Comparing the Use of Self and Peer Assessment: A Case Study in a Statistics Course

Kyung Soo Han[a], Gil Seong Mun[a], Jeong Yong Ahn[1,a]

[a]Department of Statistical Informatics, Chonbuk National University

Abstract

In this study, we compare the assessments made by self, peer and instructor in a statistics course. The goal is to investigate the following two questions: (1) Is it reasonable or fair to expect students to be responsible for assessing the work of their colleagues and themselves? (2) What are students' opinions about the learning effect after they participate in the assessment process? As part of the study investigating these questions, we designed a prototype for a Web-based assessment tool and a procedure to apply the assessment techniques in a statistics course. In addition, we collected and analyzed the data produced in the assessment processes from students and the instructor. The analysis results are summarized as follows: First, self assessment was not accord with instructor assessment, but peer assessment was similar to the assessment by instructor. This result reflected that it is reasonable or fair to expect students to be responsible for assessing the work of their colleagues. Second, peer assessment of their colleagues successfully helped students increase their understanding of the course, and the students increased their skills in the actual assessment process by assessing the work of their colleagues. Finally, many students indicated a high interest level on the assessments.

Keywords: Self and peer assessment, learning effectiveness, assessment data analysis.

1. Introduction

A traditional teaching approach is characterized by lectures requiring little or no active student involvement along with tests or exams emphasizing quantitative algorithmically solving procedures (Hake, 1998). Educators have widely acknowledged that this traditional teaching approach contributes to problems of misconception and unsatisfactory conceptual understanding in many areas (Cahyadi, 2004). Alternative approaches to solve these problems are constructivism and e-learning emphasizing active learning over passive learning. The search for more effective alternatives has been accompanied by intense research and evaluation efforts (Ahn *et al.*, 2009, 2005; Doignon and Falmagne, 1999; Jain *et al.*, 2007).

A recent issue to improve learning effectiveness is the practical use of assessment techniques. Assessment is the process of evaluating the extent to which participants in education have developed their knowledge, understanding, and abilities. It is one of the basic tasks of educational work and includes a variety of activities such as testing, problem solving, collaborative or individual projects' development, and participation in discussions. For the student, the evaluation is the moment for verifying knowledge acquired in the learning process. In addition, the evaluation process allows the student to compare his acquired knowledge with the expected amount of knowledge in a formative action and thus to determine if the educational objectives have been correctly achieved (Barchino *et al.*, 2006).

[1] Corresponding author: Professor, Department of Statistical Informatics(Institute of applied statistics), Chonbuk National University, Jeonju, 561-756, Korea. E-mail: jyahn@chonbuk.ac.kr

The assessment takes an important role, as much in education based in e-learning systems, as in traditional or actual education. Online assessment gives convenient feedback to all participants in the e-learning process, helping to improve the learning and teaching experience (Barbosa and Garcia, 2005; Dureva and Tuparov, 2006).

In this study, we compare the assessments made by self, peer and instructor in a statistics course. The goal of this case study is to investigate the following two questions: First, is it reasonable or fair to expect students to be responsible for assessing the work of their colleagues and themselves? Second, what are students' opinions about the learning effect after they participate in the assessment process? As part of the study investigating these questions, we designed a prototype for a Web-based assessment system and analyzed the data produced in the assessment process from students and the instructor.

This paper is structured as follows: In Section 2, we review briefly traditional and new forms of assessment techniques. In Section 3, we introduce a prototype of a Web-based assessment system we designed in this study, and describe an example of a learning group and a procedure to apply self and peer assessment. In Section 4, we report on the analysis results of the collected assessment data, and discuss the assessment ability of students and the learning effectiveness of the assessment.

2. Assessment Types

There are many assessment types. Traditional assessment is a term used to describe the means of gathering information on student learning through techniques such as testing students with multiple-choice, fill-in-the-blank, and matching questions, as well as essays (Bouchillon *et al.*, 1999). The approaches are generally used as part of the process of assessment of students by an instructor in off-line environments.

There are many alternatives to traditional assessment types that can be used to broaden the scope of the teacher's classroom assessment activities. The typical techniques of the alternatives are self, portfolio, and peer assessment.

- Self assessment

  Self assessment is an assessment technique where students perform self evaluations of their work. It allows the student to verify his/her level of competence, to evaluate his/her learning process, and to develop autonomous learning skills (Gentile & Plantamura, 2003). An advantage of self assessment is that it enables students to take greater responsibility for, and to become actively involved in, the assessment part of the learning process that is now regarded as such an important part of the education process (Ellington *et al.*, 1997). Because most students initially lack the experience and judgment to carry out effective self assessment without special training, however, instructors can implement strategies to support the development of students' abilities to assess their own work.

- Portfolio assessment

  Portfolio assessment is a technique of using a learning portfolio on assessment. A learning portfolio is a purposeful collection of student work that exhibits the student's efforts, progress, and achievements in one or more areas (Paulson *et al.*, 1991; Love and Cooper, 2004). Traditional assessment measures a student's ability at one time and is conducted outside instruction. On the other hand, portfolio assessment measures a student's ability over time and the process is embedded in instruction. The main advantage of portfolio-based assessment is that a portfolio provides a means for students to learn to manage their own professional development because portfolio

development provides a straightforward means for students to collect evidence of professional or generic graduate skills (Cooper and Love, 2002; Love and Cooper, 2004).

- Peer assessment

  Peer assessment is the process of assessment of students by other students, both formative reviews to provide feedback and summative grading (A'ali, 2007; Lai and Lan, 2006). The peer assessment actively involves students in the assessment process, thus giving them a greater sense of ownership of and empathy for the judgments produced by such an assessment. It also encourages students to reflect on the standards of their own work compared to that of others - a benchmarking process that can, in itself, be a powerful aid to learning (Ellington *et al.*, 1997).

  Teacher-only assessment limits the opportunity for individual evaluation skills and the opportunity for students to understand the educational objectives and how these relate to the educational experiences. Peer assessment, however, is intended to help students plan their learning, identify their strengths and weaknesses, and develop transferable skills. It is a reliable and valid method for assessment and teaching (A'ali, 2007; Topping, 1998).

Traditional assessment techniques are heavily influenced by old paradigm such as the behaviourist learning theory, the belief in objective and standardized testing, and testing being separated from instruction (Shepard, 2000; Watering *et al.*, 2008). New assessment techniques, however, are continually being developed to measure students' progress toward achieving new performance standards and benchmarks (Bouchillon *et al.*, 1999). These techniques have characteristics: First, new forms of assessment are typically embedded in an educational context and rely on higher levels of student involvement (Sluijsmans, 2004). Second, various types of tasks are performed in assessment processes. Tasks are arranged for candidates to demonstrate their progress and capabilities such as answering a questionnaire, writing an article, providing a portfolio, conducting a performance, and so on. Third, a large quantity of information is produced by participants in performing various tasks in different phases (Miao *et al.*, 2007). Finally, these new forms of delivering education need the practical use of information technologies for real adoption. If information technologies are not used, it is very hard to successfully monitor the learning attitudes of students, obtain real-time feedback, and so on.

3. Designing and Executing Assessment for a Statistics Course

There are good reasons to involve students with assessment. Self and peer assessment can lead to interesting and interactive lessons (Orpen, 1982; Wen and Tsai, 2006). They provide a rapid way for instructors to assess large volumes of student work (Wheater *et al.*, 2005). In this study, we used the assessments in an introductory statistics course. The reason for choosing the assessments is: It is generally agreed that the assessments enhance the quality and effectiveness of the learning process since assessing is a powerful aid to mastering educational materials (Ellington *et al.*, 1997). The assessments, however, are often used for assessing essays and have seldom been applied to mathematics courses. We wanted to investigate whether the advantages of the assessments used in areas such as essays applied to mathematics courses, too.

3.1. Designing a Web-based assessment tool

Self and peer assessment techniques can be completed both on-line and off-line. Off-line, however, is not as rewarding because the sharing of the learning works of students is very difficult and communication about assessment results requires large amounts of time in terms of peer assessment. An online

system, using information technologies, for effective and practical use of the assessment technique, can substantially help alleviate these issues.

In this study, we designed a prototype of a Web-based assessment system that can support assessment activities of students in a statistics course. Students submit assignments, participate in assessment, and share the learning works of other students through this system. This process helps students improve the quality of their work thus enlarging learning effectiveness. This form of learning can enlarge the students' frame of reference for problem solving.

On the other hand, the tools for mathematical courses such as the introductory statistics have to be able to handle numerical formulae on the Web. Many Web-based education systems in the past, however, handle numerical formulae in a restrictive manner using images and HTML code. The prototype developed in this study displays mathematical formulae using MathPlayer from Design Science. In addition, we designed the prototype so that students can share freely other students' work after the peer assessment.

Some features of the system are: First, the system provides a platform where students describe the solving process of problems assigned to each of them. This platform encourages students to participate in learning and it helps improve learning effectiveness. Second, the data, for example the learning works of students, the results of peer assessment, and the opinions of the students, were then delivered, in real time, to the instructor. The instructor, after checking the results, gave feedback on all of the students' work. The information is used for monitoring the learning attitudes of students and real-time feedback.

3.2. Executing peer assessment in a statistics course

The introductory statistics course is generally taken by freshman students. We executed the assessment in the second semester, 2008. The participants in this study were 24 first year undergraduate students of the department of statistical informatics at Chonbuk National University. 13 students in all were female and others were male.

Students find this course difficult; therefore, additional support for teaching the material is useful. In this course, a benchmarking process, their own work compared to that of others, can be a powerful aid to learning. An assessment exercise in this study was divided into four stages:

- stage 1: The instructor explained the self and peer assessment procedure and assigned a task to the students. In this study, we were given the same assignment to all students. One example problem from the course was: Calculate $E(X)$ when X is binomially distributed with parameters n and p.

- stage 2: The students performed the assignment on their own time. Then, they submitted the assignment with a self assessment grade on the Web-based assessment system (Figure 1(a)).

- stage 3: Each student was assigned to review three other students' assignments. In other words, a student's work was reviewed by three other students selected in random. Students reviewed other students' work, provided specific feedback on any points raised by commenting, and assigned an assessment grade on the work (Figure 1(b)).

- stage 4: The instructor then provided the assessment grade and commented on all students' work, respectively (Figure 1(c)).
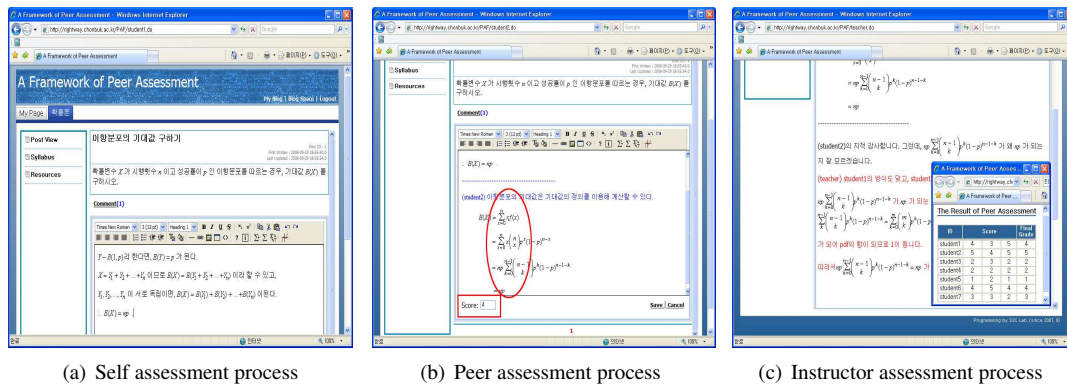
(a) Self assessment process (b) Peer assessment process (c) Instructor assessment process

Figure 1: *Web-based assessment system*

Table 1: Assessment data

| student ID | self | peer | | | | instructor |
		peer1	peer2	peer3	final	
1	4	5	4	5	5	5
2	5	3	5	4	4	4
3	3	3	3	2	3	2
⋮	⋮	⋮	⋮	⋮	⋮	⋮

4. Analysis of the Assessment Data

New online peer-assessment programs allow the reintroduction of open-ended written assignments in a manner that is logistically and economically efficient. However, even more important than that is the fact that these assignments provide extremely rich contexts for supporting deep analysis skills (Joordens *et al.*, 2009). In this study, assessment was experimentally applied using 3 open-ended problems and the students' work was assessed using a 1-to-5 scale. The final grade of the peer assessment, in stage 3, was assigned by using the mode or median of three grades. If the work of a student received the marks 4, 5 and 4, for example, the final grade of the work was assigned as 4. Similarly, the final grade on the marks 4, 2 and 1 was assigned as 2. Therefore, we collected 3 data sets as Table 1.

4.1. Comparison between self assessment and instructor assessment

Most students of our course have not had the experience and judgement to carry out effective self assessment. Therefore, there has come out a natural tendency for students to over-value their own efforts. Table 2 shows the result of comparison between self and instructor assessment grade. In the first assessment, many students (75%) assigned high grade on their own efforts than instructor. However, the differences were on the decrease according as the assessment advances. In particular, it is interesting to note that 16.7% of students assigned low score than instructor in the third assessment.

On the other hand, each Kappa statistic of three assessments, a generic term for several similar measures of agreement used with categorical raters, was not statistically significant. This analysis result shows that self assessment grades are not accord with the grades by instructor. We think, therefore, self assessment can be used to encourage a climate of self criticism rather than assessment in itself.

Table 2: Comparison of grades by self and instructor assessment

	Assessment		
	first	second	third
self > instructor	75.0%	62.5%	45.8%
self = instructor	25.0%	37.5%	37.5%
self < instructor	0.0%	0.0%	16.7%
Kappa statistic	0.083	0.180	0.167
(*p*-value)	(0.325)	(0.073)	(0.135)

Table 3: Comparison of grades by peer and instructor assessment

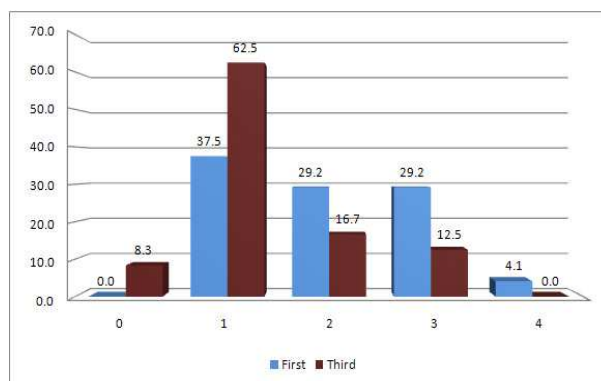|	Assessment		
	first	second	third
peer > instructor	66.7%	50.0%	33.3%
peer = instructor	25.0%	41.7%	54.1%
peer < instructor	8.3%	8.3%	12.5%
Kappa statistic	0.029	0.228	0.387
(*p*-value)	(0.780)	(0.031)	(0.001)

Figure 2: *Differences in peer assessments*

4.2. Comparison of peer assessments

In the Table 1, the work of the first student (student ID: 1) received the marks 5, 4 and 5. The difference between maximum and minimum grade in three peer assessments is 1. Figure 2 shows the distribution of the differences in the first and third assessment. In the first assessment, the differences of 2 or above were 63.5%. In the third assessment, however, the differences have fallen off to 39.2%. As a result, the differences were on the decrease according as the assessment advances, too.

4.3. Comparison between peer assessment and instructor assessment

The results of the peer assessment (the final grade) were compared with the instructor assessment as shown in Table 3. In the first assessment, only 25% of the final grades were the same with the grade by instructor. However, the proportion has increased by 54.1% in the third assessment. The Kappa statistic on the third assessment was 0.387 (*p*-value: 0.001). Therefore, the analysis results reflected the following: If we provide precise criteria and training environment, it is reasonable or fair to expect students to be responsible for assessing the work of their colleagues.

At the end of the process, on the other hand, we surveyed with a simple online questionnaire. Fig-
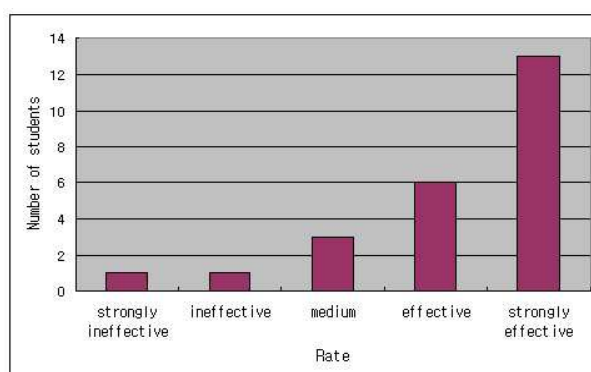
Figure 3: *Learning effectiveness*

ure 3 shows the response of students on learning effectiveness after they participated in the assessment process. 79% of students (19 among 24 students) replied that the process provided a powerful aid for understanding the learning contents.

5. Conclusions

Recent advances in Internet technologies have helped motivate extensive use of various assessment techniques. Peer assessment is a special form of collaborative learning, where peer students learn through assessing others' work (Miao and Koper, 2007). We, in this study, applied self and peer assessment techniques in a statistics course and investigated its learning effectiveness. To apply the techniques, we designed a prototype for a Web-based assessment system. As a result, most of students evaluated their colleagues on a reasonable basis nearly similar to the instructor's evaluations and expressed an affirmative answer on learning effectiveness. In addition, many students indicated a high interest level. Online assessment in an e-learning course is considered a very important tool as it can provide convenient feedback and help in improving learning effectiveness. There, however, are several remaining problems for future explorations. First, for example, some students may not wish to be responsible for assessing their colleagues. We have to consider how to cope with the situation by understanding the causes. Second, students generally lack experience in self and peer assessment. We, therefore, need very careful planning and precise criteria. Third, the assessments also requires a complete system that can support these assessment activities.

References

A'ali, M. (2007). Adoption of peer-to-peer assessment in a computing skills course, *American Journal of Applied Sciences,* **4**, 828–833.

Ahn, J. Y., Han, K. S. and Han, B. S. (2005). Web-based education: Characteristics, problems, and some solutions, *International Journal of Innovation and Learning,* **2**, 274–282.

Ahn, J. Y., Mun, G. S. and Han, K. S. (2009). Comparing self and peer assessment with instructor assessment in a statistics course, *Proceedings of the IMS-APRM & 2009 Conference of the Korean Statistical Society,* 183.

Barbosa, H. and Garcia, F. (2005). Importance of online assessment in the e-learning process, *Proceedings of the Conference on Information Technology Based Higher Education and Training,* F3B/1-F3B/6.

Barchino, R., Gutierrez, J. M., Oton, S., Martinez, J. J., Hilera, J. R. and Gutierrez de Mesa, J. A. (2006). E-learning model for assessment, *Proceedings of the Virtual Multi Conference on Computer Science and Information Systems.*

Bouchillon, W. H., Meyer, T. and White, J. (1999). *21st century assessment (One in a series of products for Preparing All Learners PAL for the Workforce),* The University of West Florida.

Cahyadi, V. (2004). The effect of interactive engagement teaching on student understanding of introductory physics at the faculty of engineering, University of Surabaya, Indonesia. *Higher Education Research* & *Development,* **23**, 455–464.

Cooper, T. and Love, T. (2002). Online portfolios: Issues of assessment and pedagogy, *In Jeffrey, P.(Ed.), AARE 2001: Crossing Borders: New Frontiers of Educational Research.*

Doignon, J. P. and Falmagne, J. C. (1999). *Knowledge Spaces,* Springer, Germany.

Dureva, D. and Tuparov, G. (2006). Assessment models in e-learning environments, *Proceedings of the Conference on Computer Systems and Technologies,* IV.7-1 - IV.7-5.

Ellington, H., Earl, S. and Cowan, J. (1997). Making effective use of peer and self assessment, *Technical Paper,* The Robert Gordon University, England.

Gentile, E. and Plantamura, P. (2003). Assessment and self-assessment in e-learning courseware, *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education,* 1587–1590.

Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *American Journal of Physics,* **66**, 64–74.

Jain, L. C., Tedman, R. A. and Tedman, D. K. (2007). *Evolution of Teaching and Learning Paradigms in Intelligent Environment*, Springer, Germany.

Joordens, S., Desa, S. and Pare, D. (2009). The pedagogical anatomy of peer-assessment: Dissecting a peer Scholar assignment, *The Journal on Systemics, Cybernetics and Informatics,* **7**, 1–6.

Lai, K. R. and Lan, C. H. (2006). Modelling peer assessment as agent negotiation in a computer supported collaborative learning environment, *Journal of Educational Technology and Society,* **9**, 16–26.

Love, T. and Cooper, T. (2004). Designing online information systems for portfolio-based assessment: Design criteria and heuristics, *Journal of Information Technology Education,* **3**, 65–81.

Miao, Y. and Koper, R. (2007). An efficient and flexible technical approach to develop and deliver online peer assessment, *Proceedings of the Conference on Computer Supported Collaborative Learning,* 502–511.

Miao, Y., Tattersall, C., Schoonenboom, J., Stevanov, K. and Petrova, A. A. (2007). Using open technical e-learning standards and service-orientation to support new forms of e-assessment, *Proceedings of the TENCompetence.*

Orpen, C. (1982). Student versus lecturer assessment of learning: A research note, *Higher Education,* **11**, 567–572.

Paulson, F. L., Paulson, P. R. and Meyer, C. (1991). What makes a portfolio a portfolio?, *Educational Leadership,* **48**, 60–63.

Shepard, L. A. (2000). The role of assessment in a learning culture, *Educational Researcher,* **29**, 4–14.

Sluijsmans, D., Brand-Gruwel, S., van Merrienboer, J. and Martens, R. (2004). Training teachers in peer-assessment skills: effects on performance and perceptions, *Innovations in Education* & *Teaching International,* **41**, 60–79.

Topping, K. J. (1998). Peer assessment between students in colleges and universities, *Review of Educational Research,* 68, 249–276.

Watering, G., Gijbels, D., Dochy, F. and Rijt, J. (2008). Students' assessment preferences, perceptions of assessment and their relationships to study results, *Higher Education,* **56**, 645–658.

Wen, M. L. and Tsai, C. (2006). University students' perceptions of and sttitudes toward (online) peer assessment, *Higher Education,* **51**, 27–44.

Wheater, P., Langan, A. M. and Dunleavy, P. (2005). Students assessing student: Case studies on peer assessment, *Planet,* **15**, 13–15.