

The Relationships among Online Question-Generation, Peer-Assessment and Academic Achievement

Fu-Yun Yu, National Cheng Kung University, Taiwan, fuyun.ncku@gmail.com
 Chun-Ping Wu*, TamKang University, Taiwan, cpwu303@gmail.com

Abstract: This study examined the relationships among student performances in question-generation, peer-assessment, and academic achievement in four different online peer-assessment learning environments. Eight fifth-grade classes (N=253) were randomly assigned to one of four different identity revelation modes (real-name, anonymous, nickname, and self-choice) and participated in the study for six weeks. An online learning system that allows students to contribute to and benefit from the process of question-generation and peer-assessment was adopted. Data analysis revealed significant relationships among the three examined variables. Additionally, the identity revelation modes moderated the predictive power of online question-generation performance on academic achievement. Specifically, the predictive power in the self-choice mode was the greatest. Finally, self-choice identity mode strengthened the relationships between peer-assessment and question-generation performances. The empirical significance of the study and suggestions for learning system development and instructional implementation are provided.

Introduction

In response to the call for the diversification of question sources and an emphasis on engaging students in the process of learning and knowledge construction, student question-generation has been increasing in popularity over the past decades (Brown & Walter, 2005; English, 1998; Silver, 1994). Studies on student question-generation have substantiated its efficacy in enhancing learners' retention of information, reading comprehension, motivation, in-group communication and interaction, satisfaction with past learning experiences, and problem-solving abilities (Abramovich & Cho, 2006; Barlow & Cates, 2006; Brown & Walter, 2005; Dori & Herscovitz, 1999; Leung & Wu, 1999; Silver, 1994; Yu & Liu, 2005; Yu, Liu, & Chan, 2005). Even though most studies have supported the student question-generation approach, investigations into factors that may affect learning performances during student question-generation activities are scarce.

The effects of differences in personal characteristics and technology integration methods on adoption and use of student question-generation have been examined and yielded useful suggestions for instructional implementation (Wu & Yu, 2009; Wu & Yu, 2010). In view of the fact that students normally have limited experience generating questions during formal schooling (Moses, Bjork, & Goldenberg, 1993; Vreman-de Olde & de Jong, 2004; Yu & Liu, 2009), their performance in the task itself (i.e., question-generation) and its relationship to academic achievement in the applied contents is one area worth examining.

Moreover, to allow question-authors to receive timely and personalized feedback about contributed question items, peer-assessment is coupled with question-generation activities in most systems (Yu, 2009). Empirical evidence, which spans more than two decades, supports its facilitative effects for learners' motivation, higher-order thinking, cognitive re-structuring, level of performance and attitudes (Brindley & Scofield, 1998; Falchikov & Goldfinch, 2000; Gatfield, 1999; Hanrahan & Isaacs, 2001; Purchase, 2000; Topping, 1998; Tsai, Lin, & Yuan, 2002; van Gennip, Segers, & Tillema, 2009; Venables & Summit, 2003; Wen & Tsai, 2006; Wen, Tsai, & Chang, 2006). However, students' performances in peer-assessment activities itself in relation to their learning (as reflected in their academic achievement) have rarely been investigated.

In an attempt to gain some knowledge on these areas, a study was conducted to examine whether students' performances in question-generation are related to academic achievement. Moreover, issues on what the relationships between peer-assessment performances and academic achievement and whether students with better question-generation performances offer peers better comments were also examined. Finally, since different forms of user identification in student question-generation and peer-assessment has been reported to lead to different psychological reactions (Yu & Liu, 2009), whether user identification moderates the relationships among question-generation, peer-assessment and academic achievement was investigated. The research questions of this study are:

1. Are learners' question-generation performances correlated with academic achievement? If the hypothesized correlation is supported, will the relationship vary with the use of different identity revelation modes?
2. Are learners' peer-assessment performances correlated with academic achievement? If the hypothesized correlation is supported, will the relationship vary with the use of different identity revelation modes?
3. Are learners' question-generation performances correlated with their peer-assessment performances?

Method

Two hundred and fifty-three 5th graders from eight intact classes participated in the study for six consecutive weeks. Students were informed that the introduced activity was intended to support their science education. To ensure that participants possessed the fundamental skills for the activity, a training session about generating questions and peer-assessment, including hands-on activities, was held at the beginning of the study. For the duration of the study, students were directed to individually generate at least one question in accordance with the instructional content covered and assess at least two questions each week after attending three instructional sessions allocated for science.

A learning environment that allows students to contribute to and benefit from the process of constructing question items, as well as exchanging ideas with their peers about the composed questions, was used. The system allowed interacting parties to communicate back and forth electronically about a specific question item. Essentially, assessors gave evaluative feedback using an online assessment form (with assessment criteria and type-in open space) (see Fig. 1). Alternatively, question composers were able to provide an elaborated explanation for their assessors to further respond to (the very moment the feedback form was submitted) (see Fig. 2). In short, the system allowed both of the interacting parties to engage in argumentative dialogue.

Peer Assessment Form		
Assessor's identity revelation mode	<input checked="" type="radio"/> Pao-Da Huang (Real Name) <input type="radio"/> Thanks God it's Friday (Nickname) <input type="radio"/> Anonymity	
The overall quality of the question:	<input type="radio"/> Well above average <input type="radio"/> Slightly above average <input type="radio"/> Average <input type="radio"/> Slightly below average <input type="radio"/> Well below average	
Rating: How would you recommend this question:	<input type="radio"/> Highly recommend <input type="radio"/> Recommend with conservations <input type="radio"/> Not recommend	
Pros:	Cons:	Comments to the author: B I U abc x₁ x₂
Concise question-stem and options Important concepts Well-explained notes	Unclear question-stem Overly-complicated question-stem Excessively verbose options Multiple correct answers Elusive phrasing	
Submit		

Figure 1. Assessment Form for Assessors to Provide Feedback to Question-Authors.

Question-Author's Assessment of Peer's Comments/Feedback Form	
Author's identity revelation mode	<input checked="" type="radio"/> Shaw-Mei Wai (Real Name) <input type="radio"/> X-Ray (Nickname) <input type="radio"/> Anonymity
Your degree of agreement with assessor's comments:	<input type="radio"/> Strongly agree <input type="radio"/> Agree <input type="radio"/> No opinion <input type="radio"/> Disagree <input type="radio"/> Strongly disagree
The usefulness of assessor's comments:	<input type="radio"/> Very useful <input type="radio"/> Useful <input type="radio"/> No opinion <input type="radio"/> Useless <input type="radio"/> Complete useless
Comments to the assessor:	B I U abc x₁ x₂
Submit	

Figure 2. Assessment Form for Question-Authors to Communicate with Assessors.

Students' performances in question-generation and peer-assessment were defined as all questions students generated and assessed during the study. To assess students' performances in each respective activity, a set of criteria were developed. Specifically, in reference to the Torrance creativity index, King's question cognitive levels (1992) and questions generated by students, the following criteria were adopted: fluency, flexibility, elaboration, originality, cognitive level and importance. Each of the indexes was further operationally defined to ensure objective assessment. For example, the fluency index (score ranging from 0 to 3) was determined by correctness, clarity and conciseness of the composed question. A composite score was created for each generated question. As for peer-assessment, it was evaluated in terms of 3 discrete levels: (1) quality judgment, (2) quality judgment with identification of strong and weak areas, and (3) quality judgment

with explicit suggestions for further refinement of questions. Students' academic performance was based on the mid-term and final exams, which were centrally administrated at the participating school.

To test the moderation effects of the different identity revelation modes, four treatment conditions were devised. In the real-name condition, the student's full-name was retrieved automatically from the database and shown on screen when questions and feedback were viewed (See Figure 3). In the anonymous group, information about the question-author and assessor were not shown, and only the word "anonymous" was marked (See Figure 4). In the nickname group, the student's self-created identity was shown at the top of the generated questions and rendered comments (See Figure 5). Finally, to take into account the recent finding that participants exhibited statistically significant different preferences for different user identity revelation modes when rendering comments (Yu & Liu, 2009), a self-choice group was created. In the self-choice group, rather than being assigned a specific and fixed identity revelation mode, the identity mode used was chosen by the user (see Figs. 1 & 2 top portion: identity revelation mode choice space). Students were free to change to whichever mode each time they generate or assess a new item. Eight participating classes were randomly assigned to four different treatment conditions (i.e., two classes per group).

Questions Author:	WANG Hong-Min (real-name)
Question Type:	Multiple-choice
Question Stem:	The correct sequence for the appearance of the following creature on earth from the Paleozoic to the Cenozoic era is: (a) reptiles, (b) mammals, (c) algae
Alternative A:	abc
Alternative B:	cba
Alternative C:	cab
Alternative D:	bac
Answer:	3
Number of assessment:	3
Submit Time:	2008-05-07 16:45:31

Latest assessment record:		
Assessor: Shang-Chun Tsai (real-name)	Average rating: 90	Time: 2008-05-07
comment:		
The quality of the question is very good!		
What a coincidence, your questions was being assigned to me again.		

Figure 3. Real-Name Mode.

Questions Author:	Anonymous
Question Type:	Multiple-choice
Question Stem:	Which two parts evolve into the limbs of amphibian?
Alternative A:	a tail fin and a pectoral fin
Alternative B:	pectoral fin and pelvic fin
Alternative C:	pelvic fin and tail fin
Alternative D:	None of the above
Answer:	2
Number of assessment:	3
Submit Time:	2008-05-08 16:53:40

Latest assessment record:		
Assessor: Anonymous	Average rating: 90	Time: 2008-05-08
comment:		
There are some typos. Please carefully check and revise them accordingly.		

Figure 4. Anonymity Mode.

Questions Author:	sucking Jun (nickname)
Question Type:	Multiple-choice
Question Stem:	Which of the following pairing (biological creature and its existence on earth) is correct?
Alternative A:	Dinosaur - Paleozoic
Alternative B:	Gymnosperms - Mesozoic
Alternative C:	Trilobites - Cenozoic
Alternative D:	None of the above
Answer:	2
Number of assessment:	2
Submit Time:	2008-05-06 16:38:34

Latest assessment record:		
Assessor: Goldfish (nickname)	Average rating: 70	Time: 2008-05-06 16:54
comment:		
The composed question is too simple.		
It is almost verbatim from the textbook.		

Figure 5. Nickname Mode.

Results

The means and standard deviations of students' performances in generating questions, peer-assessment, and midterm and final exams are listed in Table 1. The potential effects of gender and prior experience in the introduced strategies have been reported (Topping, 2010). As can be seen in Table 1, gender and prior experience in either online question-generation or online peer-assessment were approximately the same across different treatment conditions. Data was analyzed using partial correlation while controlling for the influence of the mid-term exam scores.

Table 1: Descriptive statistics of the examined variable.

Treatment	Gender		Online QG Experience		Online PA Experience		QG Performance	QA Performance	Mid-term	Final exam
	Male	Female	Yes	No	Yes	No	Averaged Mean(SD)	Averaged Mean(SD)	Mean (SD)	Mean (SD)
Real-name (n=65)	33	32	25	40	26	39	5.02 (2.37)	4.50 (2.91)	85.48 (7.74)	91.32 (6.75)
Anonymous (n=63)	32	31	35	27	35	26	6.07 (2.48)	5.23 (3.20)	86.10 (8.83)	92.24 (7.08)
Nickname (n=64)	32	32	25	39	27	37	4.52 (2.25)	3.96 (2.74)	84.41 (9.6)	91.75 (8.29)
Self-choice (n=61)	28	33	29	32	27	34	4.54 (2.39)	4.15 (2.65)	87.39 (9.86)	90.57 (10.07)
Total (n=253)	125	128	114	139	115	138	5.04 (2.44)	4.46 (2.91)	85.82 (9.04)	91.48 (8.10)

Note: QG refers to Question-Generation; PA refers to Peer-Assessment

Relationships between Question-generation Performances and Academic Achievement

The partial correlation analysis results indicated a significant relationship between performances in generating questions and academic achievement ($r = 0.14$, $p = 0.02$). A hierarchical regression analysis was conducted to explore the moderation effect of different identity revelation modes on the strength of the relationship. The analysis showed that performances in question-generation and the different identity revelation modes interacted significantly with regard to their relationships with achievement ($\beta = 0.34$, $p = 0.03$). As shown in Table 2, the significant effects of the mid-term exam, the treatment and the interaction between question-generation performances and treatments on achievement were found. However, a direct effect of question-generation performances on achievement was not found. In short, identity revelation modes fully moderated the predictive power of question-generation performances on achievement. A series of follow-up regression tests, which uncovered the moderation phenomena, indicated that the question-generation performances of students in the self-choice mode significantly positively predicted achievement scores ($\beta_{\text{opt}} = 0.20$, $p = 0.04$) and that the predictive power of the self-choice group was the greatest of the four modes.

Table 2: Hierarchical regression model summary of the question-generation performance.

Variable	Model	Achievement (Dependent Variable)		
		Model 1	Model 2	Model 3
Controlled Variable				
Mid-term Exam		0.56**	0.56**	0.56**
Independent Variable				
Question-generation performance		0.12*	0.11*	-0.14
Treatment			-0.05	-0.28**
Interaction				
Treatment X Question-generation performance				0.34*
R-square		0.36	0.36	0.36
F		70.99	47.60	37.56
Sig		0.00	0.00	0.00
ΔR -square		0.36	0.00	0.01*
ΔF		70.99	0.88	5.09
Sig		0.00	0.35	0.03
Note:				
a. predictor:(constant), Mid-term Exam (controlled), Question-generation performance				
b. predictor:(constant), Mid-term Exam (controlled), Question-generation performance, Treatment				
c. predictor:(constant), Mid-term Exam (controlled), Question-generation performance, Treatment, interaction				

Relationships between Peer-assessment Performances and Academic Achievement

The partial correlation between performances in peer-assessment and achievement is statistically significant ($r = 0.15$, $p = 0.02$). However, the moderation effects of the different identity revelation modes on the strength of the relationship was not supported by the hierarchical regression analyses, as reported in Table 3 ($\beta = 0.1$, $p = 0.48$).

Table 3: Hierarchical regression model summary of peer-assessment performance.

Variable	Model	Achievement (Dependent Variable)		
		Model 1 ^a	Model 2 ^b	Model 3 ^c
Controlled Variable				
Mid-term Exam		0.56**	0.56**	0.56*
Independent Variable				
Peer-assessment performance		0.13**	0.12*	0.04
Treatment			-0.05	-0.11

Interaction			
Treatment X Peer-assessment performance			0.1
R-square	0.36	0.37	0.37
F	71.44	48.00	36.05
Sig	0.00	0.00	0.00
ΔR -square	0.36	0.00	0.001
ΔF	71.44	1.07	0.49
Sig	0.00	0.30	0.48
Note:			
a. predictor:(constant), Mid-term Exam(controlled), Peer-assessment performance			
b. predictor:(constant), Mid-term Exam(controlled), Peer-assessment performance, Treatment			
c. predictor:(constant), Mid-term Exam(controlled), Peer-assessment performance, Treatment, interaction			

Relationships between Question-generation and Peer-assessment

The correlation analysis result showed a significant relationship between performances in generating questions and peer assessment ($r = 0.66$, $p < 0.01$). Additionally, real-name, anonymous, nick-name, and self-choice all reached statistical significance with scores of $r = 0.69$ ($p < 0.01$), $r = 0.58$ ($p < 0.01$), $r = 0.61$ ($p < 0.01$) and $r = 0.71$ ($p < 0.01$), respectively. The relationship with the self-choice mode was the strongest.

Conclusion and Implications

In this study, relationships among question-generation, peer-assessment and academic performances were first explored, followed by an examination of how different identity revelation modes strengthen or weaken the relationships. The current study confirmed that the correlations among the three examined variables were significant. First, this study substantiated that students who performed better at question-generation tended to perform better in academic achievement. Additionally, the results, which supported the moderation effects of the identity revelation modes, evidenced that students in the self-choice mode who performed better in question-generation could be expected to perform better in academic achievement. Second, correlations between students' performances in peer-assessment and academic performance were also indicated for all four different identity revelation modes; however, the intensity of the relationships was similar in different modes. Third, the study confirmed that there was a high degree of correlation between question-generation and peer-assessment, and the intensity of the relationship was strongest in the self-choice mode.

These findings have important empirical significance as well as implications for online system developments and instructional implementation. First, the current study substantiated the relationships between performances in question-generation, peer-assessment and academic achievement on the applied content. In light of these findings, instructors interested in using the student question-generation approach for promoting academic performance are advised to address the issue of supporting question-generation and peer-assessment performances by, for instance, including deliberate training or building question-generation or peer-assessment scaffolds (Yu, 2009; Rosenshine, Meister, & Chapman, 1996), etc.

Furthermore, moving away from assessee's subjective perceptions of gains and attitudes towards peer-assessment, this study included objective measures (learning gains in academic contents) and evaluate its relationship with students' performances in question-generation and peer-assessment. In this study, students with better question-generation performances tended to offer peers better feedback and excelled better in academic assessment afterwards. A word of caution, however, is warranted. Since the activities of question-generation and peer-assessment were implemented simultaneously, no cause-and-effect but correlations can be inferred from the study.

Third, to the best of the researchers' knowledge, this is the first study examining and supporting the moderation effect of self-choice identity revelation mode on the relationship of question-generation, peer-assessment activity and academic achievement. It was found that the predictive power of question-generation performances for academic achievement in the self-choice identity group was the greatest, and that the self-choice identity mode strengthened the relationships between question-generation and peer-assessment performances. Taking psychological safety into consideration (van Gennip, Segers, & Tillema, 2010), the results of this study, and the fact that almost all online peer-assessment systems adopt a fixed user identity mode

(rather than dynamic mode that is adjustable by the user), designers of online systems should consider embedding functions that permit users to choose their own identity revelation mode.

References

- Abramovich, S., & Cho, E. K. (2006). Technology as a medium for elementary preteachers' problem-posing experience in Mathematics. *Journal of Computers in Mathematics and Science Teaching*, 25(4), 309-323.
- Barlow, A., & Cates, J. M. (2006). The impact of problem posing on elementary teachers' beliefs about mathematics and mathematics teaching. *School Science and Mathematics*, 106(2), 64-73.
- Brindley, C., & Scofield, S. (1998). Peer assessment in undergraduate programmers. *Teaching in Higher Education*, 3(1), 79-90.
- Brown, S.I., & Walter, M. I. (2005). *The Art of Problem Posing* (3rd ed). New Jersey: Lawrence Erlbaum Associates.
- Dori, Y. J., & Herscovitz, O. (1999) Question-posing capability as an alternative evaluation method: analysis of an environmental case study. *Journal of Research in Science Teaching*, 36(4), 411-30.
- English, L. D. (1998). Children's problem posing within formal and in formal context. *Journal for Research in Mathematics Education*, 29(1), 83-106.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70 (3), 287-322.
- Gatfield, T. (1999). Examining student satisfaction with group projects and peer assessment. *Assessment & Evaluation in Higher Education*, 24 (4), 365-377.
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research & Development*, 20 (1), 53-70.
- King, A. (1992). Facilitating elaborative learning through guided student-generated questioning." *Educational Psychologist* 27, 111-126.
- Leung, S. S., & Wu, R. X. (1999). Problem posing with middle grades mathematics: Two real classroom examples. *Mathematics teaching in the middle school*. Reston, VA: National Council of Teachers of Mathematics.
- Moses, B.M., Bjork, E., & Goldenberg, E.P. (1993). Beyond problem solving: problem posing. In S.I. Brown, and M.I. Walter. *Problem posing: reflections and applications* (pp. 178-188). NJ, Hillsdale: Lawrence Erlbaum Associates.
- Purchase, H. C. (2000). Learning about interface design through peer assessment. *Assessment & Evaluation in Higher Education*, 25(4), 341-352.
- Rosenshine, B., Meister, C., & Chapman, S. (1996). Teaching students to generate questions: a review of intervention studies. *Review of Educational Research*, 66, 181-221.
- Silver, E. A. (1994). On mathematical problem posing. *For the Learning of Mathematics*, 14(1), 19-28.
- Topping, K. J. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, 20(4), 339-343.
- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249-276.
- Tsai, C. C., Liu, E. Z. F., Lin, S. S. J., & Yuan, S. M. (2001). A networked peer assessment system based on a vee heuristic. *Innovations in Education and Teaching International*, 38, 220-30.
- Tsai, C. C., Lin, S. S. J. & Yuan, S. M. (2002). Developing science activities through a networked peer assessment system. *Computers & Education*, 38, 241-252.
- Van Gennip, N. A. E., Segers, M., & Tillema, H. H. (2010). Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. *Learning and Instruction*, 20 (4), 280-290.
- Venables, A., & Summit, R. (2003). Enhancing scientific essay writing using peer assessment. *Innovations in Education and Teaching International*, 40, 281-290.
- Vreman-de Olde, C., & de Jong, T. (2004). Student-generated assignments about electrical circuits in a computer simulation. *International Journal of Science Education* 26 (7), 859-873.
- Wen, M. L., & Tsai, C. C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education*, 51, 27-44.
- Wen, M. L., Tsai, C. C., & Chang, C. Y. (2006). Attitudes toward peer assessment: a comparison of the perspectives of pre-service and inservice teachers. *Innovations in Education and Teaching International*, 43, 83-92.
- Wu, C. P., & Yu, F. Y. (2010). An innovation diffusion approach to online student question-generation and its effects on the relationship of perceived task value and learning approach. Paper accepted for presentation at the 18th International Conference on Computers in Education (ICCE 2010), November 29 to December 3, Putrajaya, Malaysia.

- Wu, C. P., & Yu, F. Y. (2009). Changing students' perceived value and use of learning approaches for online student-generated questions via an integrative model. *Workshop Proceedings of the International Conference on Computers in Education*, p. 30-34. November 30 to December 4. Hong Kong Institute of Education, Hong Kong.
- Yu, F. Y. (2009). Scaffolding student-generated questions: Design and development of a customizable online learning system. *Computers in Human Behavior*, 25 (5), 1129-1138.
- Yu, F. Y., & Liu, Y. H. (2009). Creating a psychologically safe online space for a student-generated questions learning activity via different identity revelation modes. *British Journal of Educational Technology*, 40 (6), 1109-1123.
- Yu, F. Y., & Liu, Y. H. (2005) Potential values of incorporating multiple-choice question-construction for physics experimentation instruction. *International Journal of Science Education*, 27 (11), 1319-1335.
- Yu, F. Y., Liu, Y. H. & Chan, T. W. (2005). A Web-based learning system for question-posing and peer assessment. *Innovations in Education and Teaching International*, 42 (4), 337-348.

Acknowledgments

This paper was funded by research grants from the National Science Council, Taiwan, ROC (NSC 96-2520-S-006-002-MY3; NSC 99-2511-S-006 -015 -MY3). The author would like to thank research assistants, Meiju Chen and Knem Chen for their assistance during the data collection process.