# Implementation Of Peer Review Assessment To Engineering Students

Hairuzila Idrus[1], Herri Mulyono[2], Raihan Mahirah Ramli[1], Ena Bhattacharyya[1], Siti Zulaiha[2]

[1]Universiti Teknologi PETRONAS, Malaysia
[2]UHAMKA, Jakarta, Indonesia

**Abstract**. Online assessment has been increasingly popular in the tertiary education setting nowadays and has had significant changes to the process of assessment. Several studies have been conducted on online assessment, particularly on online peer assessment and found that it benefits students in teaching, learning and assessment. This paper will discuss the result of a study conducted in a Chemical Engineering class which uses online peer assessment as one of the tools to conduct assessment. The objective of this study was to investigate the effectiveness of online peer assessment to students' learning from the students' perspective. The perspective of the students was under study because they are the ones who are going through the process, thus, the effectiveness of the assessment to their learning can only be known if the information is obtained from them. The students were divided in small groups and asked to conduct online peer assessment using any free online survey tools. Each group had to conduct 2 or 3 cycles of peer assessment to ensure validity of the rating. The peer rating was used to produce individual marks for group work. At the end of the semester, the students were asked to write a reflection on the effectiveness of online peer assessment in their learning process of the course. The result of this study found that students have positive perspectives on the effectiveness of this assessment. This will be further discussed in the paper. In addition, pedagogical implications will also be discussed on the use of such online peer assessment in student's learning task.

## 1 Introduction

Peer assessment is an activity in a learning process which requires students giving feedback to each other to grade their work or performance using relevant criteria. In the past few years, peer assessment has been implemented as a strategy for formative assessment (Cheng & Warren, 1999; Sadler, 1989) and for engaging students as active learners (Gielen et al., 2009; Sadler, 1989; Topping et al., 2000). Research on peer assessment has gathered significant testimony on the pedagogical (Falchikov & Blythman, 2001), meta-cognitive (Butler & Winne, 1995; Topping 1998), cognitive (Nelson & Schunn, 2009; Tseng & Tsai, 2007), and affective benefits (Strijbos et al., 2010) of peer assessment on student learning. These efforts have resulted in peer assessment being effectively designed and applied in higher-education contexts (Topping, 2003).

Online assessment is a kind of assessment conducted over the Internet by using available web technologies. It has been progressively prevalent in tertiary education setting these days and has had considerable changes to the process of assessment. Numerous studies have been conducted on online assessment and found that it benefits students in teaching, learning and assessment (Chen, 2010; Freeman & Mc Kenzie, 2002; Lu & Law, 2012). Online peer assessment enables students to communicate with peers, to reflect on and continuously revise their work based on feedback. This type of assessment can also increase the willingness of students to engage in peer assessment by allowing them to anonymously grade and provide feedback outside classroom time. Students will also be able to provide reciprocal peer review amongst each other (Cho & Schunn, 2007).

The assessment has become progressively trendy since the introduction of the Internet and has significantly changed the process of assessment (Tseng & Tsai, 2007). Online systems initiate functions like assignment submission, storage,

communication and review management (Kwok & Ma, 1999; Liu et al., 2001), which has a number of advantages over face-to-face assessment (Tsai, 2009; Tsai & Liang, 2009; Yang & Tsai, 2010). Online assessment allows students to communicate with their peers, reflect on and continuously revise their work based on the feedback given (Yang, 2010). The willingness of students to engage in peer assessment can be enhanced by the online system through allowing them to anonymously provide feedback and grade whenever they like (Lin et al., 2001; Tsai, 2009). These systems also enable teachers to supervise the online activities and progress of their students more closely (Lin et al., 2001), and allow researchers to collect information about students by automatically recording data about assignments, online participation and communication (Tsai, 2009).

The objective of this paper is to discuss the findings on students' perspectives on the effectiveness of online peer assessment. The perspective of the students was under study because they are the ones who are going through the learning process, hence, the effectiveness of the assessment to their learning can only be known if the information is obtained from them.

## 2 Methodology

The participants were asked to conduct online peer assessment for about 2 -3 times in a semester. At the end of the semester, they were given a set of self-developed survey questionnaire to gauge their perception on the assessment activity. They were also asked to write a reflection on the effectiveness of online peer assessment in their learning process of the course. But in this paper, the reflection will note discussed. The questionnaire will be analyzed quantitatively using Rasch model.

### 2.1 Instrumentation

A total of 17 questionnaire items were used to conduct a survey of students' perceptions of the use of peer assessment (PA). The questionnaire acquired four dimensions including 8-items of Positive Attitude Subscale (POS), 3-items of Understanding and Action Subscale (UAS), 4-items of Negative Attitude Subscale (NAS), and 2-items of Interrelation Factor (IF) dimensions.

The instrument had been measured for validity and reliability using the Rasch model. Prior to actual data collection, the Rasch model can transform ordinal data (Likert-rating scale) into interval data called "logit" (logarithm odd units). Through Winsteps software version 4.3.4, the Rasch model can thus simultaneously measure the validity and reliability of the instrument, which also evaluated the overall interaction between person and item (Bond & Fox, 2015; Linacre, 2012).

### 2.2 Instrument Reliability

Measurement using the Rasch model allowed the acquisition of the instrument as a whole to acquire data statistics, as outlined in Table 1. The reliability and fit statistics values were designated by the logit value, considering for person and item.

**Table 1.** Instrument Statistics

|  | Mean | Reliability | Separation | Cronbach Alpha |
|---|---|---|---|---|
| **Person** | 1.97 | 0.83 | 2.24 | 0.91 |
| **Item** | 0.00 | 0.97 | 5.93 |  |

Table 1. contributed to provide a mean score, reliability, separation and Cronbach alpha. Statistically, the interpretation of Cronbach alpha (0.91) indicated a high level of reliability (Sumintono & Widhiarso, 2014). This value indicated there was excellent interaction between 79 persons and 17 items. Definitely, the person means score (1.97) showed a positive perception obliging the use of PA in learning.

Furthermore, the reliability of persons and items scored 0.83 and 0.97, respectively. This value implied that the person holds an acceptable ability to respond to items. While the items' reliability value demonstrated a great quality of the item in assessing the persons' ability.

To identify the quality of persons' and items' dissemination, Table 1. revealed the value of the separation index. The person (2.24) and item (5.93) separation index pointed out an excellent distribution. It is accepted that the instrument can distinguish the 'person abilities' (Linacre, 2012) and the 'item difficulties' (Boone, Staver, and Yale, 2014), which can be classified into several groups. This indicates that both persons and items have diverse groups.

### 2.3 Item Stratification

Regarding the information in Table 1. the item stratification was classified into 5.93 (rounded to
6) different strata based on logit value (see Table 2). The classification of strata was then divided into two broad assortments regarding the value of the item separation index (see Table 1). Each of the classifications was ranging from the most difficult item upon (item NAS4, LVI = 1.88) to the easiest item to be agreed (item POS6, LVI = -1.11).

**Table 2.** Item Stratifications

| Category | Criteria | Item/LVI |
|---|---|---|
| **More Difficult to be Considered** | | |
| **Difficulty Strata I** | LVI > 1.32 | NAS4 (LVI = 1.88) |
| | | NAS5 (LVI = 1.47) |
| | | NAS2 (LVI = 1.42) |
| **Difficulty Strata II** | 1.32 > LVI > -0.45 | IF2 (LVI = 1.31) |
| | | IF1 (LVI = 1.29) |
| | | NAS3 (LVI = 0.87) |
| **Difficulty Strata III** | -0.45 > LVI > -.062 | POS5 (LVI = -0.46) |
| | | POS1 (LVI = -0.49) |
| | | UAS2 (LVI = -0.57) |
| **Easier to be Considered** | | |
| **Difficulty Strata IV** | -0.62 > LVI > -0.81 | UAS1 (LVI = -0.63) |
| | | UAS3 (LVI = -0.66) |
| | | POS2 (LVI = -0.76) |
| **Difficulty Strata V** | -0.81 > LVI > -0.91 | POS3 (LVI = -0.82) |
| | | POS7 (LVI = -0.82) |
| **Difficulty Strata VI** | LVI < -0.91 | POS4 (LVI = -0.92) |
| | | POS8 (LVI = -0.99) |
| | | POS6 (LVI = -1.11) |

Referring to the strata classification, three items (POS4, POS8, and POS6) appeared as the easiest item categories agreed upon by students. They perceived peer assessment as an assessment method that can help them develop a sense of participation in learning (POS4, LVI = -0.92). Also, the use of peer assessment in evaluating student performance was considered to be fair (POS8, LVI =-0.99), so that it can trigger their motivation in learning (POS6, LVI = -1.11).

## 3 Findings

### 3.1 Item-person Maps

To address the issue pertaining to the items which are agreed based upon the person, the Rasch model administered the output in the form of variable maps. This output allowed researchers to see which items were considered most important by the persons (see Figure 1)

```
          MEASURE                                        Person - MAP - Item
                                    <more>|<rare>
                        4 03A 04A 05A 11C 13C 28G 45K 47K 48K 53M 58N 63P
                          66P 74R 77T                                    +
                                         |
                                         |
                                         |
                                         |
                                        30G ||
                                         |
          3                              +
                                         |
                                         |
                                        29G |
                                         T|
                                         |
                                        02A |
                                        21E |T
          2                              +
                                      49L | NAS4
                                      15D S|
                               16D 17D 50L 60O 68Q  |
                                  44K 54N 65P | NAS1
                       06B 07B 09B 12C 18D 27F 34H 37I 40J 52M 56N 59O
          62O 64P 75S                 | NAS2
                               57N 69Q | IF1    IF2
                                  25F 31G 33H |
          1                 01A 19E 22E 41J M+S
                               26F 35I 36I | NAS3
                          32G 38J 51M 73R 76S 78T 79T  |
                                     55N |
                                         |
                                 42J 70Q 71R |
                                        S|
                                         |
          0             08B 10C 23F 39J 43K 72R +M
                                     20E 46K |
                                         |
                                     61O 67Q |
                                         |
          T| POS1  POS5
          24F | UAS1  UAS2  UAS3
                                     14D | POS2
          | POS3  POS4  POS7
          -1                           +S POS8
          | POS6
                                         |
                                         |
                                         |
                                         |
                                         |
                                         |
          -2                             +
                                    <less>|<freq>
```

**Figure 1**. Variable Maps

Figure 1 above displays the distribution of agreements between persons and items, indicated by the logit value item (LVI). The results were obtained regarding the response of the person and item, respectively. The left part considering as a person distribution map, while the right part performed the result for the item distribution map. To the end, 5 items (POS3, POS4, POS7, POS8, and POS6) revealed as the most approved by 79 persons.

Referring to the person mean score of 1.97 (>0.00) in Table 1, in general, students perceived a positive perception toward the use of peer assessment in learning. This assumption was written from the perspectives that learning activities assisted by peer assessment can improve verbal communication skills (POS3, LVI = -0.82). It also helped students in developing their sense of participation (POS4, LVI = -0.92) so they found it to be motivated in learning (POS6, LVI = -1.11). The impact might be to increase the arousal of interaction between students (POS7, LVI = -0.82). Most

importantly, peer assessment defined as a fair method of evaluating students' performance (POS8, LVI = -0.99) since they could provide constructive criticism and suggestions.

### 3.2 Item Fit Order

**Table 3**. Item Statistics-Misfit Order

| |ENTRY NUMBER | TOTAL SCORE | TOTAL COUNT | MEASURE | MODEL S.E. | INFIT MNSQ | ZSTD | OUTFIT MNSQ | ZSTD | PTMEASUR-AL CORR. | EXP. | EXACT MATCH OBS% | EXP% | Item |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 12 | 245 | 79 | 1.47 | .13 | 1.86 | 4.34 | 2.19 | 5.51 | A .65 | .77 | 31.3 | 38.0 | NAS1 |
| | 13 | 248 | 79 | 1.42 | .13 | 1.46 | 2.54 | 1.59 | 3.12 | B .70 | .76 | 32.8 | 37.8 | NAS2 |
| | 14 | 279 | 79 | .87 | .13 | 1.47 | 2.52 | 1.50 | 2.68 | C .63 | .69 | 32.8 | 39.6 | NAS3 |
| | 15 | 222 | 79 | 1.88 | .14 | 1.28 | 1.67 | 1.37 | 1.99 | D .77 | .81 | 32.8 | 41.2 | NAS4 |
| | 17 | 254 | 79 | 1.31 | .13 | 1.33 | 1.88 | 1.30 | 1.71 | E .71 | .75 | 45.3 | 38.3 | IF2 |
| | 16 | 255 | 79 | 1.29 | .13 | 1.24 | 1.44 | 1.25 | 1.45 | F .72 | .75 | 39.1 | 38.5 | IF1 |
| | 8 | 358 | 79 | -.99 | .19 | 1.09 | .48 | .97 | -.08 | G .45 | .44 | 65.6 | 58.1 | POS8 |
| | 6 | 361 | 79 | -1.11 | .20 | .69 | -1.58 | .85 | -.62 | H .48 | .43 | 68.8 | 59.3 | POS6 |
| | 5 | 341 | 79 | -.46 | .17 | .80 | -1.10 | .73 | -1.49 | I .56 | .51 | 54.7 | 45.8 | POS5 |
| | 3 | 353 | 79 | -.82 | .18 | .74 | -1.36 | .67 | -1.75 | h .52 | .47 | 57.8 | 54.0 | POS3 |
| | 1 | 342 | 79 | -.49 | .17 | .65 | -2.08 | .70 | -1.70 | g .56 | .51 | 53.1 | 46.8 | POS1 |
| | 2 | 351 | 79 | -.76 | .18 | .65 | -1.93 | .66 | -1.81 | f .54 | .47 | 62.5 | 52.7 | POS2 |
| | 9 | 347 | 79 | -.63 | .17 | .59 | -2.44 | .66 | -1.90 | e .57 | .49 | 57.8 | 48.7 | UAS1 |
| | 11 | 348 | 79 | -.66 | .17 | .61 | -2.24 | .65 | -1.91 | d .56 | .49 | 64.1 | 48.8 | UAS3 |
| | 7 | 353 | 79 | -.82 | .18 | .57 | -2.48 | .62 | -2.07 | c .55 | .47 | 59.4 | 54.0 | POS7 |
| | 10 | 345 | 79 | -.57 | .17 | .57 | -2.58 | .56 | -2.60 | b .60 | .50 | 68.8 | 46.8 | UAS2 |
| | 4 | 356 | 79 | -.92 | .19 | .46 | -3.31 | .48 | -2.95 | a .57 | .45 | 67.2 | 56.0 | POS4 |
| MEAN | | 315.2 | 79.0 | .00 | .16 | .94 | -.4 | .98 | -.1 | | | 52.6 | 47.3 | |
| P.SD | | 49.1 | .0 | 1.04 | .02 | .41 | 2.3 | .46 | 2.4 | | | 13.5 | 7.3 | |

The utilization of the instrument frequently formed an unexpected result, such as the issue of misfit items. Measurement using the Rasch model was able to identify misfit items (outliers). To confirm the misfit items, the mean scores and standard deviations of the MNSQ INFIT (see Figure 2) were then calculated. The INFIT MNSQ mean score of 0.94 is added to 0.41 standard deviations, which is 1.35. This value alters as the preeminent limit value for items deemed appropriate. The INFIT MNSQ value that exceeds the 1.35 limit, then the item was categorized as misfit (outliers).

The information shown in Figure 2, presented three misfit items (> 1.35). These items are 'I think students should not be responsible for making assessments' (NAS1, LVI = 1.86), 'PA is time- consuming' (NAS2, LVI = 1.46), and 'My marks giving to classmates are affected by the marks given to me' (NAS3, LVI = 1.47). Students find it difficult to figure out and responded to the items appropriately.

## 4 Discussion

The results indicate that the students like the peer assessment activity because it gives them some empowerment in the learning process because their opinion in the assessment does matter. This is very important to make them in favour of the learning that took place in the course. Hence, they will also like the course and will be motivated to attend the class. A fair assessment of a course is very important to the students because it makes them feel safe and not being treated unduly. This is because their grades come from several assessors. If they are being assessed but only the lecturer, they might think that the lecturer might be bias towards them. However, with peer assessment being implemented, the lecturer cannot be accused of biasness. In addition, according to McGourty (2000), online peer assessment can provide learners with an anonymous environment to freely express their thoughts and ideas about other students' work, hence the safe feeling among the students.

The finding in this study also shows that the students believe that giving meaningful, constructive, and tangible comments to their peers appears to be more conductive to learning than solely assigning grades. In peer assessment activity, the assessors develop better understandings of the criteria used for the assessment and project tasks by evaluating and giving comments on the quality of work produced by the peers (Liu et al. 2001; Reuse-Durham 2005). They gained better comprehensions into their own work by assessing the projects of their peers. Consequently,  peer assessment

activity is a good exercise for the students to develop and improve their own critical thinking skills. This idea is supported by Roberts (2006) when he refers to peer assessment as "a process which allows learners to reflect critically upon the learning of their peers. Peer assessment is also a reciprocal process in that the student providing feedback also benefits from increasing their own understanding. This is achieved by students having to critique and review someone else's work and thereby reflect on their own understanding or performance." (p.19).

One of the findings in this study is that students feel that peer assessment has helped them improve on their communication skills thus, made them develop their sense of participation in the learning process. In other word, what the students have developed in this activity is actually their teamwork ability. Throughout each peer assessment activity students worked as a team by taking the roles as assessors and assesses alternately. As assessors, the students assess their peers' work, with their own expectations in mind. They gave comments and as they review, they had to discipline their thoughts to concentrate on the elements of a good piece of work that they have learned in the course. If the piece of work they read is comprehensible to them, it is deemed to be a good work. At the same time the assessors were able to learn from the work that they assessed as they encountered good pieces of work. Meanwhile, as assessees, the students read the feedback or comments and through this process, they were reminded of the elements of good work that they have learnt in class. They were informed about the assessors' expectations from the feedback given on their work. The assessee then eventually met such expectations through either incorporating the suggestions or revising their pieces of work. This indicates that peer assessment is a team effort.

## 5 Conclusion

In conclusion, the peer assessment activity is perceived by the engineering students as an effective method to help them develop a sense of participation in learning, a fair way of evaluating their performance, which has led to enhancing their motivation to learn. Ultimately, it gives an opportunity for students to enhance their soft skills such as critical thinking and communication skills, as well as teamwork. Besides understanding the criteria used for the evaluation, peer assessment activity allows students to read, respond and reflect what they read from the work of their peers to their existing knowledge. In the process of assessing the work of their peers, students were prompted to think analytically and critically of what they read and communicate their opinions to peers. In short, peer assessment activity can be conveniently incorporated in education to propel the development of the students' learning process and relevant skills.

## References

Chen, C.H. (2010). The implementation and evaluation of a mobile self-and peer-assessment system. *Computers & Education*, 55(1), 229-236.

Cheng, W., & Warren, M. (1999). Peer and teacher assessment of the oral and written tasks of a group project. Assessment & Evaluation in Higher Education, 24(3), 301–304.

Cho, K. & Schunn, C.D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, 48(30, 409-426.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. Review of Educational Research, 65(3), 245–281.

Falchikov, N., & Blythman, M. (2001). Learning together: Peer tutoring in higher education (1st ed.). New York: Routledge.

Freeman, M., & McKenzie, J. (2002). SPARK, a confidential web-based template for self and peer assessment of student teamwork: benefits of evaluating across different subjects. *British Journal of Educational Technology*, 33(5), 551-569.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2009). Improving the effectiveness of peer feedback for learning. Learning and Instruction, 20(4), 304–315.

Kwok, R. C. W., & Ma, J. (1999). Use of a group support system for collaborative assessment. Computers & Education, 32(2), 109–125

Lin, S., Liu, E., & Yuan, S. (2001). Web-based peer assessment: Feedback for students with various thinking styles. Journal of Computer Assisted Learning, 17(4), 420–432.

Liu, E. Z. F., Lin, S. S. J., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer review: The learner as both adapter and reviewer. IEEE Transactions on Education, 44(3), 246–251.

McGourty, J. (2000). 'Using multisource feedback in the classroom: A computer-based approach', IEEE Transactions on Education 43, 120–124

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.

Roberts, T. S. (2006) Self, peer, and group assessment in e-learning. Hershey, PA, Information Science Pub.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. Instructional Science, 18(2), 119–144

Topping, K. J. (1998). Peer assessment between students in colleges and universities. Review of Educational Research, 68(3), 249–277.

Topping, K. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy, & E. Cascallar (Eds.), Optimising new modes of assessment: In search of qualities and standards (pp. 55–87). Dordrecht, Netherlands: Kluwer.

Topping, K. J., Smith, E. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. Assessment & Evaluation in Higher Education, 25(2), 149–169.

Tsai, C.-C. (2009). Internet-based peer assessment in high school settings. In L. T. W. Hin & R. Subramaniam (Eds.), Handbook of research on new media literacy at the K-12 level: Issues and challenges (pp. 743–754). Hershey, PA: Information Science Reference.

Tsai, C.-C., & Liang, J.-C. (2009). The development of science activities via on-line peer assessment: The role of scientific epistemological views. Instructional Science, 37(3), 293–310.

Tseng, S. C., & Tsai, C. C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. Computers & Education, 49(4), 1161–1174.

Yang, Y.-F. (2010). A reciprocal peer review system to support college students' writing. British Journal of Educational Technology. doi:10.1111/j.1467-8535.2010.01059.x.

Yang, Y.-F., & Tsai, C.-C. (2010). Conceptions of and approaches to learning through online peer assessment. Learning and Instruction, 20(1), 72–83.