

## The Influence of SGRs on Self, Peer, and Instructor Evaluation in Higher Education

Seongyoun HONG

Seoul National University

Korea

This research has three purposes: The first purpose is to verify whether SGRs(Student-Generated Rubrics) are effective for learner achievement in higher education or not. The second one is to analyze the agreements between an instructor's evaluations and peer group assessments with SGRs in real classroom situation. The final purpose is to propose the possibility of students' more objective self-evaluation with the use of SGRs. It appears that rubrics are effective tools to improve learner achievement in adult education and to promote the agreement between the instructor's and peer's evaluations, even though learners do not develop SGRs. In addition, this research has demonstrated that most students are not very accurate self-assessors, and they often overestimate themselves. This study can be helpful for novice instructors, especially for those who try to assess their students more fairly, to develop the SGRs as an evaluation tool appropriate for their own situation, and to give more reliable feedback to students.

*Keywords : SGRs(Student-Generated Rubrics), self-evaluation, instructor's evaluations, peer's evaluations*

## **Introduction**

Rubrics are tools for evaluating product-oriented assignments and informing students of instructor's expectation (Han, 2004). However it is not easy for instructors to design rubrics suitable for each purpose due to differences in students' intellectual levels, and various teaching and learning environments. In particular, novice instructors have inadequate experience to evaluate students and have few assignment samples. Novice instructors, therefore, need to find a way to design the rubrics which are appropriate for students' intellectual standards and their circumstances.

Student-generated rubrics (SGRs) are made through negotiation between a teacher and students. Thus, teachers can estimate students' standards and consider unique characteristics of diverse teaching and learning environments. Moreover, inviting students to develop their own evaluation structure enhances students' motivation, interests, and performances in projects. According to previous research, the SGRs have a significant effect on motivation to learn in terms of enhancing students' level of attention and satisfaction (Ainsworth & Christinson, 1997; Skillings & Ferrell, 2000; Han, 2004). However, this research has been only conducted to K-12, not to adult learners.

Furthermore, well-made rubrics are easy to be marked not only for novice instructors but also for students. In higher education, self and peer assessments give huge potential benefits in teaching and learning; both for the assessor and for the assessee. Self and peer assessments encourages students' degree of autonomy and higher order thinking skills. In adult higher education, self and peer assessment techniques are used widely. Through the experience of self and peer assessments, students can objectify themselves and compare their own assignment to those of others (Kang & Lee, 2006). In addition, if the developed SGRs work, the agreement between instructor and students rating should have similar results.

As a method of improving students' performance and confidence in assessment,

this study suggests the practice of SGRs in higher education settings, investigation of the objectivity of students' self evaluation, and comparison of instructor's evaluation with that of their peers. This research has three purposes: The first purpose is to verify whether SGRs are effective for learner achievement in higher education or not. The second one is to analyze the agreements between an instructor's evaluations and peer group assessments with SGRs in real classroom situation. The final purpose is to propose the possibility of students' more objective self-evaluation with the use of SGRs.

## **Theoretical Background**

### **Student-Generated Rubrics**

Student-generated rubrics (SGRs) are the rubrics which consist of specific pre-established performance criteria which are developed by negotiation between a teacher and students (Han, 2004). Involving students in the development of performance targets can be an effective instructional tool because students who are given the task of analyzing quality work and its critical components become better performers themselves (Stiggins, 1997).

Stix (1997) refers to this rubric co-construction process as negotiable contracting. His research shows that students given a role in assessment process and provided with appropriate direction by teachers are able to more accurately evaluate their strengths and weaknesses and better pinpoint areas where they need to focus their efforts for improvement. In aspects of intellectual factors, students understand learning goals and concrete elements of proficiency, by creating rubrics. In terms of motivational factors, by knowing the scoring guide, SGRs enhance students' confidence on performance tasks, and students try reaching criteria which they created beforehand. Moreover, instructors can also develop the rubrics appropriate to

their situations.

Han (2004) modified Taggart & Wood (2001)'s process of rubric development into six steps. Firstly, instructors verify criteria of achievement in knowledge, skill, and motivation. Next, instructors decide their expectation of the performance task. Thirdly, instructors choose the form of the rubrics. And then, concrete indicators (descriptors) of proficiency are created by negotiation between instructors and students. Fifthly, instructors and students decide on the performance criteria. Lastly, SGRs are applied to the real assessment and adjustment.

The negotiation decides elements of assessment and describes performance as criteria. Most of the instructors help students to understand what rubrics are and how students can perform the rubric process. Then, instructors give the data related to students' performance task, and remind the students' learning and future performance. After that, students discuss important elements of the performance and choose elements of assessment. Finally, students suggest positive and negative traits of performer's behavior and instructors adjust the criteria by taking students' description into account.

### **Self and Peer Assessing in Higher Education**

Peer evaluation is widely used as an evaluation method in school. Students assess others' performance, and then results are considered into their grade. Peer evaluation can be a good alternative measurement, especially if the student size is too large for instructors to assess all of the students objectively (Kang & Lee, 2006; Baek, 2002).

Topping (1998) reviewed the 31 studies of the reliability and validity of peer assessment compared marks or grades. Out of the 31 studies, he focused on 18 studies and suggested that peer assessment was of adequate reliability and validity in a wide variety of applications. However, 7 studies have been found with unacceptably low level of reliability and validity for peer assessment in particular projects (Topping, 1998).

Not only that, according to research conducted before Toppings, self and peer assessments has educational effects. Peer assessments enhance students' opportunities to discuss their work with others (Boud, 1995). In the process of evaluation, students try to review, summarize, clarify, give feedback, diagnose misconceived knowledge, identify missing knowledge, and consider deviations from the ideal as assessors (Topping, 1998). From this process, students develop responsibilities and a sense of ownership for their peers' learning (Dochy, Segers, & Sluismans, 1999; Orsmond, Merry, & Reiling, 1996; Topping, Smith, Swanson, & Elliot, 2000).

In previous research, most students were not very accurate self-assessors (Bjork, 1999). Salden, Paas, Van der Pal, and Van Merriënboer (2006) asked students to self-assess their overall task performance on a rating scale. Analyzing students' self-assessment score log files and comparing the "objective" performance scores to the self-assessment scores, Salden et al. (2006) found that 67% of participants tended to overestimate their performance during training. Even in test situations based on the lists of scoring criteria, the result of self-assessments is likely to turn out to be higher compared to that of teacher's or peer assessments (Miller, 2003; Topping, 2003). Lorraine (1994) discovered peculiar phenomena of high achieving students' underestimating their performance and low achieving students' overestimating their performance (Lorraine A. J., 1994). However, Han (2006) suggested that students estimated themselves more objectively and strictly with SGRs, so self-assessments scores were lower than peer and instructor assessment scores.

## Method

### Participants

Three classes titled as 'Technical Communication' in a university in Seoul were

examined for two semesters at a university in Seoul. More than 100 graduate students in each phase with all mixed-up level participated in this experiment.

In phase one, total number of 113 undergraduate students in a class were targeted. Each class consisted of 8 groups. After about one and a half months of oral presentation lessons, the first 4 groups performed oral presentation in the first week out of two-weeks' presentation test.

The first class developed their own SGRs and evaluated themselves and their peer presenters. The second class just used the SGRs developed by the first group for self and peer group evaluation. In contrast, the third class rated with a form which focused on evaluation point. Every student assessed the self group and peer groups individually using rubrics or evaluation form.

In phase 2, the participants were 103 undergraduate students of the 3 'Technical Communication' classes. The process of research followed nearly the same as the phase one. The only difference was that the third class received the rubrics 1 week before the presentation to reconfirm the gap between the SGR groups and the just used groups.

## Procedure

First, a teacher and the students of SGR groups negotiated and developed the Student-Generated Rubrics based on the proceeding research (Taggart & Wood, 2001; Ainsworth & Christinson, 1998; Han, 2004) one week before the first oral presentation.

The process of creating SGRs was as such: (1) Instructor explained purpose of creating SGRs and informed participants that it would be used as assessment tools. (2) Students could use ready-made rubrics to understand what rubrics were. (3) Instructor gave examples of rubrics for presentation which were created by other teachers and research. (4) Instructor explained and shared their expectation of the task which students would do. (5) Students decided standards and levels of rubrics

and gave score according to standards. (6) Concrete indicators (descriptors) of proficiency were devised by negotiation between instructors and students. (7) Other two instructors of the same subjects reviewed the student-generated rubrics. (8) Students and instructor applies SGRs to real assessment and adjustment.

Second, after six-week-presentation classes, oral presentation test was held for 2 weeks.

In phase 1, SGR class used its own rubrics. Another class received the rubrics developed by the SGR class. And the other class was presented rubrics with only standards eliminated levels.

In phase 2, in next semester, similar research was performed. The only difference was that the third class received the rubrics one week before the presentation test to reconfirm the gap between the SGR groups and the just used groups.

Thirdly, the students of the SGR class and the class used the rubrics participated in survey of the satisfaction of their rubrics.

## Measurements

To analysis the difference in the learners' achievements among three experimental groups, One-way ANOVA was performed. Independent variables are the difference of using rubrics among classes. Dependent variable is score which is assessed by instructor and students. Additionally, to examine which group's score is different from others, Scheffe post-hoc analysis was carried out to divide the classes.

To understand the linear relationship between instructor's evaluation and peer groups' evaluation, the correlation between instructor's evaluation and peer evaluation was analyzed. Pearson's correlation method was adopted in order to acquire a coefficient.

## Result

### The Development of Student-Generated Rubrics

In phase 1 and phase 2, the participants developed the different SGRs which was appropriate their teaching and learning environment. The result of SGRs is shown on Table 1 and Table 2.

Table 1. The Result of SGRs in Phase 1

Standards	5	4	3	2	Score
Attitude & Skill (15)	① Constant eye contact ② Appropriate expression ③ Natural hand gestures ④ Presentation time management ⑤ Stands up straight with both feet on the ground	Satisfied 4 items	Satisfied 2-3 items	Satisfied less than 1 item	
Visualization (20)	① Relative effectiveness of text and image ② Appropriate aids(photo, animation, sound etc) ③ Attractive design ④ Intuitive design ⑤ Clear expression	Satisfied 4 items	Satisfied 2-3 items	Satisfied less than 1 item	
Contents (20)	① Consistent focus on a clear and compelling thesis ② Logical and coherent structure with claims ③ Evidence and interpretations supporting the thesis	① Consistent focus on a clear thesis and a logical sequence of claims ② Lack of evidence to support the thesis	Lack the coherence of consistent claims, evidence, and interpretations	Lack of organization and coherence and fail to maintain focus	
Creativity (15)	Unique idea which is not look like the others	Good idea which can be improved	Temporary expedient and complementary measures	Tried idea and no originality	
Usability (10)	① Usability in real situation ② Effective and opportune suggestion	① Realization with modification ② Effective and opportune suggestion	Realization with modification, but ineffectiveness	Irrelevant suggestion	
Understanding (20)	Accurate and in-depth understanding of the topic, audience, and purpose	Accurate although somewhat basic understanding of the topic, audience, and purpose	Confused or largely inaccurate understanding of the topic, audience, and purpose	No evidence of understanding and incoherent explanations	

The Influence of SGRs on Self, Peer, and Instructor Evaluation in Higher Education

Table 2. The Result of SGRs in Phase 2

Standards	Items	( ) ( ) ( ) ( )
Attitude & Skill (30)	Constant eye contact	5 4 3 2 1
	Natural hand gestures	5 4 3 2 1
	Appropriate expression	5 4 3 2 1
	presentation time management	5 4 3 2 1
	Stands up straight with both feet on the ground	5 4 3 2 1
	Decent appearance	5 4 3 2 1
	Understanding of audience	5 4 3 2 1
Visualization (20)	① Relative effectiveness of text and image	5 4 3 2 1
	Appropriate aids(photo, animation, sound etc)	5 4 3 2 1
	③ Attractive design	5 4 3 2 1
	Intuitive design	5 4 3 2 1
	Clear expression	5 4 3 2 1
Contents (30)	① Consistent focus on a clear and compelling thesis	5 4 3 2 1
	② Logical and coherent structure with claims	5 4 3 2 1
	③ Evidence and interpretations supporting the thesis	5 4 3 2 1
	Creativity	5 4 3 2 1
	Usability	5 4 3 2 1
Understanding (10)	Accurate and in-depth understanding of the topic	5 4 3 2 1
	Accurate and in-depth understanding of the purpose	5 4 3 2 1
	Accurate understanding the question	5 4 3 2 1
	Persuasive answer	5 4 3 2 1

Results of the Learner achievement

Phase 1, as indicated in Table 3, the highest class of learner achievements was the class which was given SGRs before the presentation (SGR class=74.4919, Ontime

class=77.6071, noSGR class=70.1688). The result of one-way ANOVA tests, Table 4, indicated there were statistically significant differences on three groups ( $F=10.352$ ,  $p<0.01$ ). Follow-up Scheffe post hoc tests, Table 5, indicates the means of achievement in two class using SGRs and developing SGRs are statistically significantly higher than the means of not using SGRs.

Table 3. The Result of Descriptive Statistic Analysis in Phase 1

	N	Mean	Std.Deviation	Std.Error
SGR	31	74.4919	7.2656	1.3049
Ontime	42	77.6071	7.0352	1.0856
noSGR	40	70.1688	7.9114	1.2509
Total	113	74.1195	8.0142	.7539

Table 4. The Result of ANOVA Analysis in Phase 1

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1139.510	2	569.755	10.352	.000
Within Groups	6054.002	110	55.036		
Total	7193.512	112			

Table 5. The Result of Scheffe Post hoc Tests in Phase 1

	N	Subset for alpha = .05	
		1	2
noSGR	40	70.1688	
SGR	31		74.4919
Ontime	42		77.6071
Sig.		1.000	.200

These results confirm that the SGRs have significant effect on learning achievement in adult education. However, it was unexpected that the means of achievement of using SGR class was higher than that of developing SGR class.

In phase 2, a one-way ANOVA was conducted to compare achievement among the developing SGR class, the presenting SGR class of 1 week before the presentation, and the presenting SGR class on the time of presentation. As indicated in Table 6, the highest class of learner achievement is the class which was given SGRs 1 week before the presentation (SGR class=69.9310, 1 week before class=70.0972, on time class=68.1711). Nevertheless, the result of one-way ANOVA tests, Table 7, indicates that there are not statistically significant classes differences as using SGRs ( $F=10.352$ ,  $p>0.5$ ).

Table 6. The result of descriptive statistic analysis in phase 2

	N	Mean	Std.Deviation	Std.Error
SGR	29	69.9310	8.9967	1.6707
1weekbf	36	70.0972	8.2504	1.3751
ontime	38	68.1711	8.3296	1.3512
Total	103	69.3398	8.4590	.8335

Table 7. The result of ANOVA analysis in phase 2

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	82.698	2	41.348	.573	.566
Within Groups	7215.910	100	72.159		
Total	7298.607	102			

These results confirm that the SGRs have significant effect on learning achievement in adult education in unrelated to the method of presentation.

### Agreement between instructor's evaluation and peer evaluation

Pearson's correlation coefficients were used to examine the relationships of agreement between instructor's evaluation and peer evaluation. In phase 1, the positive correlation between instructor's evaluation and peer evaluation when they developed and presented SGRs, Table 8, is statistically significant ( $p < 0.01$ ,  $p < 0.05$ ). However, the correlation between instructor's evaluation and peer evaluation when they do not use SGRs is not statistically significant ( $p > 0.05$ ). As shown in Table 8, the degree of linear dependence between instructor's evaluation and peer evaluation in SGR class is the strongest than others ( $r = .879^{**}$ ).

Table 8. The result of Pearson Correlation analysis in phase 1

	Pearson Correlation	Sig.(2-tailed)
SGR class	.879**	.004
Ontime class	.721*	.044
noSGR class	.655	.078

In phase 2, the correlation between instructor's evaluation and peer evaluation when they developed and presented SGRs one week before, Table 9, is statistically significant ( $p < 0.05$ ). As shown in Table 9, the strength of a linear relationship between instructor's evaluation and peer evaluation is the greater ( $r = .737^*$ ,  $r = .800^*$ ). However, the correlation between instructor's evaluation and peer evaluation when they presented SGRs is not statistically significant ( $p > 0.05$ ) in contrast with phase 1.

Table 9. The result of Pearson Correlation analysis in phase 2

	Pearson Correlation	Sig.(2-tailed)
SGR class	.737*	.037
1 weekbf class	.800*	.017
Ontime class	.690	

### Difference self evaluation with peer evaluation

T-test was conduct to compare achievement between self evaluation and peer evaluation. If SGRs allowed students to evaluate themselves more objectively, the scores of self evaluation were not significantly different in the score of peer evaluation. Even though instructor-announcing self evaluation scores does not influence the grade of students and students checks the pledge to be evaluate impartially, the T-test, Table 10, indicates statistically significant difference between self evaluation and peer evaluation in all classes of phase 1 and 2 ( $p < 0.05$ ). As shown in graph 1, the scores of self evaluation are higher than that of peer evaluation (red line-self evaluation, green line-peer evaluation).

Table 10. The Result of T-test Analysis in Phase 1 and Phase 2

		Mean	Std. Deviation	Std. Error Mean	t	Sig.(2-tailed)	
Phase 1	SGR	self	78.0412	3.7665	1.3317	3.923	.006
		peer	68.0838	8.8060	3.1134		
	Overtime	self	84.6750	3.1453	1.1120	6.807	.000
		peer	73.4950	5.6791	2.0078		
	noSGR	self	79.3888	4.7894	1.6933	3.601	.009
		peer	70.8950	5.5016	1.9451		
Phase 2	SGR	self	40.0275	4.7968	1.6959	2.822	.026
		peer	34.1750	3.1234	1.1043		
	1weekbf	self	41.2100	3.9600	1.4001	4.451	.003
		peer	33.8825	4.2615	1.5067		
	Overtime	self	41.5775	4.9636	1.7549	5.052	.001
		peer	33.3613	2.6862	.9497		

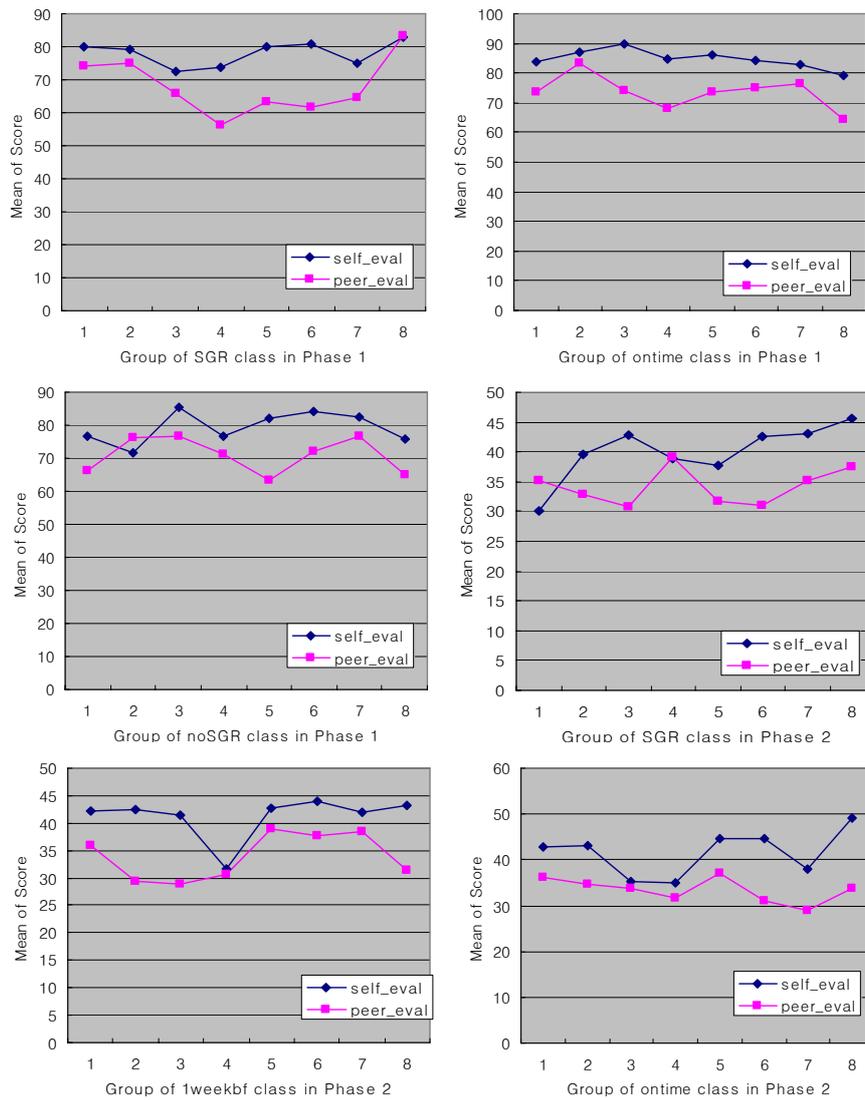


Figure 1. The Score Comparison between Self Evaluation and Peer Evaluation

## Conclusion

It appears that rubrics are effective tools to improve learner achievement in adult education, even though learners do not develop SGRs. By comparing, SGRs with general rubrics, the former is made by negotiation between instructors and learners in consideration of each situation and students can easily understand SGRs and evaluate the result of it (Han, 2004). Actually, Han (2004) suggested that students estimate themselves more objectively and strictly with SGRs in K-12. However, in the case of adult learners, the latter, which, like SGRs, are concerned with the intellectual level and situation of student, help students to understand the standard of evaluation and to perform well. In this research, the rubrics created by students were functionalized as a well made rubrics for the other students. It is probable that the adult learners can clearly understand the standard and assess their presentation, even if they do not develop the rubrics. In phase 2, well made rubrics were surely effective when instructors presented the rubrics at least one week before the performance test.

It is also evident that the rubrics are effective tools to promote the agreement between the instructor and peers for evaluations. As rubrics set standards for good performers and appropriate levels of evaluation, students actually can receive more fair assessment, as well as sense the reliability of their evaluation and a feeling of impartiality according to the result of the satisfaction survey.

Most students are not very accurate self-assessors, and they often overestimate themselves. In spite of the fact that learners have checked the pledge to evaluate impartially, the scores of self evaluation are far higher than the scores of peer evaluation. Because of this, when learners receive the feedback about their tests, sometimes they cannot agree on the result. Therefore, instructors give students their feedback with the results of peer evaluation.

However, it does not mean the SGRs are useless in adult education. If instructors do not have much experience to evaluate students or know thoroughly the characteristics of learners, it is too hard to develop the well made rubrics. In these

cases, instructors can make SGRs by negotiating with students. Furthermore, novice instructors are encouraged to employ rubrics to rate more justly and to give students a sense of reliability.

This study can be helpful for novice instructors, especially for those who try to assess their students more fairly, to develop the SGRs as an evaluation tool appropriated to their own situations, and to give more reliable feedback to their students.

## References

- Ainsworth, L. & Christinson, J. (1998). *Student-Generated Rubrics: An Assessment Model to Help All Students Succeed*. Orangeburg, New York: Dale Seymour Publications.
- Andrade, H. G. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership*, 57(5). Retrieved April 5, 2008 from <http://www.ascd.org/readingroom/edlead/0002/andrade.html>.
- Arter, J. A. & McTighe, J. (2001). *Scoring Rubrics in the Classroom: Using Performance Criteria for Assessing and Improving Student Performance*. Calif.: Corwin Press.
- Baek, S. G. (2002). *Performance Evaluation: Theoretical Approach*. Seoul: Educational Science Press.
- Dunbar, N. E. Brooks, C. E. & Kubicka-Miller, T. (2006). Oral Communication Skills in Higher Education: Using a Performance-Based Evaluation Rubric to Assess Communication Skills. *Innovative Higher Education*, 21(2). 115-128.
- Glass, K. T. (2005). *Curriculum design for writing instruction : creating standards-based lesson plans and rubrics*. Calif. : Corwin Press.
- Greguras, G. J., Robie, C. & Born, M. P. (2001). Applying the Social Relations Model to Self and Peer Evaluations. *Journal of Management Development*, 20(6). 508-525.
- Hafner, J. C. & Hafner, P. M., (2003). Quantitative Analysis of the Rubric as an Assessment Tool: an Empirical Study of Student Peer-Group Rating, *International Journal of Science Education*, 25(12). 1509-1528.
- Han, J. Y. (2004). Influence of Student-Generated Rubrics on Learning Process and Academic Achievement in Technology Education Assessment. Unpublished doctoral dissertation, Seoul National University, Seoul.
- Kang, A. & Lee, G. (2006). A Generalizability Theory Approach to Investigating the Generalizability of Performance Assessment Using Student Peer Reviews. *Journal of Education Evaluation*, 19(3). 107-121.
- Han, J. Y. (2006). The Influence of Student-Generated Rubrics on Reliability of Grade Result in Technology Education Assessment. *The Journal of Vocational Educational Research*, 25(3). 87-110.

- Lorraine, A. L. (1994). Peer, Self and Tutor Assessment: Relative Reliabilities. *Studies in Higher Education, 19*(1). 69-75.
- Luft, J. A. (1999). Rubrics: Design and Use in Science Teacher Education. *Journal of Science Teacher Education, 10*(2). 107-121.
- Lundberg, R. (1997). Student-Generated Assessment. *The Science Teacher, 64*(1). 50~53.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. Practical Assessment, *Research & Evaluation, 7*(25). Retrieved July 8, 2005 from <http://PAREonline.net/getvn.asp?v=7&n=25>.
- Olina, Z. & Sullivan, H. J. (2004). Student Self-Evaluation, Teacher Evaluation, and Learner Performance. *Educational Technology Research and Development, 52*(3). 5-22.
- Ozogul, G. Olina, Z. & Sullivan, H. (2006). Teacher, Self and Peer Evaluation of Lesson Plans Written by Preservice Teachers. *Education Tech Research, 56*. 181-201.
- Skillings, M. J. & Ferrell, R. (2000). Student-Generated Rubrics: Bringing Students into the Assessment Process. *The Reading Teacher, 53*(6). 452-458.
- Stiggins, R. J. (1994). *Student-Centered Classroom Assessment*, New York: Macmillan Publishing Company.
- Stix, A. (1997). Creating rubrics through negotiable contracting and assessment. US Department of Education: ERIC #TM027246.
- Taggart, G. L. & Wood, M. (2001). Rubrics: A Cross-Curricular Approach to Assessment. In Taggart, G. L., Phifer, S. J. Nixon, J. A. & Wood, M. ed. (2001). *Rubrics: A Handbook for Construction and Use*. The Scarecrow Press, Inc.
- Topping, K. (1998). Peer Assessment between Students in Colleges and Universities, *Review of Educational Research, 68*(3). 249-276.



**Seongyoun HONG**

Ph.D. candidate, Educational Technology, Seoul Nat'l Univ. & Researcher, Center for Teaching and Learning, Seoul Nat'l Univ.  
Interests: Distance Education, Constructive Learning Environment Design, e-Learning and Learning Support Design in Higher Education  
E-mail: hsy06@snu.ac.kr