Chen, W. et al. (Eds.) (2016). Proceedings of the 24th International Conference on Computers in Education. India: Asia-Pacific Society for Computers in Education

Provision of Latitude for Target Selection during Online Peer-Assessment

Fu-Yun YU
Institute of Education, National Cheng Kung University, Taiwan
fuyun.ncku@gmail.com

Abstract: The aim of this study was to investigate the potential of the provision of latitude for target selection during online peer-assessment. One online system was expanded to support the associated learning activities. Fifty-four undergraduates participated and experienced different ranges of items available for online peer-assessment: full range (i.e., all items listed), partial range (half or one-fourth items listed), and the exact items to be assessed. Two major findings were obtained. First, Pearson's chi-squared test of goodness of fit found that the observed frequency distribution among the three arrangements was statistically significant. Second, while the majority of the participants preferred the system automatically assigning the exact number of items to be assessed, still nearly half of all participants appreciated having some latitude in choosing their targets during online peer-assessment, with explicit and legitimate reasons supporting their revealed preference. Suggestions for instructors and system developers are provided.

Keywords: Choice of targets, online learning activity, peer-assessment, revealed preference

1. Introduction

Peer-assessment (PA) has attracted increasing attention among researchers and practitioners since the 1990s, along with the greater acceptance of the notion of interconnectedness of teaching, learning and assessment. PA is the practice of students giving evaluative feedback to their peers about the work they have produced and their overall performance (Topping, 2009). Through the resulting processes of mutual support and assistance, it is suggested that peers with similar learning status are able to assist each other to achieve their respective learning goals, expand their knowledge bases and enhance their skill levels (Topping and Ehly, 2001).

With empirical evidence generally confirming the use of PA to support various cognitive, affective and social growth (e.g., self-regulated learning, critical thinking, performance, learner motivation, sense of responsibility, and attitudes) (Topping, 2009; van Gennip, Segers and Tillema, 2009; Yu and Wu, 2013), issues regarding how such activities can be further made varied and versatile have been the focus of several recent research endeavors. For instance, the potential of embedding the features of anonymity and created identities for assessors (Yu and Wu, 2011), allowing multiple PA modes (Yu, 2011), and incorporating different types of feedback (Gielen, Peeters, Dochy, Onghena and Struyven, 2010; Nelson and Schunn, 2009; Yang, Badger and Yu, 2006) have all been explored. Along this same line of thought and in consideration that students being asked to assess a specific item or a fixed set of items seems to be the main form of online PA, the aim of this study was to investigate whether the provision of latitude for target selection would be a viable approach for online PA. Specifically, the two research questions examined in this work are as follows:

(1) Would students exhibit preferences for different ranges of latitude for target selection?

(2) What were the students' underlying reasons supporting their revealed preferences for a specific PA arrangement?

2. Methods

To serve the research purposes of this study, an online system supporting students to generate different types of questions using different media formats, and engage in followed-up PA was extended. The expanded version of the system allows the instructor to determine different ranges of latitude for online

PA—full and limited. In the 'full range' arrangement, all questions generated are presented to users as candidates for their assessment targets. On the other hand, for the 'limited range' the instructor can specify a specific number/percentage of items produced to be sent to respective peer-assessors as assessment targets to be considered.

Fifty-four student teachers from different colleges (liberal arts: 30; science: 9; social sciences: 9; engineering: 1; medical: 1; planning and design: 1: life science: 2; electrical engineering and computer science: 1) and levels (undergraduates: 37; graduates: 17) enrolled in one course (effective instructional principles) participated in this study. In the first class, the instructor explained the general arrangement, requirements, course format and the reason for incorporating student-generated questions (SGQ) and PA in this course. A training session was then arranged to equip students with essential skills associated with the focal tasks. Information on the basic concepts related to SGQ and PA with examples and the operational procedures of the adopted system were explained and practiced. As a routine practice, following the instructor's explanation of each instructional principle, students were given twenty minutes to generate two multiple-choice questions pertaining to the delivered instruction and assigned text in class. They were then asked to assess four randomly assigned questions so that individual feedback from peers could be obtained, and the questions could be revised with reference to peers' feedback out of class. At the next class session, group feedback was given by the instructor to highlight exemplary SGQ and PA practices.

To enable all participants to experience different ranges for target selection, the number of items listed for the participants for online PA was set as follows: full range (i.e., all questions generated by the participants during the current SGQ activities), partial range (50%, 25%), and the exact number of items to be assessed (i.e., four items). In this study, each participant experienced each of the four ranges (i.e., 100%, 50%, 25%, and exact items) on two separate occasions. Regardless of whichever ranges students were exposed to, during online PA they were directed to click on the item to be assessed from the list of SGQ window (Figure 1), before being directed to the online PA form (Figure 2), where they provided both quantitative and qualitative feedback to the question-author. In sum, the only difference among different PA arrangements in this study is the number of items shown in the 'list of SGQ window.'

Figure 1. List of SGQ window

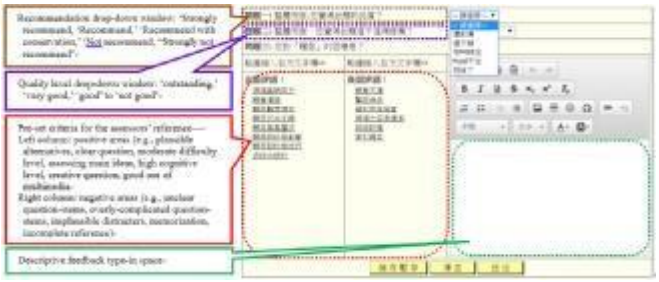

Figure 2. The online PA form

To collect the participants' views of different PA arrangements (i.e., the number of items available as potential targets for PA), they were asked to respond to one question in the last instructional session: 'In this semester, you were directed to assess SGQ where different numbers of items were listed as candidate targets. Which of the following arrangements did you prefer: full range (listing all SGQ), partial range (listing half or one-fourth of SGQ), and the exact items to be assessed (i.e., four items). Please explain your selection.'

3. Results and Discussion

Descriptive statistics and Pearson's chi-squared test of goodness of fit were used to analyze the quantitative data, and the constant comparative data analysis method proposed by Lincoln and Guba (1985) was adopted to analyze students' descriptive responses. Of those completing the question (n=53), it was found that the majority of the participants (n=27, 50.94%) preferred the system automatically assigning the exact number of items to be assessed. Despite this, there were still nearly half of all the respondents (n=26, 49.06%) expressing a preference to having some latitude in selecting their targets, with half of those preferring the full range (n=13, 24.53%), and the other half the partial range (n=13,

24.53%). A X^2 test further indicated that the observed frequency distribution among the three arrangements was statistically significant, $X^2 = 7.40$, $p < .05$.

Two salient themes emerged from students' written responses explaining their preference for the full range arrangement. Of the thirteen respondents, almost all (n=12, 92.31%) highlighted the feature of 'being able to observe questions with various styles and focus,' which helped to enhance their abilities. In addition, more than two-thirds (n=9, 69.23%) stressed the feature of 'being given choices regarding which item to assess.'

For those preferring the partial range, all thirteen respondents acknowledged the strengths associated with the other two arrangements while appreciating what this eclectic approach could do by mitigating problems frequently associated with each of the other two arrangements. More explicitly, users were given a choice in selecting their assessment targets and could learn from observing peers' work with various styles (which was not possible with the exact item approach), while not feeling overwhelmed (as would be the case in the full range arrangement).

Finally, the reason of 'ensuring each question generated receives some feedback from peers for further improvement' was highlighted most frequently by those preferring the exact item arrangement (n=17, 62.96%), followed up by 'the opportunity to be challenged since students could no longer target items of their liking' (n=11, 40.74%). Also, quite a few (n=8, 29.63%) pointed out the reason of 'time being saved and better spent on giving feedback rather than on reviewing and deciding which item to assess.'

Existing studies have found that students based their choice of items to assess on a number of reasons, including the characteristics of the assessed item (e.g., the interestingness, quality, and difficulty level), its relation to the key ideas of the study material, the number of times the item was assessed, item submission time, and the length of the question item (Yu and Sung, in press). Also, as recognized by humanists, each learner has different motivations, interests, needs, and preferences (Ediger, 2006). In light of these, as well as the findings of this study, instructors are advised to consider giving their students some freedom in selecting targets to create more versatile PA activities. In addition, developers of online learning systems supporting PA may consider incorporating such a feature to allow the provision of some latitude in target selection.

References

Ediger, M. (2006). Present day philosophies of education. *Journal of Instructional Psychology. 33*(3), 179-182.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304-315.

Lincoln, Y. S. & Guba, E. G. (1985). Naturalistic inquiry. Beverly Hills, CA: Sage Publications.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science, 37*(4), 375-401.

Topping, K. J. (2009). Peer assessment. *Theory into Practice, 48*(1), 20-27.

Topping, K. J., & Ehly, S. E. (2001). Peer-assisted learning. *Journal of Educational and Psychological Consultation, 12*(2), 113-132.

van Gennip, N. A. E., Segers, M. & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal and structural features. *Learning and Instruction, 4*(1), 41–54.

Yang, M., Badger, R., & Yu, Z. (2006). A comparative study of peer and teacher feedback in a Chinese EFL writing class. *Journal of Second Language Writing, 15*, 179-200.

Yu, F. Y. (2011). Multiple peer assessment modes to augment online student question-generation processes. *Computers & Education, 56*(2), 484-494.

Yu, F. Y. & Sung, S. (in press). A mixed methods approach to the assessor's targeting behavior during online peer-assessment: Effects of anonymity and underlying reasons. *Interactive Learning Environments*.

Yu, F. Y. & Wu, C. P. (2013). Predictive effects of online peer feedback types on performance quality. *Educational Technology and Society, 16*(1), 332-341.

Yu, F. Y. & Wu, C. P. (2011). Different identity revelation modes in an online peer-assessment learning environment: Effects on perceptions toward assessors, classroom climate and learning activities. *Computers & Education, 57*(3), 2167-2177.