Towards Adaptive Social Comparison for Education

Sergey Sosnovsky$^{(\boxtimes)}$, Qixiang Fang , Benjamin de Vries, Sven Luehof, and Fred Wiegant

Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands
{s.a.sosnovsky,q.fang,b.e.devries,s.luehof,f.a.c.wiegant}@uu.nl

Abstract. Informing students about their progress in comparison to their peers has been widely used in educational research as a strong motivational factor, effective gamification technique and means for adaptive guidance to learning material. A typical social comparison interface helps students weight their individual levels against the average levels of other students. However, such uniform approach may not be effective for every category of students and every learning situation. Underachieving students might find the displayed social goal impossible, while overachieving students might decide that the learning goal has been attained and stop investing time and efforts. An alternative approach is an adaptive social comparison strategy that chooses different levels of the social goal for different categories of students. This paper presents one of the first steps towards developing such a strategy.

Keywords: Social comparison · Learning analytics

1 Introduction

When students engage in self-regulated learning, they often experience difficulties with identifying strong and weak areas of their knowledge, planning their learning, and finding useful educational resources [2]. A large class of technologies supporting students in such a context is based on providing them with information about their own progress. Several positive effects of letting students view and interact with indicators of their progress have been discovered including improvement in self-reflection and planning, higher levels of engagement with educational content and increase in trust in the educational tool [3].

More recently, researchers have begun to combine indicators of individual progress with social information reflecting average progress of students' peers thus promoting social comparison (SC) and using it as a vehicle to further increase motivation and facilitate navigation. Bretzke and Vassileva have explored SC from the point of building learning communities where members are aware of each other's status (defined by learning activity) and strive to raise their own status [1]. Domínguez et al. have focused on the gamification aspects of SC [4]. They have developed a BlackBoard plugin using badges and leaderboards

and observed an overall positive effect on motivation; yet, they have also noticed that a large category of students have found the system demotivating as they have not found it "fun to compete with their classmates". Open social student modelling systems allow students to compare their progress against the progress of their peers using structured representations of knowledge broken into concepts and topics. Often, such systems use student models as navigational tools guiding students to supplementary learning resources [9].

Overall, this type of tools aims at leveraging our general tendency to view the world through a social lens and self-evaluate ourselves, "our opinions and abilities", with others [6]. However, while the overall tendency is undeniable, social psychologists have observed over the years that the amount, direction (upwards versus downwards) and impact of SC varies depending on different individual factors [8]. Yet, the existing technology-enhanced learning (TEL) systems leveraging SC predominantly implement one single level of social goals corresponding to the average level of the group. As a result, for overachieving students, SC interfaces provide little motivational value once they surpass the course average. At the same time, underachieving students might find the uniform social goal unreachable and the interface utilising it discouraging. Finally, some students can react stronger to other motivational factors besides performance orientation.

An alternative approach is an adaptive SC strategy that optimises the social goal for individual students. This research presents the first step in this direction. We have implemented a student-facing dashboard that allows students to choose a level of SC. We have examined the influence of different initial SC levels on students' activities and learning. We focused on three research questions: What are the preferences of SC levels among students? What are the effects of different SC levels on students' activities and learning outcomes? Are there sub-groups of students that respond deferentially to different SC levels?

2   StudyLens Platform and Its SC Interface

StudyLens is a learning analytics software developed at Utrecht University (UU). It implements a suite of interfaces visualising results of learning activity for both teachers and students. StudyLens has been integrated with several educational tools used in UU courses. In this project, we used its integration with the digital assessment system Remindo. The project was conducted within the framework of one undergraduate course on Evolutionary Biology. During the first three weeks of the course, students were offered a set of six formative tests on crucial concepts in Evolutionary Biology. The tests were based on validated inventories of multiple choice questions [7].

Once students took formative tests in Remindo, their results were sent to the learning record store of StudyLens, aggregated and visualised. In order to support meaningful representation of student knowledge in StudyLens, test questions were associated with the concepts and misconceptions they diagnose. A knowledge level for an individual concept (misconception) was computed as an average score of all questions associated with it. The overall domain model of

the target part of the course consists of 7 misconceptions and 21 concepts (on two levels of hierarchy) and was used as the framework for visualising student knowledge. Figure 1 shows a screenshot of StudyLens. Student's progress is visualised for both, concepts (blue bars) and misconceptions (red bars). Individual progress can be compared against the SC indicators (grey bars). Once a student clicks on a concept or a misconception, links to corresponding learning material are provided below the charts. Finally, above the charts, buttons for selecting desired levels of SC are displayed. Students could freely choose to be compared with the "average" course progress, the progress of "upper 50%" of students in the class or the progress of "lower 50%" students in the class.
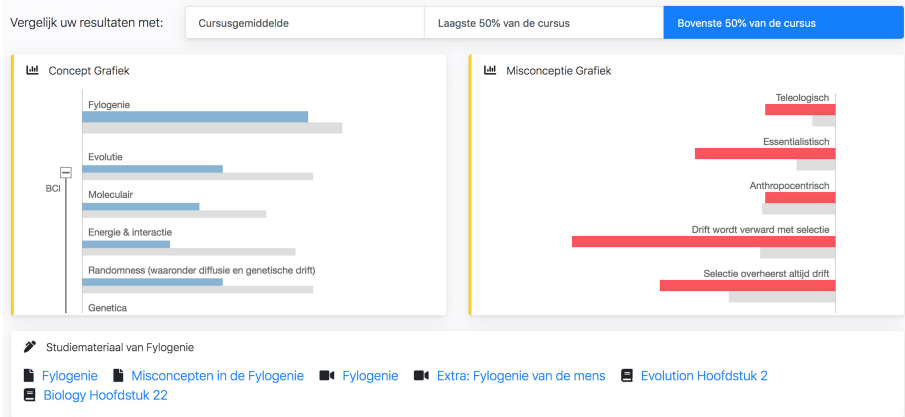

Fig. 1. SC dashboard of StudyLens (Color figure online)

3   Evaluation

We adopted a within-subject experimental design, where we divided the entire duration of the course into six periods (based on the number of formative tests) and the 30 students in the class into three SC level groups. In each period, students in the same group started their StudyLens sessions with the same SC level. After every week, the groups alternated. This way, students were equally exposed to all starting SC levels over time, while still being able to freely choose and switch SC levels. At the end of the course, the students were asked to fill in a questionnaire measuring their achievement goal orientation [5].

An average student spent 94.10 s on the "course average" SC level, 84.66 s on the "upper 50%" and 36.20 s on the "lower 50%" per session.

Considering the hierarchical data structure (i.e. session records clustered within subjects), we used multilevel regression models to study the relationships between the manipulated factor (i.e. starting SC levels) and three outcome variables: the amount of time spent on StudyLens during a session, the number

of clicks on learning materials during a session and the amount of change in the upcoming test or exam grade.

First, using multilevel Gaussian regression, we found that being assigned to the "lower 50%" level, compared to the "average" level, led to on average 3.21 min longer time spent on a session, after controlling for individual test grades ($B = 3.21, t(134.41) = 2.83, p < .01$). The "upper 50%" level, compared with the "average" level, also predicted a sizeable (but statistically insignificant) increase of 1.48 min in the amount of time spent in a session ($p > .05$).

Second, using multilevel Poisson regression, we found that being assigned to the "lower 50%" level led to on average 2.58 times the rate of clicking study materials for the "average" level ($B = 0.95, e^B = 2.58, z = 6.37, p < .01$), while controlling for individual test grades. In contrast, the difference in the rate of clicking study materials between the "upper 50%" and the "average" level was small and insignificant ($B = 0.05, e^B = 1.05, z = 0.28, p > .05$).

Lastly, we tested whether previous exposure to a specific starting SC level would influence the next test grade of a student. To do so, we matched each test grade with the immediately previous preset SC level. Using multilevel Gaussian regression, we found that, in comparison to those starting with the SC level "average", students starting with "upper 50%" score about 0.58 lower in an upcoming test (while controlling for test difficulty) ($B = -0.58, t(83.01) = -2.28, p < .05$). In contrast, the effect of "lower 50%" (compared with "average") on test grades was small and insignificant ($B = -0.01, t(86.15) = -0.03, p > .05$).

We highlight two interesting findings from the subgroup analysis. First, we can see from Fig. 2A that when students with high (i.e. above-average) test scores spent the least amount of time in StudyLens when they were assigned the "average" SC level, compared to the other SC levels. In contrast, students with low (i.e. below-average) test performance spent the least amount of time when they were assigned the "upper 50%" level. Second, in Fig. 2B and 2C, students with high performance approach or avoidance orientation clicked study materials the most when they were assigned the "lower 50%" level, while those with low performance (approach or avoidance) orientation engaged with the study materials the most when they started with the "upper 50%" level. Despite these interaction effects being not statistically significant, they are sizeable.
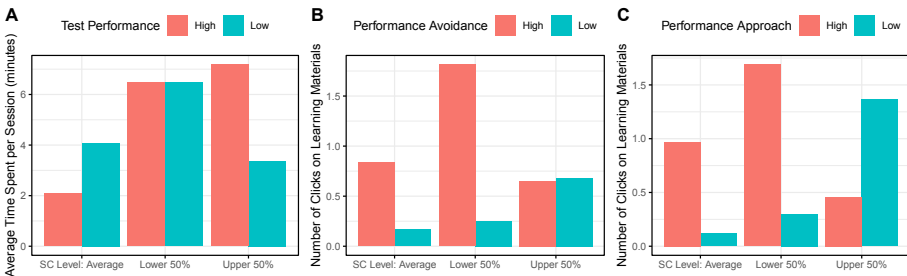

Fig. 2. Interaction between SC levels and student subgroup status on outcomes

4 Conclusion

There are three important findings. First, in terms of SC preferences, students favoured the "average" SC level the most, followed by "upper 50%" and lastly "lower 50%". Second, we have observed that the starting SC level substantially influence students' activities and learning. Specifically, students who started with "lower 50%" spent more time in StudyLens, clicked on more learning materials and did not drop in test grades (compared to students who started with the "upper 50%" SC level). Finally, we have found evidence for differential responses of different student subgroups to different SC levels.

These findings support two important claims. First, students' SC preferences do not necessarily align with their best interests (e.g. learning outcomes). Students might favour "average" SC level the most; however, they seem more engaged with learning materials when they started with "lower 50%". On the other hand, while students might prefer "upper 50%" over "lower 50%", starting out with "upper 50%" resulted in lower grades the next time. Second, for different subgroups of students, different SC levels maybe more optimal. Variables like test performance and performance orientation can play a role here. For instance, when the goal is to engage with study materials, it might be best if we assign above-average performance-oriented students to the "lower 50%" starting level and below-average performance-oriented students to the "upper 50%" level. On the other hand, when the goal is to make students explore the learning platform longer, we should avoid assigning students with above-average test scores to the "average" SC level or students with below-average test scores to the "upper 50%" SC level. Overall, more work is required to identify a pathway towards building an adaptive SC strategy.

References

1. Bretzke, H., Vassileva, J.: Motivating cooperation on peer to peer networks. In: Brusilovsky, P., Corbett, A., de Rosis, F. (eds.) UM 2003. LNCS (LNAI), vol. 2702, pp. 218–227. Springer, Heidelberg (2003). https://doi.org/10.1007/3-540-44963-9_30
2. Broadbent, J., Poon, W.L.: Self-regulated learning strategies & academic achievement in online higher education learning environments: a systematic review. Internet High. Educ. **27**, 1–13 (2015)
3. Bull, S.: There are open learner models about!. IEEE TLT **13**(2), 425–448 (2020)
4. Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., Martínez-Herráiz, J.J.: Gamifying learning experiences: practical implications and outcomes. Comput. Educ. **63**, 380–392 (2013)
5. Elliot, A.J., Church, M.A.: A hierarchical model of approach and avoidance achievement motivation. J. Pers. Soc. Psychol. **72**(1), 218–232 (1997)
6. Festinger, L.: A theory of social comparison processes. Hum. Relat. **7**(2), 117–140 (1954)
7. Furrow, R.E., Hsu, J.L.: Concept inventories as a resource for teaching evolution. Evo. Educ. Outreach **12**(1), 1–11 (2019). https://doi.org/10.1186/s12052-018-0092-8

8. Gibbons, F.X., Buunk, B.P.: Individual differences in social comparison: development of a scale of social comparison orientation. J. Pers. Soc. Psychol. **76**(1), 129 (1999)
9. Hsiao, I.H., Brusilovsky, P.: Guiding and motivating students through opensocial student modeling: lessons learned. Teach. Coll. Rec. **119**(3), 33 p. (2017)