Impact of Open Social Student Modeling on Self-Assessment of Performance

Sibel Somyürek, Peter Brusilovsky
Computer Education and Instructional Technologies Department, Information Science Department
Gazi University, University of Pittsburgh
Turkey, United States
ssomyurek@gazi.edu.tr, peterb@pitt.edu

Abstract: This study examines the impact of Open Student Modeling (OSM) and its extension known as Open Social Student Modeling (OSSM) on students' self-assessment of their SQL programming performance. It also explores the relationship between self-assessment and normalized gain. The study was performed with graduate students of University of Pittsburgh taking Database Course for the first semester 2014/2015 and ran for 11 weeks. The results demonstrated that both OSM and OSSM positively affect students' regular self-assessment ability. However, only OSSM was able to positively impact students' relative self-assessment ability (i.e., compare their SQL knowledge to others). The results also indicated that, students' self-assessments are positively and highly correlated to their normalized gain.

Introduction

Self-assessment toward individuals' learning and their performance is one of the important competencies for people that lead them to develop a correct self-image. The study of Kruger and Dunning (1999), which was awarded the 2000 satirical Ig Nobel Prize in Psychology, focused on the people's lack of self-assessment skills and their tendency of overestimating their abilities. Although their study is one of the first successful scientific studies focused on "unskilled and unaware people", similar phenomena is put into words by several philosophers and have been existing in proverbs of several nations for a long time. Kruger and Dunning (1999) stated that people are generally unaware of their incompetence, especially poor performers has illusory superiority. They interpreted this situation as the deficiency of metacognition of unskilled individuals that hinder their recognition about their real knowledge and performance level in comparison to others.

Metacognition, "thinking about thinking" is a fundamental skill for academic performance (Garner & Alexander, 1989). Metacognitive knowledge which is one of two main components of metacognition, refers to the ones knowledge about her/his cognitive processes and how to control them (Flavel, 1979; Jubran et al., 2014). Self-assessment is a part of metacognitive knowledge that enables being aware about one's ability level (Stankov & Crawford, 1996; Pintrich, 2002). Having ability to monitor learning is seen as a trigger to become an independent learner and lead students to take their own responsibility (Garrison, 1997; Borkowski et al, 1990).

Arguably, E-learning systems could provide critical help to students in improving their learning monitoring and self-assessment abilities. Especially promising in this aspects are adaptive e-learning systems. These systems include a *student model* which is the internal representation of student knolwegde used by personalized systems to monitor their student and to adapt learning activities them (Kay, 2000; Somyürek, 2009). To open this knowledge monitoring ability to students themselves, some adaptive system use Open Student modeling (OSM) making some parts of student model explicit. The research shows that OSM could support some important meta-cognitive abilities (Bull & Kay, 2013) although very little work has done specifically to determine the impact of OSM on students' self-assessment ability.

Even more promising in the sense of self-assessment is a more recent extension of OSM approach known as Open Social Student Modeling (OSSM) (Hsiao et al. 2013; Loboda et al., 2014). This approach suggest to visualize not just student own knowledge, but also knowledge of other student and peers in a form that enables the students to compare themselves with others. OSSM is motivated by Social Comparison theory Festinger (1954), which states that "self-knowledge is fulfilled not just getting information about oneself but also comparing oneself to another" which is thought is a critical aspect of social interaction (Buunk & Gibbons, 2007; Brickman & Bulman, 1977, p.50).

This paper presents an attempt to investigate to what extent metacognitive promises of OSM and OSSM are fulfilled, in other words, to what extent these technologies could improve self-assessment (which is an important part of self-awareness). We report the results of a study that engaged students to work with an adaptive system extended with OSM and OSSM and assessed the impact of exposure to this technologies on self-assessment.

The Research Questions

The purpose of this study is to examine the effects of OSM and OSSM on students' self-assessment skills. The research questions that guide the study are:

1. Do OSM and OSSM improve students' self-assessment skills of their *absolute* SQL performance?
2. Do OSM and OSSM improve students' self-assessment skills of their *relative* SQL performance?
3. Is there any significant relation between students' self-assessment abilities and their normalized gain?

The Study

In order to examine the impact of OSM and OSSM on students' self-assessment skills, we conducted a classroom study in a graduate Database Course. The students were able to access a variety of database learning content through a social visualization system Mastery Grids (Loboda et al. 2014) which was developed by the Personalized Adaptive Web Systems (PAWS) Lab of the School of Information Sciences at the University of Pittsburgh. The learning content was focused on teaching SQL programming and included problems and examples from a Database Exploratorium system (Brusilovsky et al., 2010).

We used two versions of Mastery Grids environment for this study. One of them (called simply OSM) provided only OSM functionality, i.e., students were able to see an up-to-date model of their SQL knowledge. As it can be seen in Figure 1, in OSM interface all content topics are shown as grid cells that represents the progress of the user. The grids color are changing from a continuum gray to dark green to show the growth of student's knowledge of the topic. When students mouse over grid cells, user can see her numerical progress percentile on that topic. When they click on a grid cells they can see what kind of content (examples and problems) is available for that topic and also see their knowledge progress for each example and quizz in that topic.

Figure 1. OSM interface

Another version (called simply OSSM) provided full Mastery Grids interface that supports both OSM and OSSM functionality (note that OSSM is an extension of OSM!) Figure 2 shows OSSM interface which includes the same row of topic-by-topic knowledge visualization as the OSM interface and additionally two more grid rows and a button "Load the rest of learners". In OSSM interface, students can see the average progress of the class in Group row in a continuum of colors from gray to blue. It also compares students' progress with the average of class in 'Me vs Group' column. In this row if the user has a higher progress in any topic, grid color becomes green, if the user has a lesser progress comparing to class average grid colors become blue. If the student clicks the 'Load the rest of learners' button, she can see the knowledge models or all student in class ordered by progress as well as her exact position in this list as in Figure 3.

Figure 2. OSSM interface

Figure 3. Student list in OSSM interface

Participants

The study was performed in two sections of a graduate Database Course offered in the Fall 2014 at the University of Pittsburgh. One section of 47 graduate students was assigned to work with OSM interface and another section with 42 students was assigned to OSSM interface. The system use was not mandatory, so some students used the system to different extent and some didn't. For the purpose of the study we identified two subgroups in each section - students who used the system to reasonable extent to be affected by it and students who did not use the system. The criteria for inclusion in the first group was solving at least 5 questions, the criteria for the second group was never log in system.

Data Collection Tools

For data collection, same academic achievement test including 10 questions about SQL statements is used as a pretest and posttest. After each SQL questions' students are also asked to check 'Yes' or 'No' to state whether they are confident their answer is correct. After answering all questions students are also asked to write their opinions about their position (percentage) among other classmates if ranked by the results of the sql test. According to students' actual answers and their self-assessment data, four different scores were calculated for both pretest and posttest to determine student's self-assessment abilities. These scores are:

1. Total correct assessments: This score shows students total number of correct assessments about their answers in the test. It is computed for total number of questions where a student checked she give the correct answer (+) and indeed performed correctly (+) or they checked they did not give correct answer (-) and they really give wrong answers (-).
2. Correct ratio: This score is computed by dividing the total number of correctly answered questions to the total number of questions that were checked as correct by the student.
3. Wrong ratio: This score is computed by dividing the total number of incorrecttly answered question to the total number of questions that were not checked as correct by the student.
4. Percentile estimation: This score is obtained from student's answer for the question "If you ranked all your classmates according to their SQL knowledge (from the highest score the lowest score), which percent would you be in comparison to the rest of the class? (Student who have the best score will be %1, student who have worst score will be in %99)"
5. Actual percentile: The score obtained from converting students actual scores from the test to percentiles.
6. Percentile difference: This score is the difference between estimated percentile and actual percentile.

Findings

1. Change in Students Absolute Knowledge Prediction

To examine whether OSM or OSSM improves students self-assessments skills of their absolute performance, students pre and post *absolute* self-assessments (1-3) are compared. Paired sample t test is used for normally distributed data and if the normality assumption is violated Wilcoxon sign test is conducted. Several significant differences were found between student self-assessment scores at pre-test and post-test times.

Table 1: Paired Sample t-test Results for Students' Self-Assessments' of Their SQL Knowledge

Group	Self-assessment measures	Mean	SD	t	p-value
OSSM	pre total correct assessments	2.90	1.95	29	.000
	post total correct assessments	6.13	1.94		
	pre correct ratio	0.44	0.31	20	.010
	post correct ratio	0.66	0.16		
OSM	pre total correct assessments	3.42	2.11	11	.001
	post total correct assessments	6.25	1.60		

Table 2. Wilcoxon sign test Results for Students' Self-Assessments' of Their SQL Knowledge

Group	Self-assessment measures	negative	positive	ties	p-value
OSSM	pre wrong ratio - post wrong ratio	0	6	13	.027
OSM	pre correct ratio - post correct ratio	0	8	2	.012
	pre wrong ratio - post wrong ratio	1	1	6	.655

As it can be seen in Table 1 and 2, post self-assessments scores are always higher than pretest scores. In other words, after working with the system, students' self-assessments' abilities improved. This can be explained both by the effects of learning SQL and by studying with e-learning system including OSM. To find out the main reason, the pre-test and post-test self-assessments scores were also compared for students who took the Database course but did not use the system. The analysis showed that none of these measures are significant for these students in both OSM and OSSM group. This result indicates that the main reason of self-assessment improvement was not going through the course but using the system. In other words, both OSM and OSSM interfaces helped students to better assess their absolute knowledge.

2. Change in Students Relative Knowledge Prediction

We also analyzed how successful the students are in comparing themselves with their classmates on a relative percentile scale (4-6). When we look pretest assessments, we found significant differences between students average actual performance and estimated performances for both OSM (actual percentile=52, estimated percentile=33.78) ($t(32)=-4.70, p<.001$) and OSSM groups (actual percentile=48, estimated percentile=29.36) ($t(40)=-6.36$, $p<.001$).

In post-test assessments for OSM group there were still significant differences for relative knowledge prediction. Whereas students actual performance fell in the 43th percentile on average, they put themselves to 75^{th} percentile ($t(11)=7.67, p<.001$). In other words students were incapable of assessing their relative knowledge in post tests in OSM group. However, there were no significant differences in OSSM group. The average performance of students in OSSM group was in the 59^{th} and they also put themselves in the 59^{th} percentile ($t(26)=-.042, p>.05$). This expected result confirms when students have chance to observe their classmates performance and compare their own progress with others, they become more aware about their relative knowledge. In other words, OSSM environment provided students critical information to better assess their relative level of knowledge in class. This result supports the main principle of Social comparison theory, which states the importance of social comparison for humans for helping them evaluate their abilities and reduce uncertainty (Buunk & Gibbons, 2007).

3. The Effects of Success Quartiles on Self-Assessment of Performance

Incompetent individuals are expected to be less successful in assessing their knowledge (Kruger and Dunning, 1999). To examine whether it is true in our context, we compared student's' self-assessment abilities between student with different level of knowledge. For this test we grouped students by their test score into quartiles. Figure 4 shows students percentile estimations and actual percentiles in pre and post-test according to their success quartiles in both OSM and OSSM groups. Note that there were no students in 2nd quartile.

Figure 4. Percentile Estimation and Actual Percentiles According to Success Quartiles

As it can be seen in Figure 4, students estimated percentiles and actual percentiles in bottom quartiles are very close to each other in both groups in pretest. According to the one sample t test results, in pretest, estimations of bottom quartile students in both OSM (18th percentile) and OSSM (26th percentile) groups were not significantly different from average actual score (20th). In other words, bottom quartiles are successful to assess their relative (low) knowledge in pretest. However, for 3rd and 4th quartiles estimation were significantly different in pretest in both groups. On average users in 3rd quartile put themselves in the 33.5 percentile in OSM and 34.4 percentile in OSSM groups, which is considerably lower than their actual performance 56th percentile. Similarly 4th quartile put themselves into 41.82 percentile and 53.13 percentile, whereas their average actual performance was 85th percentile in OSM and 87th percentile in OSSM. As a result, it is for the bottom quartile, that self-assessments were adequate in pretest. Whereas it looks like an unexpected result, in our context it is reasonable. SQL is a new topic for which many student had no or very little knowledge. The bottom quartile at pre-test was fully formed by novices who answered at most 1 question correctly. The same reason lead students in 3rd and 4th quartile to considerably underestimate their knowledge. They have moderate knowledge and expressed it in self-assessment, however due to large number of total novices (the fact that stronger students were not able to predict) their true relative score was quite high.

However when we look the results in post-test we can see that after passing the course, students generally overestimate their knowledge while more successful students are better in assessing their relative knowledge. At post-test time when no one remained a total novice, students both relative and absolute (correct ratio) assessments follow the prediction from the cited literature. As Figure 5 shows, in OSM group students in bottom percentile are, indeed, the worst to assess their absolute knowledge as Kruger and Dunning (1999) suggested. When we compared students correct ratio according to their quartiles with Kruskal-Wallis test, we found significant differences ($p=0.36$). In other words upper quartiles have higher means - if users have lower domain knowledge they also have lower ability to assess their knowledge level. When we compare students correct ratio according to their quartiles with Kruskal-Wallis test, we also see significant differences for OSSM ($p=.003$) groups. In OSSM group fourth quartile significantly higher than third quartile, third quartile is significantly higher than second quartile, however first quartile is not significantly different from any quartiles and it also has not the lowest mean. However, there are only two students in this quartile.

Figure 5. Self-Assessments According to Posttest Success Quartiles

4. Normalized Gain and Self-Assessments

Correlations between the self-assessment measures and normalized gain (post-test to pre-test) were computed and corrected by the Pearson Correlation formula. A significant negative correlation was found at the .01 level between normalized gain and post percentile difference for both OSM (R=-.800, p = .000) and OSSM groups (R=-.629, p =.000). In other words, when the students' knowledge gain increases, the difference between their estimated percentiles and real percentiles is degreasing - better learners are better in assessing their knowledge. A significant positive correlation was found between correct ratio and normalized gain in both OSM (R=.910, p = .000) and OSSM (R=.515, p =.004) groups. A significant positive correlation was also found between total correct assessment's and normalized gain in both OSM (R=.717, p =.009) and OSSM groups (R=. 374, p = .042). These results also showed that when the students' success increases their ability of absolute knowledge assessments is improving. Figure 6 shows the relations between students' success and their absolute and relative assessment abilities.

Figure 6. Scatter Plots Showing Relationship between Normalized Gain and Self-Assessments

Conclusions

The results of our study indicated that students' self-assessment abilities are improved when they use Mastery Grids interface for their learning. This is another confirmation that open student modeling is beneficial for meta-cognitive skills, in this case, to support students' self-awareness about their absolute knowledge. We also obtained evidence that OSSM interface is more efficient than OSM interface for improving individuals' self-assessments about their relative knowledge in a class. Finally, we obtained evidence that students' normalized gain are correlated with their self-assessment skills. Better learners, i.e., those with higher normalized gain scores, are also better in their meta-cognitive skills, i.e., their absolute and relative knowledge assessments.

Acknowledgment

The first author is supported in part by grants from The Turkish Fulbright Commission and 2219 Postdoctoral Research Fellowship Program of The Scientific and Technological Research Council of Turkey. This research was supported by the Advanced Distributed Learning Initiative contract W911QY13C0032.

References

Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. *Dimensions of thinking and cognitive instruction*, *1*, 53-92.

Buunk, A. P., & Gibbons, F. X. (2007). Social comparison: The end of a theory and the emergence of a field. *Organizational Behavior and Human Decision Processes*, *102*(1), 3-21.

Brickman, P., & Bulman, R. J. Pleasure and pain in social comparison. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives*. Washington, D.C.: Hemisphere, 1977.

Brusilovsky, P., Sosnovsky, S., Lee, D., Yudelson, M., Zadorozhny, V., and Zhou, X. (2010) Learning SQL programming with interactive tools: from integration to personalization. *ACM Transactions on Computing Education* **9** (4), Article No. 19, pp. 1-15

Bull, S. and Kay, J. (2013). Open Learner Models as Drivers for Metacognitive Processes. In: R. Azevedo and V. Aleven (eds.): *International Handbook of Metacognition and Learning Technologies*. Springer International Handbooks of Education, Berlin: Springer, pp. 349-365.

Festinger, L. (1954). A theory of social comparison processes. *Human relations*, *7*(2), 117-140.

Flavell, J. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist*, 34, 906– 911.

Garner, R., & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist*, 24(2), 143-158.

Garrison, D. R. (1997). Self-directed learning: Toward a comprehensive model.*Adult Education Quarterly*, *48*(1), 18-33.

Hsiao, I. H., Bakalov, F., Brusilovsky, P., and König-Ries, B. (2013) Progressor: social navigation support through open social student modeling. *New Review of Hypermedia and Multimedia,* 19 (2), 112-131.

Jubran, S. M., Samawi, F. S., & Aalshoubaki, N. H. (2014). The level of Students' Awareness of the Self-monitoring Strategy of Reading Comprehension Skills in Jordan and its Relationship with the Desire to Learn. *Dirasat: Educational Sciences*, *41*.

Kay, J. (2000). User Interfaces for All, chapter User Modeling for Adaptation, p.p. 271–294. *Human Factors Series*. Lawrence Erlbaum Associates, Inc., http://www.cs.usyd.edu.au/~judy/Homec/Pubs/ch18.pdf

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, *77*(6), 1121.

Loboda, T., Guerra, J., Hosseini, R., and Brusilovsky, P. (2014) Mastery Grids: An Open Source Social Educational Progress Visualization. In: S. de Freitas, C. Rensing, P. J. Muñoz Merino and T. Ley (eds.) Proceedings of 9th European Conference on Technology Enhanced Learning (EC-TEL 2014), Graz, Austria, September 16-19, 2014, pp. 235-248.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into practice*, *41*(4), 219-225.

Somyürek, S. (2009). Student modeling: Recognizing the individual needs of users in e-learning environments. *International Journal of Human Sciences*. 6(2). 429-450.

Stankov, L., & Crawford, J. D. (1996). Confidence judgments in studies of individual differences. *Personality and Individual Differences*, *21*(6), 971-986.