Learning Analytics Dashboard for Motivation and Performance

Damien S. Fleur[1,2(✉)][0000-0003-4836-5255], Wouter van den Bos[2,3][0000-0002-8017-3790] and Bert Bredeweg[1,4][0000-0002-5281-2786]

[1] Informatics Institute, University of Amsterdam, Amsterdam, The Netherlands
[2] Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands
[3] Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany
[4] Amsterdam University of Applied Sciences, Amsterdam, The Netherlands
{d.s.fleur, w.vandenbos, b.bredeweg}@uva.nl

Abstract
Deploying Learning Analytics that significantly improve learning outcomes remains a challenge. Motivation has been found to be related to academic achievement and is argued to play an essential role in efficient learning. We developed a Learning Analytics dashboard and designed an intervention that relies on goal orientation and social comparison. Subjects can see a prediction of their final grade in a course as well as how they perform in comparison to classmates with similar goal grades. Those with access to the dashboard ended up more motivated than those without access, outperformed their peers as the course progressed and achieved higher final grades. Our results indicate that learner-oriented dashboards are technically feasible and may have tangible benefits for learners.

Keywords: Learning Analytics · Motivation · Social Comparison · Goal Orientation

1 Introduction

While the potential of Learning Analytics (LA) for improving learning has been shown, for example through the use of dashboards (e.g. [1, 2]), strong evidence in terms of improved learning behavior and learning outcomes remains scarce. Only few studies reported significant improvements on these two factors [3], whereas most studies evaluate the success of their dashboards on usability and perceived usefulness. Recently, LA Dashboards (LAD) researchers have called for an integration of theories and models from learning sciences into the development and evaluation of LA, and for learner-centered designs [4].

In our study, we investigate to what extent a LAD designed around social comparison and goal orientation can be successful at increasing motivation and learning outcomes in the context of a university course.

2 Theoretical background

Self-regulated learning theory argues that motivation plays an essential role in academic achievement and in the self-regulatory behavior of learners during their study (e.g. planning, time management, learning strategies) [5], which is supported by empirical evidence (e.g. [6]). This can in part be explained by the fact that highly motivated students tend to be more attentive to their learning outcomes and tend to put more effort into their learning, compared to poorly motivated ones [7].

Social comparison and goal orientation have both been found to stimulate motivation and performance [8, 9]. According to the former, humans have an inherent drive to evaluate their abilities which is done through the comparison with others [10]. Comparison is typically made upwards (i.e. with better peers), in particular when threat to self-esteem is absent [11]. The individual is then pressured towards uniformity, in particular when peers are perceived as close, both in terms of ability and relatedness [11]. By setting goals and working towards them, students tend to become more motivated and have higher learning outcomes. Goal orientation is particularly effective when goals are specific, reachable and when feedback is given which shows progress in relation to the goals [12].

Therefore, a LAD intervention relying on social comparison and goal orientation needs to satisfy the following criteria to be effective: (*i*) the learner must be proposed to set a goal for the course; (*ii*) regular feedback on the learner's progress towards that goal needs to be given; (*iii*) the LAD needs to provide relevant social comparison to increase motivation; (*iv*) relatability needs to be ensured. In our study, we primed students to set a goal grade for the course. Based on the evidence discussed above, we hypothesize that comparing one's progress with those of peers with similar grade goals increases motivation to reach the set goal. It also acts as relevant and relatable feedback. Importantly, comparison needs to be made slightly upward and the number of peers is kept low. A prediction of the final grade can provide additional feedback.

3 Design and implementation

We developed a LAD designed to stimulate students' motivation to perform better (source code: github.com/UvA-FNWI/coach3). The LAD was integrated within the course details in a Learning Management System (LMS, Canvas) with help of the available API and LTI libraries. Subjects could access it via a link on the course-homepage and via a button the menu. The dashboard itself consisted of two visualizations. To generate the visualizations, three datasets were used: The first dataset concerned data processed from the LMS which were all the grades of the course published so far of the students who consented to sharing their data. The second dataset contained the grades of every assignment given in the last two years in which the course was given (332 samples). The third dataset consisted of a list of permissions to access a student's grades on the LMS, whether students had access to the dashboard and their goal grade. Each time a subject received a new grade via the LMS, the dashboard was updated with new visualizations.

Visualization 1 (Fig. 1, LHS) shows the subject's current average grade in the course together with the current average grade of 9 anonymous peers who have set similar goal grades as the subject. The sample of peers is selected such as to elicit a slight upward comparison. The sample satisfies the following criteria: (*i*) the average grade of the sample is 0.5 to 1.5 points higher than the average grade of the subject, (*ii*) 20-40% of the sample has a lower average grade than the subject, and (*iii*) 30-50% of the sample has a lower or equal average grade as the subject. The algorithm used to generate the sample is a variation of the 'knapsack algorithm' [13] and uses the computed average grades of the students in the course. Candidates for the comparison sample are students whose goal grade is equal to the goal grade of the subject within a tolerance margin. This margin is increased until a sample can be generated. In the case that no sample can be made, the comparison sample is based on the 9 closest peers in terms of average grade. Such edge cases typically apply for the subjects performing either among the top 9 or bottom 9 students in the course, or when the sample-generation algorithm timed out. Note that the subjects were oblivious to the manipulation.

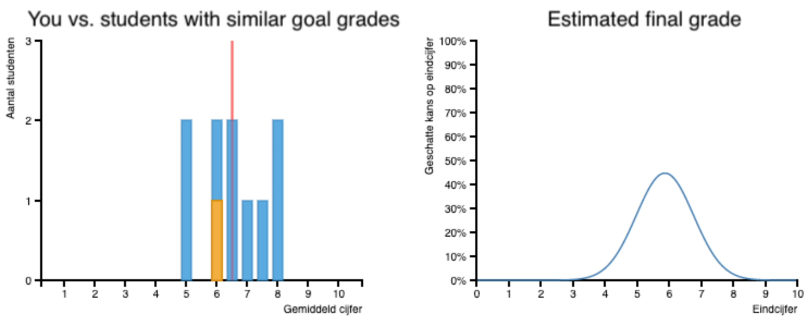

Fig. 1. Example visualization in the LAD (early in the intervention). Left: the subject's average grade (in orange) is compared to the average grade of a selected sample of peers (in blue) with similar goal grades. The red line represents the average grade of the sample of peers. Right: Estimation of the student's final grade as a normal distribution. The x-axis represents the possible final grades (0-10) and the y-axis represents the estimated probability of getting any of the grades.

Visualization 2 (Fig. 1, RHS) shows a prediction of the student's final grade based on the grades he or she has received in the course so far. The prediction is made with a Bayesian Ridge Regression with the implementation proposed in [14], with hyperparameter $\alpha_1=\alpha_2=\lambda_1=\lambda_2=10^{-6}$ (in particular, the BayesianRidge function in the Scikit-Learn package in Python was used). The Bayesian regressions have the advantage that they provide information about the uncertainty of the prediction. Both predicted mean and uncertainty are used to draw the normal distribution. In addition, this method was chosen over an algorithm proposed by Meier and colleagues [15]. Meier's algorithm was designed to timely estimate a student's final grade based on the grades of previous years' students. It resembles a K-nearest neighbors algorithm and also provides with an uncertainty of the prediction. Although this algorithm was argued to be most appropriate for this type of data, the Bayesian regression resulted in lower root-mean-square error for timed predictions. Figure 2 shows the prediction error of the two algorithms trained on the previous year's grades with final grades as target values and tested on the treatment group's final grades, which validate the choice for the Bayesian regression. Note that the confidence threshold included in Meier's algorithm was left out because the intervention design required that subjects be able to see their predicted final grade at all stages of the course.

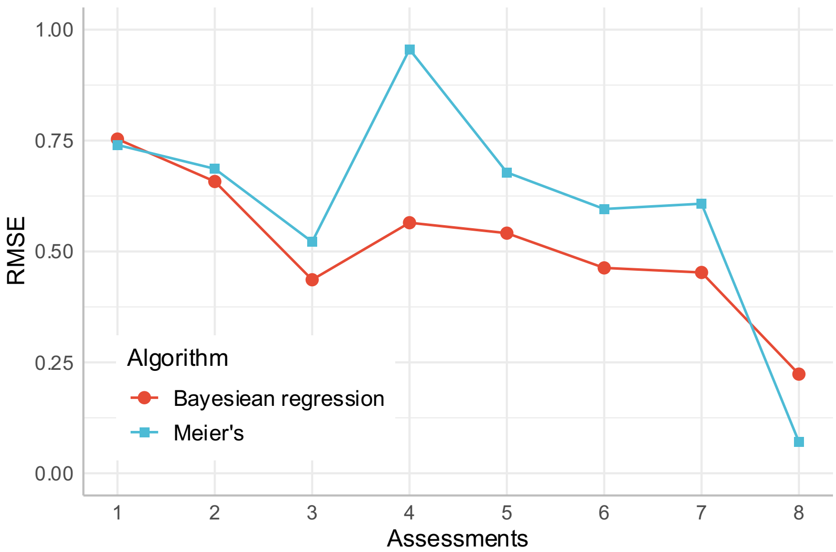

Fig. 2. Root-Mean-Square Error (RMSE) of the Bayesian regression (red) and Meier's algorithm [15] (blue) when testing on the treatment group's final grades. Meier's algorithm is set up with first radius r_1=14, which gives the lowest RMSE. A new prediction is made every time a subject receives an assessment (i.e. a new grade).

4 Results

First year Bachelor students were recruited (n=79) and randomly assigned to either the treatment group (access to LAD) or the control group (no access). The control group was told that enough participants were gathered for the experiment with the LAD but that they would still earn a compensation sharing their data and filling in questionnaires. There was no further experimental difference between the two groups. 7 subjects were excluded from the analysis because they either left the experiment or dropped out of the course, resulting in 72 subjects included in the analysis (34 treatment subjects). All subjects received a monetary compensation (€7.5) for their participation.

The intervention was carried out for the duration of the course (8 weeks). At the start and at the end of the course, all subjects were asked to fill in a self-report questionnaire. The subscales of the Motivated Learning Strategies Questionnaire (MSLQ) [16] were used to evaluate extrinsic and intrinsic motivation. A 19-item subset of the Metacognitive Awareness Inventory (MAI) was used to evaluate metacognitive knowledge and metacognitive control [17]. 26 subjects (of which 16 treatment) filled in the questionnaires at both time points. The difference between the total number of subjects and those who filled in the questionnaire is addressed in the discussion section. Additionally, at the start of the course, all subjects were asked to set a goal grade that they desired to achieved for the course (1-10, with passing grade at 6). They were given the opportunity to update their goal grade half-way through the course. 6 individual assessments and 2 group assessments were given and spread out more or less evenly throughout the 8 weeks in which the course took place.

To assess for the effect of the tool on motivation, a linear mixed-effects analysis was done in R (version 3.5.2; R Core Team, 2018) using the R package 'lme4' [18], with the two-level factors 'group' (treatment/control) and 'time' (begin/end). The latter factor reflects the moment in the course at which motivation was evaluated. Random intercepts for subjects were included. The outcome variables were the scores obtained in the MSLQ for extrinsic motivation and intrinsic motivation. The full code of the analysis is available via the authors' figshare repository (https://doi.org/10.21942/uva.12053676.v1).

We found a significant interaction effect of 'group' and 'time' (β = -1.07, *S.E.* = 0.42, t = 2.54, p = 0.01) for extrinsic motivation with both groups showing similar scores at the start of the course, and the treatment group showing increased extrinsic motivation at the end of the course, while it decreased in the control group (Fig. 3). This confirms the hypothesis that the intervention increases motivation. With regards to intrinsic motivation, while the scores significantly decreased overall (β = -0.62, *S.E.* = 0.19, t = -3.27, p < 0.01), the groups did not evolve in significantly different ways.

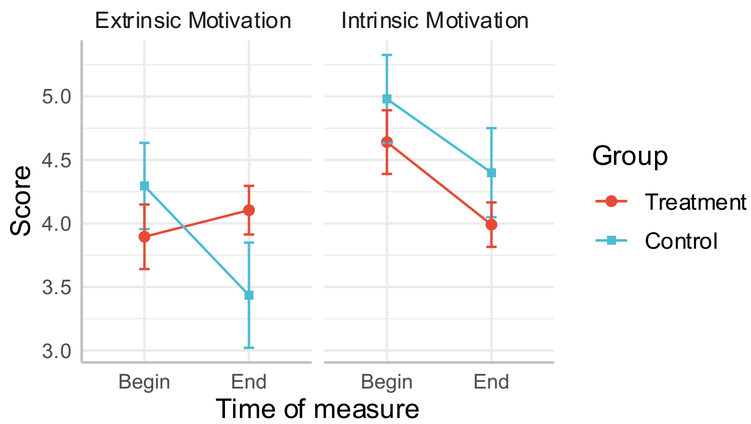

Fig. 3. Mean change (with standard error) in extrinsic and intrinsic motivation between start and end of the course for the treatment (red) and control group (blue).

We analyzed the effect of the intervention on performance in two ways: (*i*) effects on final grades, and (*ii*) how performance evolved for similar assessments throughout the course. We identified two assessment types: (1) formative assessments in the form of homework; (2) summative assessments in the form of midterm exams. Type 1 included two assessments that were given in the first half of the course. Type 2 included two assessments that were given in the middle and at the end of the course. The same analysis as for motivation was performed with the homework and midterm as outcome variables. For final grades, a linear regression was performed with "type" as independent variable and final grades as outcome variable. The treatment group achieved a higher final grade than the control group on average ($\beta = 0.36$, $S.E. = 0.17$, $t = 2.07$, $p = 0.04$; Fig. 4a).

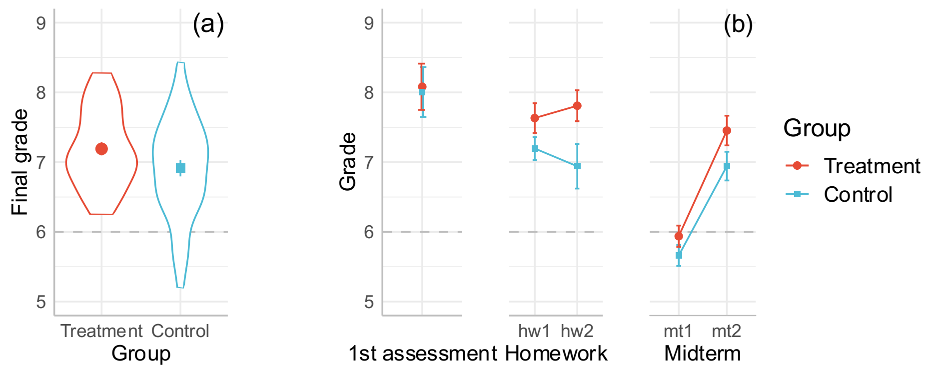

Fig. 4. (a): Final grades in treatment (red) and control group (blue). (b): Mean change in grade (with standard error) per type of assignment for treatment and control group.

Regarding performance on homework, the treatment group also performed better overall than the control group ($\beta = 0.60$, $S.E. = 0.28$, $t = 2.20$, $p = 0.03$). Fig. 4b (middle pane) shows a trend of the treatment group's performance in homework increasing whereas the control group's performance decreased. A similar trend of the treatment group ending up outperforming the control group can also be seen for the midterm assessments (Fig. 4b, right pane). However, neither of these trends were found to be significant. As for comparison, both groups performed similarly on the very first assessment of the course (Fig. 4b, left pane), showing that both groups started the course on similar bases. No relation was found between changes in (extrinsic and intrinsic) motivation and final grade, change in homework performance or change in midterm performance. This was the case both between groups and in general.

5 Discussion and conclusion

A LAD was developed that relies on social comparison and goal orientation to stimulate motivation in learners and, by extension, improve performance. As we hypothesized, the results indicate that the LAD positively influences extrinsic motivation and that subjects with LAD access perform better over time. Put together, our results indicate that a LAD designed around goal orientation and social comparison in the classroom is technically feasible and can have tangible benefits for learners.

In particular, while the extrinsic motivation of control subjects fell, treatment remained stable if slightly increasing. Intrinsic motivation, on the other hand, declined in both groups. Such general decrease in motivation is not surprising and may be related to a novelty effect which has been reported in various contexts [19–21]. Indeed, the intervention took place in the first course offered in the subjects' curriculum and we may assume that most students were excited to begin a new study. As time passes, the novelty effect dwindles. The LAD and the introduced manipulation appear to have mitigated this decrease in the case of extrinsic motivation. However, the intervention did not seem to support intrinsic motivation.

While the intervention had effects on extrinsic motivation and performance, the relationship between the two could not be identified. We propose that motivation and performance are indirectly related in that increased motivation would lead to increased engagement and effort which would translate in higher grades. Research in social comparison in MOOCs did indeed report increases in engagement [2]. There are also various accounts for the relation between goal orientation, effort and performance [12]. This idea could be investigated in future research by collecting data on students' learning behavior, such as time dedicated to study and engagement with the learning material on the LMS.

Although filling in questionnaires was a mandatory part of the intervention, we dealt with low compliance in this regard. This can be partly explained by the logistical challenges that we faced. The questionnaires were filled out online using the students' private devices during the tutorial sessions. In these sessions, the students were spread across different classrooms. Although, the Teaching Assistants gave the tutorials and oversaw the process, they may not have been able to be sufficiently present in all places.

The data used in the analysis only included subjects who filled out and completed the questionnaire at both time points. The low compliance resulted in more data on motivation from the treatment group than the control group. As a result, the effects on motivation reported above may be less robust for the control group. Lastly, the sample obtained may not be fully representative of the whole experimental population. It could be argued that only highly motivated subjects filled out the questionnaires which could bias the scores upwards.

From a technical point of view, the presented approach requires the availability of grade records and a course with a sufficient number of assignments, such that the visualizations can be often updated. Given that coursework is migrating to LMSs, and with the advent of Massive Online Open Courses (MOOCs) there will be numerous opportunities to apply similar interventions in the future. Regardless, we did not find evidence that performance and motivation were directly related to how many times subjects opened the dashboard. This may indicate that students beneficiate equally from the dashboard regardless of the degree to which they use it and, thus, that the frequency of assessments may not be essential.

In this study, we laid the groundwork for a simple intervention that can boost students' motivation and academic outcomes. Building upon these early findings, future research can improve efficacy and understanding of the relation between motivation and performance. In terms of possible improvement for this study, making study-related questionnaires an integral part of the course would improve student compliance and data collection, while larger samples of students would strengthen our findings. Proposed ideas for future research include having different types of treatment groups (e.g. access to one or both visualizations) and obtaining further insights on the evolution of students' motivation.

Reference

1. Brouwer, N., Bredeweg, B., Latour, S., Berg, A., van der Huizen, G.: Learning Analytics Pilot with Coach2 - Searching for Effective Mirroring. In: Verbert, K., Sharples, M., and Klobučar, T. (eds.) Adaptive and Adaptable Learning. pp. 363–369. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-45153-4_28
2. Davis, D., Jivet, I., Kizilcec, R.F., Chen, G., Hauff, C., Houben, G.-J.: Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In: Proceedings of the Seventh International Learning Analytics & Knowledge Conference. pp. 454–463. ACM (2017). https://doi.org/10.1145/3027385.3027411
3. Bodily, R., Verbert, K.: Trends and issues in student-facing learning analytics reporting systems research. In: Proceedings of the seventh international learning analytics & knowledge conference. pp. 309–318. ACM (2017). https://doi.org/10.1145/3027385.3027403
4. Gašević, D., Dawson, S., Siemens, G.: Let's not forget: Learning analytics are about learning. TechTrends. 59, 64–71 (2015). https://doi.org/10.1007/s11528-014-0822-x
5. Schunk, D.H., Zimmerman, B.J.: Motivation and self-regulated learning: Theory, research, and applications. Routledge (2012)
6. Singh, K., Granville, M., Dika, S.: Mathematics and science achievement: Effects of motivation, interest, and academic engagement. J. Educ. Res. 95, 323–332 (2002). https://doi.org/10.1080/00220670209596607
7. Schunk, D.H., Zimmerman, B.J.: Motivation an essential dimension of self-regulated learning. In: Motivation and self-regulated learning. pp. 13–42. Routledge (2012)
8. Lee, L.H.: Goal orientation, goal setting, and academic performance in college students: An integrated model of achievement motivation in school settings, (1998)
9. Schokker, M.C., Keers, J.C., Bouma, J., Links, T.P., Sanderman, R., Wolffenbuttel, B.H.R., Hagedoorn, M.: The impact of social comparison information on motivation in patients with diabetes as a function of regulatory focus and self-efficacy. Health Psychol. 29, 438–445 (2010). https://doi.org/10.1037/a0019878
10. Festinger, L.: A theory of social comparison processes. Hum. Relat. 7, 117–140 (1954)
11. Gerber, J.P., Wheeler, L., Suls, J.: A social comparison theory meta-analysis 60+ years on. Psychol. Bull. 144, 177 (2018). https://doi.org/10.1037/bul0000127
12. Locke, E.A.: Motivation through conscious goal setting. Appl. Prev. Psychol. 5, 117–124 (1996). https://doi.org/doi.org/10.1016/S0962-1849(96)80005-9
13. Dantzig, T.: Number, The Language of Science. Revised and Augmented. J. Nerv. Ment. Dis. 92, 546 (1940)
14. Tipping, M.E.: Sparse Bayesian learning and the relevance vector machine. J. Mach. Learn. Res. 1, 211–244 (2001)
15. Meier, Y., Xu, J., Atan, O., Schaar, M. van der: Predicting Grades. IEEE Trans. Signal Process. 64, 959–972 (2016). https://doi.org/10.1109/TSP.2015.2496278
16. Duncan, T.G., McKeachie, W.J.: The making of the motivated strategies for learning questionnaire. Educ. Psychol. 40, 117–128 (2005). https://doi.org/10.1207/s15326985ep4002_6
17. Harrison, G.M., Vallin, L.M.: Evaluating the metacognitive awareness inventory using empirical factor-structure evidence. Metacognition Learn. 13, 15–38 (2018). https://doi.org/10.1007/s11409-017-9176-z

18. Bates, D., Mächler, M., Bolker, B., Walker, S.: Fitting linear mixed-effects models using lme4. ArXiv Prepr. ArXiv14065823. (2014)
19. Berlyne, D.E., Craw, M.A., Salapatek, P.H., Lewis, J.L.: Novelty, complexity, incongruity, extrinsic motivation, and the GSR. J. Exp. Psychol. 66, 560 (1963)
20. Mutsuddi, A.U., Connelly, K.: Text messages for encouraging physical activity Are they effective after the novelty effect wears off? In: 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops. pp. 33–40 (2012). https://doi.org/10.4108/icst.pervasivehealth.2012.248715
21. Jeno, L.M., Vandvik, V., Eliassen, S., Grytnes, J.-A.: Testing the novelty effect of an m-learning tool on internalization and achievement: A Self-Determination Theory approach. Comput. Educ. 128, 398–413 (2019). https://doi.org/10.1016/j.compedu.2018.10.008