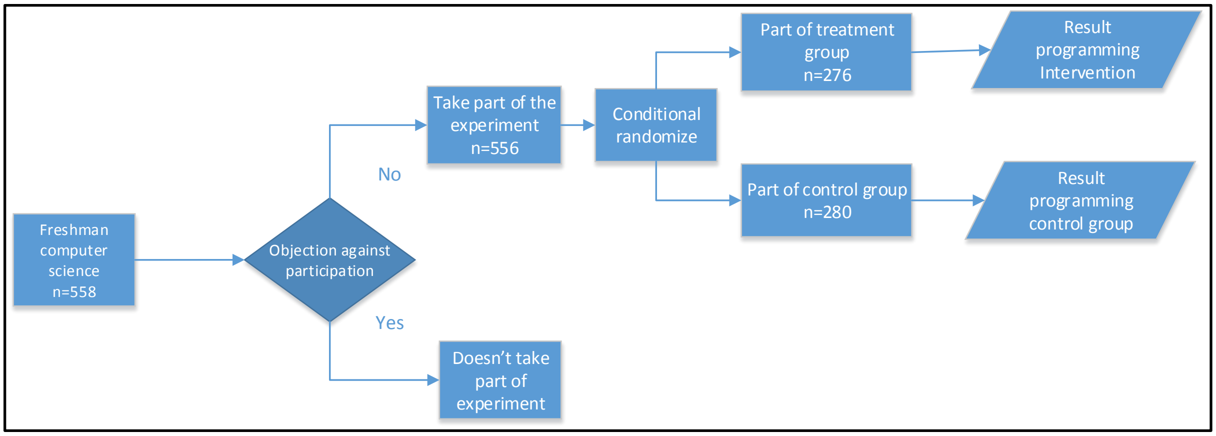

Figure 3. RCT setup for dashboard treatment

RESULTS

In this section, the results of the exams are evaluated. The article (Lauría et al., 2013) gave rise to analysis of whether the use of the dashboard had an effect on not participating in the exam. In Table 3, cross tables of the exam results are given. It indicates that 100 students of the total cohort (n = 556), which is 18%, did not take any exam. Between the means of the treatment – and control group there was no significant difference on whether or not to withdraw from the exam (Table 1).

Table 1. Means of students taking part of the exam: treatment and control group

	Control	Dashboard	n
Exam	.509	.491	456
No exam	.504	.496	100
Total	.505	.496	556

X^2 (1) = 0.27 p=0.602

The effect of the dashboard treatment on passing the programming course is shown in Table 2, where the means of passing the exams for the treatment and control groups are shown. This shows that for the treatment group there is a small improvement in the percentage passing at the first exam, which is not significant.

Of the entire cohort of 556 students, 332 passed the programming course. This is 59.7%. Of the 456 students who took the exam, 72.8% were successful.

Table 2. Means of passing: treatment and control group

Result	Control	Dashboard	n	p	χ2
First exam passed	.523	.533	228	.830	(1) .5
Retake passed	.155	.519	116	.599	(1) .28
All exam passed	.741	.714	456	.516	(1) .42

n=556. Missing =100. Passed: 332 Failed= 124

Table 3 shows the means of the grades of the exams. The marks are displayed in (0-100). The results in Table 3 show no significant differences between the control and treatment group.

Table 3. Means and standard deviation of the grades of the treatment and control group (n=556)

		n	M	SD	p
First exam	Control	220	57.95	28.91	.83
	Dashboard	212	57.38	27.45	
Retake exam	Control	110	57.53	26.06	.87
	Dashboard	106	56.92	26.71	
All exams	Control	232	68.62	24.91	.42
	Dashboard	224	66.71	25.13	

CONCLUSION

This study tried to increase the number of students passing a Java programming course by giving feedback on their online behaviour through a learning analytics dashboard. The treatment consisted of a randomized controlled experiment with 556 students, 276 students received for eight weeks an email with a link to their dashboard and 280 students were assigned to the control group. There was no significant difference in withdrawal between the control and treatment group. The analyses have shown no significant effect of offering the dashboard, on either the percentage passed or the mark of the exam.

An important aspect of the dashboard intervention is the use of the online environments by the students. The analyses of the online use did not provide any significant differences between the control- and the treatment group, although a small effect is found at Myprogramminglab exercises. Students from the treatment group passed about 5% more exercises on a significance level p = .1. The analysis showed that online activity strongly declined over time. This was due to the flip-the-classroom principle being no longer compulsory. In the academic year 2015 the students practised online significantly less compared to those in the academic year 2014. The dashboards will only be able to predict well if the student practise in the online environments, otherwise the dashboard will predict very badly. This may cause the low interest of the students in the dashboard at the end of the course. Maybe the dashboard should be incorporated in courses where a part of the course result is determined in the online environments. In these courses the students will practice more and the dashboard will be more functional.

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR

REFERENCES

Davies, R., Dean, D., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development*, *61*(4), 563–580. http://doi.org/10.1007/s11423-013-9305-6

Hu, Y.-H., Lo, C.-L., & Shih, S.-P. (2014). Developing early warning systems to predict students' online learning performance. *Computers in Human Behavior*, *36*(0), 469–478. http://doi.org/http://dx.doi.org/10.1016/j.chb.2014.04.002

Lauría, E. J. M., Moody, E. W., Jayaprakash, S. M., Jonnalagadda, N., & Baron, J. D. (2013). Open Academic Analytics Initiative: Initial Research Findings. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge* (pp. 150–154). New York, NY, USA: ACM. http://doi.org/10.1145/2460296.2460325

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2014). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior*. http://doi.org/10.1016/j.chb.2014.05.038

In: Ferguson, R., Sharkey, M. & Mirriahi, N. (eds) (2016). *Practitioner Track Proceedings of the 6th International Learning Analytics & Knowledge Conference (LAK16)*. University of Edinburgh, Edinburgh, UK (25–29 April). SoLAR