Original article

doi: 10.1111/jcal.12125

Journal of Computer Assisted Learning

From integrative to game-based integrative peer response: high ability versus low ability

J.H. Wang,* S.Y. Chen,* B. Chang† & T.W. Chan*

*Graduate Institute of Networking Learning Technology, National Central University, Taoyuan, Taiwan
†Graduate Institute of Learning and Instruction, National Central University, Taoyuan, Taiwan

Abstract Peer response is useful to improve student writing. However, traditional peer response takes a single mode, which has some problems, such as effort for preparation of documents or ambiguous feedback. To address these problems, this study presents two peer response approaches, that is, an integrative approach and a game-based integrative approach. Additionally, whether students with these two peer response approaches and those with a non-peer response approach performed differently is examined in this study, where students' ability levels were also considered. The findings suggest that students with the peer response generally demonstrated better writing performance than those without the peer response. Furthermore, students with the game-based integrative approach showed better writing quality than those with the integrative approach. Moreover, the former was more helpful in giving direct feedback on surface features and criticism than the latter. However, the usefulness of peer response approaches was associated with ability levels. More specifically, the integrative approach was beneficial for the high-ability students to improve on their written expression in the aspect of length of composition, while the game-based integrative approach was advantageous to enhance the low-ability students' writing quality in the aspect of clear paragraph.

Keywords game-based learning, interactive learning environments, peer response.

Introduction

Peer response (Elbow, 1973; DiPardo & Freedman, 1988), also known as peer review, refers to a collaborative activity in which learners work together to improve the quality of their works by providing comments for each other. As highlighted by Hwang, Hung and Chen (2014), peer response could facilitate students to enhance their in-depth thinking, problem-solving skills and creativity. In other words, peer response possesses many benefits, which generally can be classified into four aspects: social, cognitive, affective and linguistic (Min, 2006).

Accepted: 26 September 2015
Correspondence: Sherry Y. Chen, Graduate Institute of Network Learning Technology, National Central University, No. 300, Zhongda Rd, Zhongli District, Taoyuan City 32001, Taiwan. Email: sherry@CL.ncu.edu.tw

- Social aspect: negotiations used in the process of peer response could foster the skills of students' communication and collaboration (Mendonça & Johnson, 1994).
- Cognitive aspect: students' critical and analytical skills could be improved (Stoddard & MacArthur, 1993), and they could develop the sense of audience (Lockhart & Ng, 1993).
- Affective aspect: peer response could reduce students' anxiety and increase their confidence (Leki, 1990) and make them have a sense of ownership of their works (Tsui & Ng, 2000).
- Linguistic aspect: students could create new ways to express their thoughts (Lockhart & Ng, 1993), so their writing quality could be enhanced (Cho & Schunn, 2007).

Because of such benefits, peer response has been widely applied to support student learning. In particular, peer response is useful to facilitate students to develop their writing capabilities. For example, Sims (2001) found that students' writing fluency could be enhanced by peer response, and subsequently, Boscolo and Ascorti (2004) found that peer response could cultivate students' writing abilities. The aforementioned studies demonstrated the effectiveness of peer response. However, most of these studies implemented a single-mode peer response, that is, a face-to-face (FTF) mode or computer-mediated communication (CMC) mode, each of which has different advantages and disadvantages (Table 1), so they can be complementary to each other. Therefore, we develop an integrative peer response (IPR) approach, which combines the advantages of the FTF and CMC modes. However, the IPR is still implemented in a traditional educational context, where students may have low motivation (Tüzün, Yılmaz-Soylu, Karakuş, İnal & Kızılkaya, 2009). Therefore, there is a need to use a mechanism that can increase students' motivation.

Among various mechanisms, digital games have transformed people's learning patterns and make them have enjoyable experience (Prensky, 2001; Marsh, 2011). This may be the reason why game-based learning (GBL) could improve students' learning achievement (Pivec, 2007) and enhance their learning motivation (Gee, 2003; Papastergiou, 2009). Regarding learning achievement, Chang, Wu, Weng and Sung (2012) and Maratou, Chatzidaki and Xenos (in press) indicated that GBL could enhance students' learning achievement. The former was demonstrated in the context of problem solving, while the latter was undertaken in the context of software project management. Moreover, Ku, Chen, Wu, Lao and Chan (2014) found that GBL could improve students' performance in mathematics. Regarding learning motivation, Sabitzer (2013) employed GBL to help primary and secondary students learn core concepts of informatics. The results revealed that students' motivation was enhanced. Subsequently, Sung, Hwang and Yen (2015) used GBL to help the learning of health education, and their results indicated that students' learning motivation was improved.

As shown in these studies, GBL had positive effects, so it could be applied to address the problem of the IPR. Therefore, we also propose a game-based IPR (G-IPR) approach by incorporating GBL into the aforementioned IPR approach. However, game elements presented in the G-IPR may distract student learning. On the other hand, traditional teacher-centred instruction, belonging to a non-peer response (NPR) approach, provides feedback via a single channel, where students may be less distracted but they may receive insufficient feedback (Cho & Schunn, 2007). In brief, the NPR, IPR and G-IPR have different advantages and disadvantages, so they may be valued by different learners.

In particular, learners are diverse, so they may prefer different peer response approaches. In other words, it is necessary to consider individual differences. Among various individual differences, learners' ability levels greatly influence students' perceptions (Cheng, Lam & Chan, 2008), which, in turn, also affect their learning achievement. For instance, low-ability students (LAS) may enjoy interacting with more capable peers because the former can obtain assistance from the latter (Ghaith, 2001). Conversely, high-ability students (HAS) may be

Table 1. Summary of Advantages and Disadvantages of FTF Mode and CMC Mode

	FTF mode	CMC mode
Advantages	Direct discussion of meanings with more visual cues (Tiene, 2000) Low cost and complexity of implementation (Peckham, 1996)	Low cost of document and feedback delivery and scale extension of peer response (Peckham, 1996) Less psychological pressure (Ciftci & Kocoglu, 2012) Equality of participation (Palloff & Pratt, 2007) Ease of access of peers' feedback (Yang, 2011)
Disadvantages	Effort for preparation of documents (DiPardo & Freedman, 1988) Psychological pressure (Ho & Savignon, 2007)	Ambiguous feedback (Ho & Savignon, 2007; Lin, Liu, & Yuan, 2001) Lack of visual cues (Liang, 2010)

FTF, face-to-face; CMC, computer-mediated communication.

reluctant to learn with less capable peers because the former may need to devote additional time to assisting the latter so that the learning progress of the HAS may be delayed (Fuchs, Fuchs, Mathes & Simmons, 1997). Thus, such ability levels may affect how students react to the NPR, IPR and G-IPR approaches, and there is a need to examine the impacts of students' abilities on their reactions to these three approaches.

To this end, this study aims to examine the effects of these three peer response approaches from a perspective of ability levels. To correspond to this aim, three research questions are examined in this study:

(1) whether students with different peer response approaches performed differently;
(2) whether HAS with different peer response approaches performed differently; and
(3) whether LAS with different peer response approaches performed differently.

The answers to these research questions can contribute to developing a deep understanding of how to implement peer response that can accommodate students' ability differences. By doing so, both HAS and LAS can benefit from peer response.

Methodology design

To correspond to the research questions, an empirical study was conducted. The details are described in this section, including the implementation of the IPR and G-IPR, participants, an experimental procedure and data analysis.

The implementation of IPR and G-IPR

Two peer response approaches, that is, the IPR and G-IPR, were implemented. Furthermore, writing was chosen as the subject content for the employment of these two approaches because writing has become a critical competence for students' success at schools and in careers (National Commission on Writing, 2004).

Integrative peer response

The IPR, a forum-based environment, was implemented by making the best use of the advantages of the FTF and CMC modes. Furthermore, solutions were proposed to overcome the problems of each mode. The problem of the FTF mode is that additional effort is required to prepare hard copies of student writing. To remove the necessity of such effort, the IPR provided a role-based access mechanism (Figure 1), with which students not only could play as writers to compose and revise their own drafts in their own online writing workspace but also could act as readers (i.e., responders) to access their peers' writing works and give feedback to them. The other problem is psychological stress on students who are afraid of providing FTF feedback. To alleviate such fear, the IPR provided an online responding mechanism (Figure 2), with which students could give and edit their comments, instead of delivering FTF comments. The problem of the CMC mode is that the meanings of online

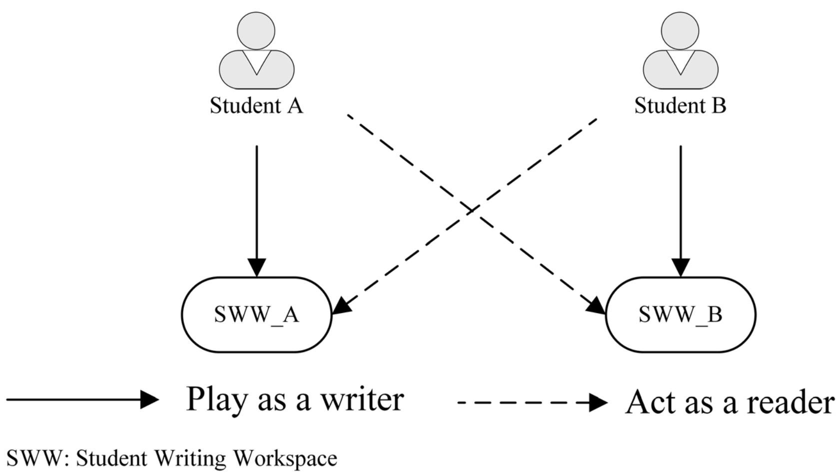

Figure 1 The Role-based Access Mechanism of the Integrative Peer Response

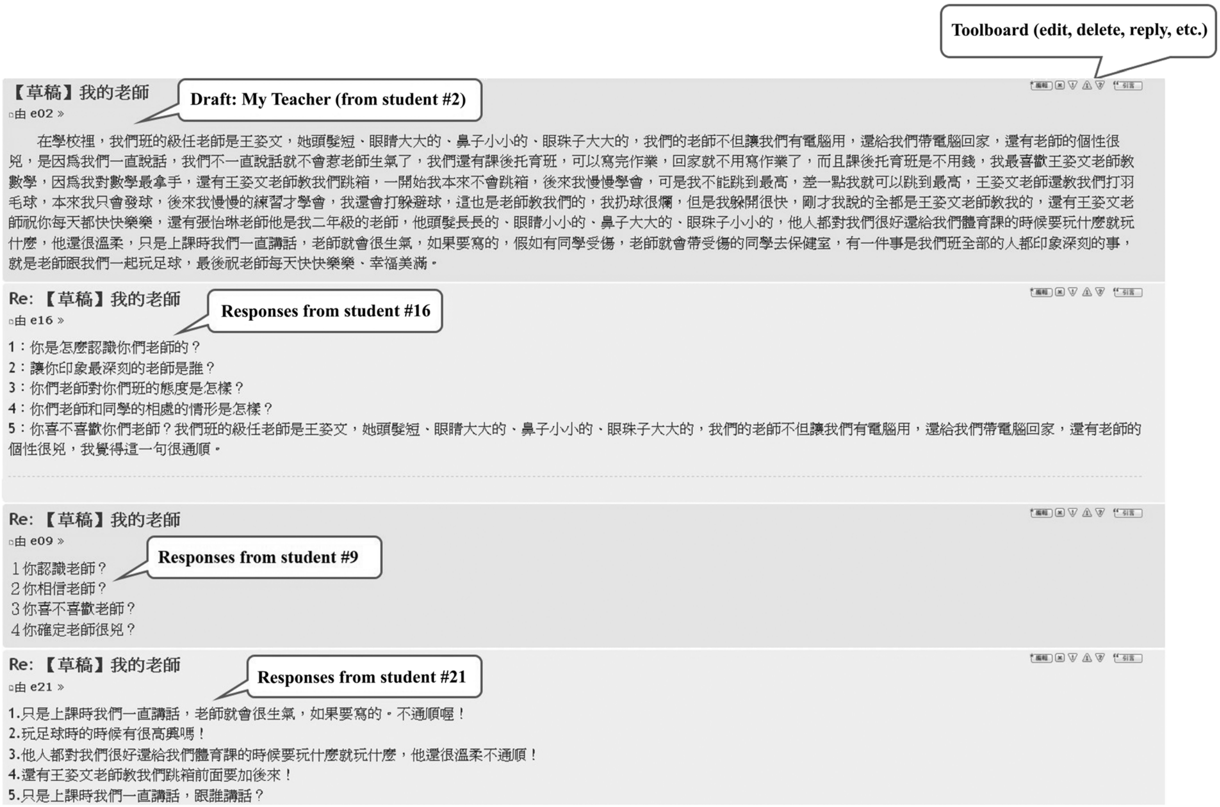

Figure 2 The Online Responding Mechanism of the Integrative Peer Response

feedback may be too vague to be understood so that some misunderstandings may be caused between writers and readers. To address this problem, the IPR provided peers with FTF discussion, clarification and negotiation to enhance their communication.

Game-based integrative peer response

As described in the Introduction, digital games can enhance students' learning motivation. Accordingly, the G-IPR was developed by integrating a digital game and the IPR. Among various types of digital games, a role-playing game was implemented in the G-IPR. More specifically, students played as the head of a publisher and manage the publisher, which included a lobby, a galley and five departments (Figure 3):

- *Lobby*: to work as a main entry for the gallery and departments of the G-IPR;
- *Gallery*: to work as a sharing platform and provide students with more opportunities to exhibit their published works;
- *Edit department*: to draft a writing composition;
- *Review department*: to deliver feedback for peers' works and to do the revision based on comments received from peers;
- *Publish department*: to publish students' revised works in a digital magazine;
- *Marketing department*: to enhance the visibility of the digital magazine; and
- *Head office*: to manage the status of every department and corresponding rewards.

Among these five departments, the review department was deeply involved in peer response activities, where a structured checklist was provided to facilitate readers to deliver comments and the usefulness of the comments was assessed by writers. Additionally, an entire writing task was decomposed into several small tasks, each of which had a clear sub-goal and was achieved by different departments. When students completed a small task, they also achieved its sub-goal and could move to pursue the next sub-goal. Accordingly, the writing task could be completed through the accumulation of these sub-goals. In brief, the G-IPR provides a 'learning by doing'

Figure 3 The Interface of the Game-based Integrative Peer Response

environment, where the structured checklist was applied to deliver feedback.

Further, to provide a 'learning by doing' environment, students were also engaged in a 'learning by playing' scenario, where game elements were used to enhance their engagement. Among various game elements, virtual currency, leader boards and trophies (Figure 4) were considered in this study because they are the most commonly employed to support GBL (Hamari, Koivisto, & Sarsa, 2014). In addition, they could serve different purposes (Seaborn & Fels, 2015):

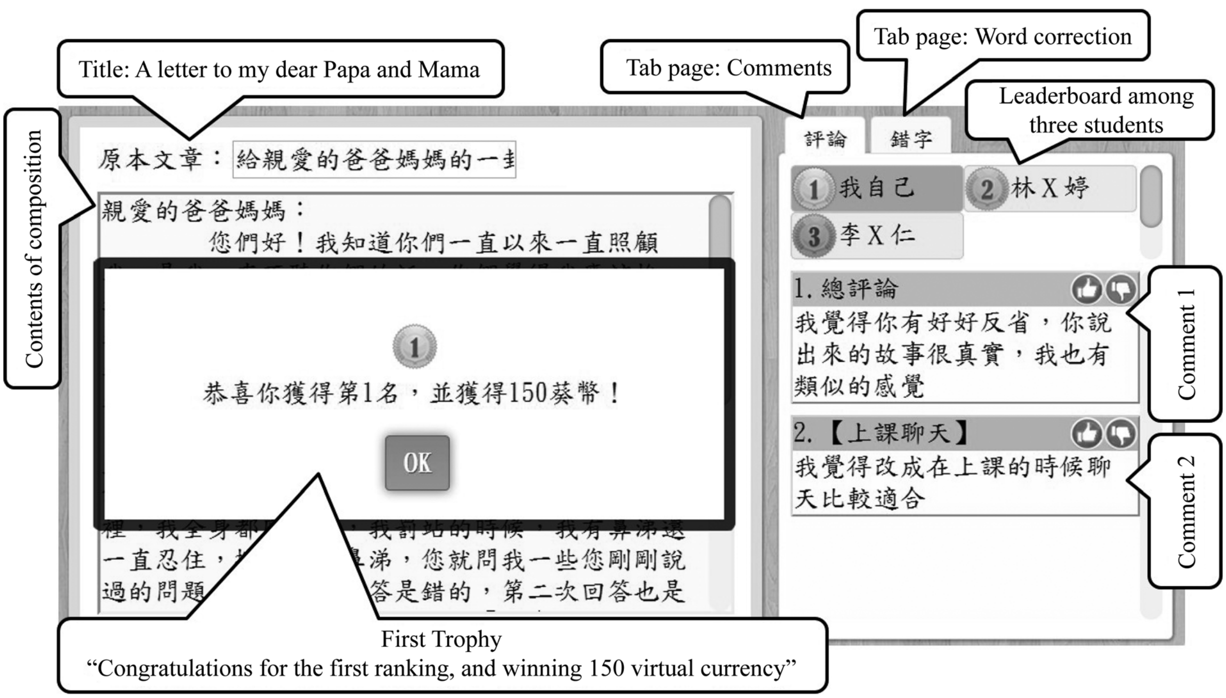

Figure 4 The Snapshot of the Game Elements

- Virtual currency: rewards for students' effort in the accomplishment of tasks that they have performed and resources for subsequent promotion activities
- Leader board: directs students' attention to good comments delivered by peers so that they could learn how to produce useful comments
- Trophy: social recognition of achievement to motivate students to produce useful feedback

In summary, the design rationale of the G-IPR lies within two aspects: 'learning by doing' and 'learning by playing'. The former was achieved by the structured checklist, while the latter was realized by various game elements. The ultimate goal of these aspects was to enhance the effectiveness of peer response so that their writing performance could be improved.

Participants

Seventy-five third-grade elementary students participated in this study. Such a sample size was in line with previous studies in digital learning, for example, Bolzer, Strijbos and Fischer (2015) and Bagley and Shaffer (in press). Our participants were divided into three groups, that is, the experimental group I (EGI, $N = 27$), the experimental group II (EGII, $N = 21$) and the control group (CG, $N = 27$). The IPR and G-IPR were implemented in the EGI and EGII, respectively, while the NPR approach was applied in the CG. Apart from such a difference, these three groups were equally treated with the same curriculum, writing assignments and instruction.

Experimental procedures

Figure 5 illustrates the experimental procedures that each group followed. As shown in this figure, all of the groups initially needed to take a pre-test to identify their prior writing abilities and were finally assessed by taking a post-test. Regardless of the pre-test and post-test, participants were given a topic to make a narrative composition within 80 min. However, these three groups undertook different pedagogical activities between the pre-test and post-test. Regarding the CG, students were provided with a conventional teacher-centred pedagogy, where they were initially given a brief introduction for their writing tasks. Then, students completed their compositions individually with paper and pencils and handed in their works. Subsequently, the teacher would return their works with some comments, and they were expected to improve their next works with these comments.

Regarding the EGI and EGII, participants were evenly allocated into small peer response groups of four or five students. Subsequently, they were trained on how to undertake peer response. The training covered the following:

- Capacities of the readers and writers. Instructions were given to them based on two guidelines: (a) the interaction between readers and writers proposed by Elbow (1973) and (b) the guidance for peer response proposed by Hansen and Liu (2005).
- Usage of the IPR and G-IPR. Each participant of the EGI and EGII was trained on how to complete writing and reviewing tasks with the IPR and G-IPR, respectively.

After the training sessions were completed, students moved to start their writing tasks. More specifically, students wrote drafts individually in their own writing workspace. Then, each peer response group started to read an assigned group mate's draft and posted their feedback, followed by an FTF discussion to clarify and negotiate the meanings of written texts. Thereafter, students revised their own drafts based on the feedback from their peers.

The aforesaid activities happened in both the EGI and EGII. Nevertheless, the EGII needed to undertake some game activities, for example, to publish their revised works in the digital magazine and to conduct promotion activities to attract peers to access and order their published works. Furthermore, how writers perceived feedback received from readers was employed to decide what the readers could obtain, in terms of virtual currency, leader boards and trophies. Finally, they were required to express how they perceived such game elements by filling out a brief survey, which used a 5-point Likert scale ranging from 1 (*highly valueless*) to 5 (*highly valuable*).

Data analysis

The independent variables of this study are the peer response approaches and students' ability levels, while the dependent variables are writing performance and feedback performance. Writing performance, which

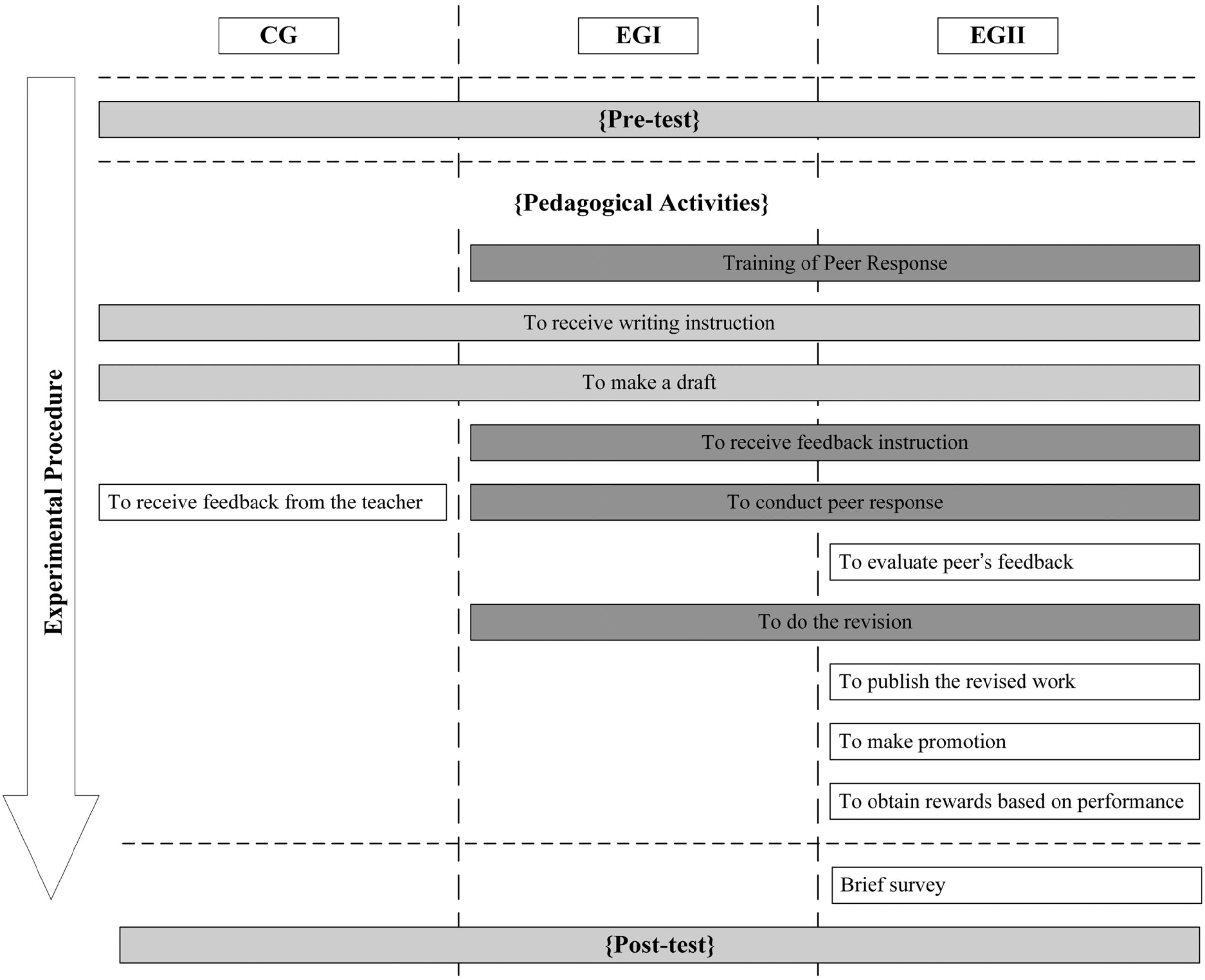

Figure 5 The Experimental Procedure of the Study. CG, Control Group; EGI, Experimental Group I; EGII, Experimental Group II

was measured based on the pre-test and post-test scores, included two aspects: writing quality and written expression. An assessment mechanism proposed by Yang, Ko and Chung (2005) was adopted to assess students' writing quality because it was designed for elementary students, to which our participants belonged. This assessment covers five items: (1) *elegant words*, (2) *clear paragraph*, (3) *coherence*, (4) *title consistence* and (5) *new and original ideas*. A 5-point Likert scale was used for each item. Thus, the total score for a composition was between the lowest score (5 points) and the highest score (25 points). Two raters were recruited to independently evaluate the participants' writing quality, so each student's final score was defined based on the mean of scores by the raters, of which the inter-rater reliability was found to be $\kappa = 0.728$ ($p < 0.001$). In other words, a substantial level of agreement between the raters was reached. On the other hand, the length of composition and the richness of vocabulary used in previous studies (e.g., Chanquoy, 2001) were also adopted to assess written expression, and they were measured based on the number of words and the number of vocabulary items, respectively.

Feedback performance was examined based on the number of various types of feedback provided by the EGI and EGII. In total, there were six types of feedback based on Cho, Schunn and Charney (2006), that is,

- directive feedback on surface features: suggestions for surface adjustments, for example, changes in punctuations;
- directive feedback on content features: suggestions for the changes of content, for example, the development of ideas or organization of the content;

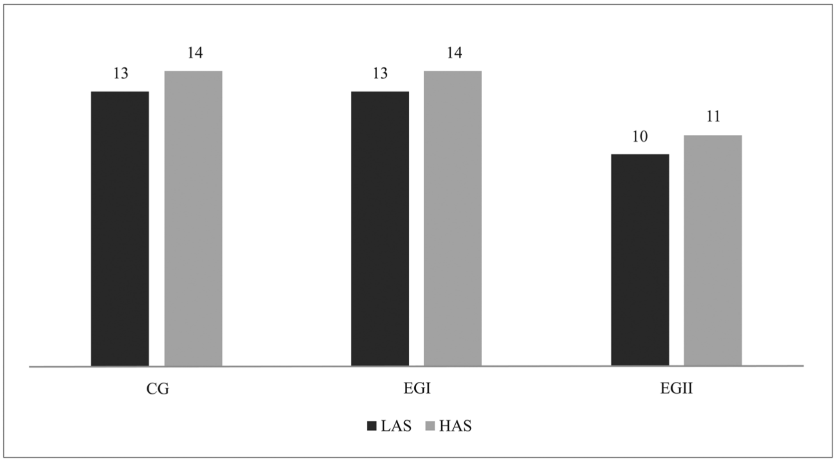

Figure 6 The Distribution of High-ability and Low-ability Students (HAS and LAS, Respectively). CG, Control Group; EGI, Experimental Group I; EGII, Experimental Group II

- praise: appreciation or encouragement about some portions of an article;
- criticism: nondirective or negative comments to some particular portions of an article;
- clarification and discussion: description of personal experience, which may be similar to or different from a writer's experience; and
- off task: comments unrelated to a writer's article.

Thereafter, the coding task was conducted by two researchers independently. During the process of doing the coding task, a feedback unit was defined as an integrated and meaningful message focusing on a single problem (Artemeva & Logie, 2002). According to this definition, all responses were initially classified into feedback units, each of which was then assigned to one of the aforementioned six types. A discussion would be undertaken to sort out disagreement on the coding task so that a consensus could be reached.

The aforementioned data were majorly analysed with two-way analysis of variance (ANOVA), where the level of significance was set at $p < 0.05$ for all comparisons. Furthermore, the effect size of η^2 was measured based on Cohen's criteria. Cohen (1988) has indicated that $\eta^2 \geq 0.059$ represents a moderate effect size and $\eta^2 \geq 0.138$ represents a large effect size; that is, the effect sizes of the two-way ANOVA results presented in this study are large. These aforementioned analyses were undertaken using SPSS 16.0 (SPSS Inc., Chicago, IL, USA).

Results

The participants of each group were classified into the HAS and LAS based on the mean of the pre-test scores. Figure 6 illustrates the distribution of the sample. On the other hand, two outcomes are examined in this study, that is, writing performance (section on Writing Performance) and feedback performance (section on Feedback Performance).

Writing performance

The results from the pre-test scores showed that no significant difference [$F(2, 72) = 1.415$, $p = 0.250 > 0.05$] existed among the CG ($M = 12.76$, SD $= 1.88$), EGI ($M = 12.39$, SD $= 2.41$) and EGII ($M = 13.48$, SD $= 2.42$). This implied that the three groups had a similar level of prior writing ability. The post-test scores, which were applied to assess students' writing quality and written expression, were analysed with the two-way ANOVA. Regarding the writing quality, the peer response approaches and the ability levels did not have significant interaction effects, apart from the aspect of clear paragraph (Table 2). Therefore, the main effects on the remaining dependent variables were analysed, and they revealed that the peer response approaches significantly affected students' performance in elegant words ($F = 8.15$, $p = 0.001 < 0.01$, $\eta^2 = 0.184$), coherence ($F = 27.33$, $p = 0.000 < 0.001$, $\eta^2 = 0.430$), title consistence ($F = 117.60$, $p = 0.000 < 0.001$, $\eta^2 = 0.761$) and new and original ideas ($F = 14.20$, $p = 0.000 < 0.001$, $\eta^2 = 0.283$).

Table 2. Results of Two-way ANOVA on Different Aspects of Students' Writing Performance

Source	SS	*df*	MS	*F*	*p*	η^2	*Post hoc* tests
Writing quality							
Elegant words							
PR approaches (A)	3.23	2	1.61	8.15**	0.001	0.184	EGII > CG EGI > CG
Ability levels (B)	0.41	1	0.41	2.05	0.157	0.023	
A × B	0.15	2	0.07	0.37	0.691	0.008	
Error	13.65	69	0.20				
Clear paragraph							
PR approaches (A)	7.15	2	3.57	10.43***	0.000	0.211	
Ability levels (B)	0.21	1	0.21	0.62	0.433	0.006	
A × B	3.05	2	1.52	4.45*	0.015	0.090	EGII_LAS > CG_LAS EGII_LAS > EGI_LAS
Error	23.64	69	0.34				
Coherence							
PR approaches (A)	14.85	2	7.42	27.33***	0.000	0.430	EGII > CG EGII > EGI
Ability levels (B)	0.38	1	0.38	1.38	0.244	0.011	
A × B	0.46	2	0.23	0.85	0.432	0.013	
Error	18.74	69	0.27				
Title consistence							
PR approaches (A)	53.04	2	26.52	117.60***	0.000	0.761	EGII > CG EGII > EGI
Ability levels (B)	0.61	1	0.61	2.70	0.105	0.009	
A × B	0.57	2	0.29	1.27	0.289	0.008	
Error	15.56	69	0.23				
New and original ideas							
PR approaches (A)	9.45	2	4.73	14.20***	0.000	0.283	EGII > CG EGII > EGI
Ability levels (B)	0.24	1	0.24	0.72	0.398	0.007	
A × B	0.75	2	0.37	1.12	0.333	0.022	
Error	22.97	69	0.33				
Written expression							
Length of composition							
PR approaches (A)	264 237.47	2	132 118.74	14.99***	0.000	0.267	
Ability levels (B)	43 999.02	1	43 999.02	4.99*	0.029	0.044	
A × B	56 683.18	2	28 341.59	3.22*	0.046	0.057	EGI_HAS > CG_HAS EGI_HAS > EGII_HAS EGI_LAS > CG_LAS
Error	608 032.63	69	8 812.07				
Richness of vocabulary							
PR approaches (A)	5 720.90	2	2 860.45	11.52***	0.000	0.228	EGI > CG EGII > CG
Ability levels (B)	823.63	1	823.63	3.32	0.073	0.033	
A × B	1 103.23	2	551.61	2.22	0.116	0.044	
Error	17 132.22	69	248.29				

ANOVA, analysis of variance; PR, peer response; SS, sum of squares; *df*, degree of freedom; MS, sum of mean squares; EGI, experimental group I; EGII, experimental group II; CG, control group; HAS, high-ability students; LAS, low-ability students. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$.

The results from the *post hoc* comparison show three interesting trends:

- Both the EGII ($M = 2.71$, SD $= 0.46$) and EGI ($M = 2.63$, SD $= 0.49$) performed significantly better than the CG ($M = 2.24$, SD $= 0.38$), in terms of elegant words.
- The EGII significantly outperformed the EGI, in terms of coherence (EGII: $M = 3.14$, SD $= 0.79$; EGI: $M = 2.26$, SD $= 0.40$), title consistence (EGII: $M = 4.57$, SD $= 0.60$; EGI: $M = 2.81$, SD $= 0.37$) and new and original ideas (EGII: $M = 2.90$, SD $= 0.70$; EGI: $M = 2.41$, SD $= 0.57$).
- In the three aspects mentioned before, the EGII was also significantly superior to the CG (coherence: $M = 2.07$, SD $= 0.33$; title consistence: $M = 2.61$, SD $= 0.49$; and new and original ideas: $M = 2.02$, SD $= 0.47$).

As shown in Table 2, peer response approaches and ability levels have a significant interaction effect on the aspect of clear paragraph ($F = 4.45$, $p = 0.015 < 0.05$, $\eta^2 = 0.090$). Therefore, the simple main effect was further conducted to identify the interaction effect of the peer response approaches and ability levels on the performance of clear paragraph. The results of the simple main effect revealed that the peer response approaches did not significantly affect the performance of clear paragraph for the HAS ($F = 1.49$, $p = 0.240 > 0.05$), while significant influences were found for the LAS ($F = 11.30$, $p = 0.000 < 0.001$). The results from the *post hoc* test indicated that the LAS within the EGII ($M = 3.40$, SD $= 0.84$) significantly outperformed those within either the EGI ($M = 2.42$, SD $= 0.49$) or the CG ($M = 2.19$, SD $= 0.56$) in the aspect of clear paragraph (Table 3).

In summary, the EGII performed better than the CG in the aspects of elegant words, coherence, title consistence and new and original ideas, and the EGII performed better than the EGI in the aspects of coherence, title consistence and new and original ideas. The EGI, EGII and CG used the IPR, G-IPR and NPR, respectively. Accordingly, these findings suggested that the G-IPR was more useful in enhancing students' writing quality in the aspects of elegant words, coherence, title consistence and new and original ideas. To further analyse which game elements could make such enhancement, Pearson's correlations were applied to identify relationships between students' responses to the brief survey and their performance in the aforementioned aspects. The results indicated that a positive relationship ($r = 0.491$, $p = 0.024 < 0.05$) existed between their performance of elegant words and reactions to the trophies and a positive relationship ($r = 0.482$, $p = 0.027 < 0.05$) existed between their performance of new and original ideas and reactions to the virtual currency. These findings revealed that the trophies and virtual currency were useful game elements to foster student writing.

Regarding the written expression, the peer response approaches and the ability levels did not have a significant interaction effect on the aspect of richness of vocabulary (Table 2), while they had a significant interaction effect on the aspect of length of composition ($F = 3.22$, $p = 0.046 < 0.05$, $\eta^2 = 0.057$). Thus, the main effect was analysed for the richness of vocabulary, whereas the simple main effect was analysed for the length of composition. The results of the main effect revealed that the peer response approaches significantly affected students' performance in the richness of vocabulary ($F = 11.52$, $p = 0.000 < 0.001$, $\eta^2 = 0.228$). The results from the *post hoc* comparison indicated that both the EGI ($M = 66.04$, SD $= 21.16$) and EGII ($M = 61.00$, SD $= 15.05$) performed significantly better than the CG ($M = 46.04$, SD $= 10.95$) in this aspect. On the other hand, the results of the simple main effect revealed that the peer response approaches significantly affected the HAS's and LAS's length of composition (Table 4). The results from the *post hoc* test suggested that the HAS within the EGI ($M = 397.07$, SD $= 131.29$) not only significantly outperformed those within the CG

Table 3. The Simple Main Effect of the Peer Response Approaches on Clear Paragraph for the LAS

Item	SS	*df*	MS	*F*	*p*	*Post hoc* tests
Clear paragraph	8.96	2	4.48	11.30***	0.000	EGII_LAS > CG_LAS EGII_LAS > EGI_LAS

LAS, low-ability students; SS, sum of squares; *df*, degree of freedom; MS, sum of mean squares; EGI, experimental group I; EGII, experimental group II; CG, control group.
***$p < 0.001$.

Table 4. The Simple Main Effect of the Peer Response Approaches on Length of Composition for the HAS and LAS

Item	Ability levels	SS	*df*	MS	*F*	*p*	*Post hoc* tests
Length of composition	HAS	241 929.86	2	120 964.93	13.11***	0.000	EGI_HAS > CG_HAS EGI_HAS > EGII_HAS
	LAS	84 931.96	2	42 465.98	5.08*	0.012	EGI_LAS > CG_LAS

HAS, high-ability students; LAS, low-ability students; EGI, experimental group I; EGII, experimental group II; SS, sum of squares; *df*, degree of freedom; MS, sum of mean squares.
$^{*}p < 0.05$; $^{***}p < 0.001$.

(M = 222.93, SD = 70.58) but also significantly outperformed those within the EGII (M = 248.73, SD = 65.77) in the length of composition. However, the LAS within the EGI (M = 279.15, SD = 111.67) only performed significantly better than those within the CG (M = 174.00, SD = 62.82) in this aspect.

The aforesaid findings suggested that no significant differences existed between the IPR and the G-IPR in the aspect of the richness of vocabulary although both of them were more helpful than the NPR. On the other hand, peer response approaches had significant effects on the length of composition. More specifically, all students with the IPR significantly outperformed those with the NPR, and the HAS with the IPR also performed significantly better than those with the G-IPR. These findings implied the IPR was relatively beneficial for the HAS in this aspect.

Feedback performance

The peer response approaches and the ability levels did not have interaction effects on directive feedback on surface features, criticism, praise, clarification and discussion, and off task. As shown in Table 5, the results of the main effect analysis, however, revealed that the peer response approaches had significant impacts on the aspects of directive feedback on surface features (F = 11.28, p = 0.002 < 0.01, η^2 = 0.195) and criticism (F = 10.96, p = 0.002 < 0.01, η^2 = 0.189). More specifically, EGII (directive feedback on surface features: M = 5.14, SD = 4.25; criticism: M = 2.05, SD = 1.18) outperformed EGI (directive feedback on surface features: M = 2.15, SD = 1.60; criticism: M = 1.13, SD = 0.83) in these two aspects. On the other hand, the peer response approaches and the ability levels had an interaction effect on directive feedback on

Table 5. Results of Two-way ANOVA on Different Types of Feedback Performance

Source	SS	*df*	MS	*F*	*p*	η^2	Comparison results
Directive feedback on surface features							
PR approaches (A)	104.00	1	104.00	11.28**	0.002	0.195	EGII > EGI
Ability levels (B)	22.68	1	22.68	2.46	0.124	0.042	
A × B	0.96	1	0.96	0.10	0.748	0.002	
Error	405.54	44	9.22				
Directive feedback on content features							
PR approaches (A)	42.06	1	42.06	5.95*	0.019	0.091	
Ability levels (B)	52.19	1	52.19	7.38**	0.009	0.114	
A × B	36.83	1	36.83	5.21*	0.027	0.080	EGI_HAS > EGII_HAS
Error	311.09	44	7.07				
Criticism							
PR approaches (A)	10.46	1	10.46	10.96**	0.002	0.189	EGII > EGI
Ability levels (B)	0.10	1	0.10	0.11	0.744	0.002	
A × B	3.09	1	3.09	3.23	0.079	0.056	
Error	42.00	44	0.96				

ANOVA, analysis of variance; PR, peer response; SS, sum of squares; *df*, degree of freedom; MS, sum of mean squares; EGI, experimental group I; EGII, experimental group II; HAS, high-ability students.
$^{*}p < 0.05$; $^{**}p < 0.01$.

Table 6. The Simple Main Effect of the IPR and G-IPR on Directive Feedback on Content Features for the HAS

Item	SS	*df*	MS	*F*	*p*	Comparison results
Directive feedback on content features	82.34	1	82.34	8.70**	0.007	EGI_HAS > EGII_HAS

IPR, integrative peer response; G-IPR, game-based integrative peer response; HAS, high-ability students; SS, sum of squares; *df*, degree of freedom; MS, sum of mean squares; EGI, experimental group I; EGII, experimental group II.
$^{**}p < 0.01$.

content features ($F = 5.21$, $p = 0.027 < 0.05$, $\eta^2 = 0.080$), which was, thus, analysed with the simple main effect. The results revealed that the peer response approaches did not affect the delivery of directive feedback on content features for the LAS ($F = 0.08$, $p = 0.893 > 0.05$), but for the HAS ($F = 8.70$, $p = 0.007 < 0.01$) (Table 6). More specifically, the HAS in the EGI ($M = 6.96$, SD = 3.70) has significantly better achievement than those in the EGII ($M = 3.30$, SD = 1.99) in this aspect.

Figure 7, a framework drawn from the results presented in this section, illustrates how students with different levels of abilities reacted to the IPR and G-IPR. In summary, the G-IPR was useful in motivating students to give directive feedback on surface features and critical comments, while the IPR was helpful for students to provide directive feedback on content features, especially for the HAS. In other words, the game context may not be beneficial for all situations and all students. Such findings are coherent with those of Domínguez *et al.* (2013), which indicated that the value of GBL might not happen for everyone anytime.

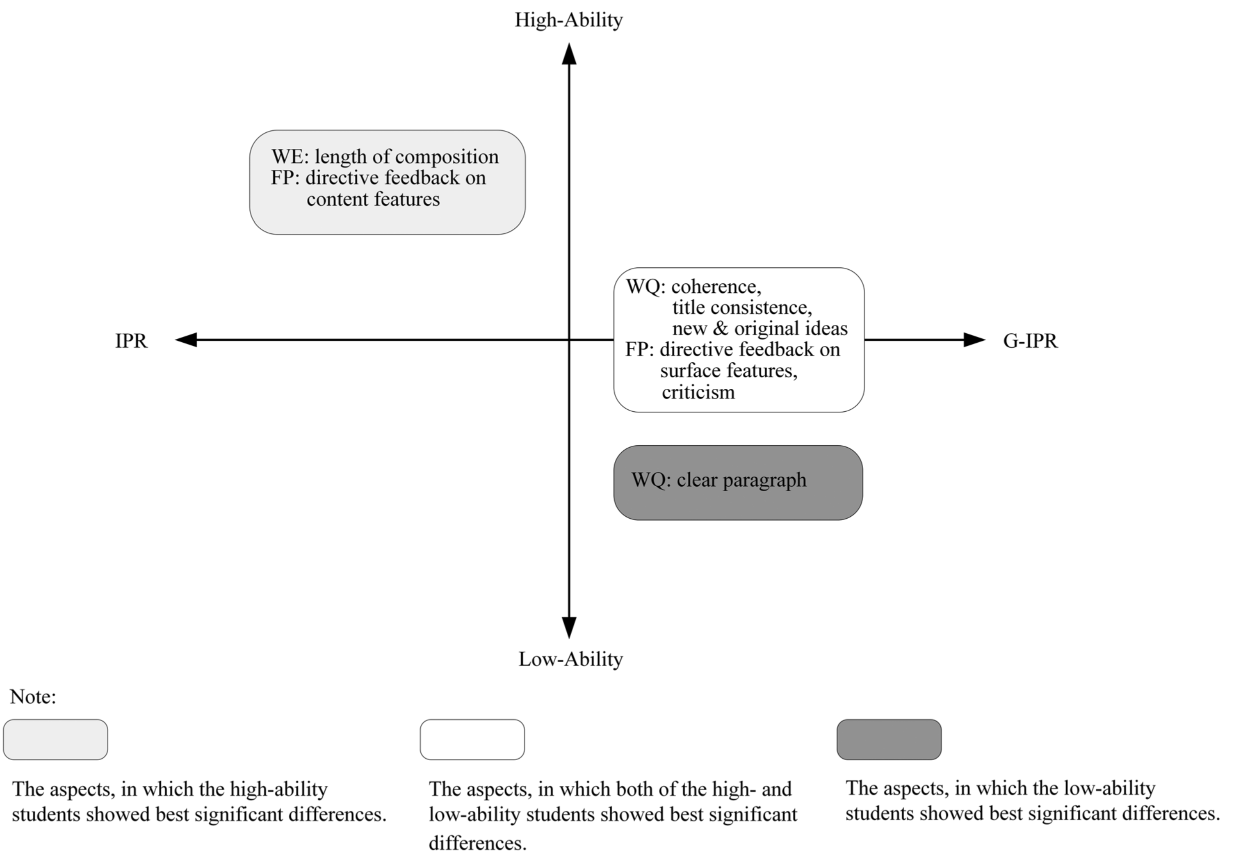

Figure 7 A Framework: Integrative Peer Response (IPR) versus Game-based Integrative Peer Response (G-IPR). WE, Written Expression; FP, Feedback Performance; WQ, Writing Quality

Discussions

The results presented in the Results section revealed that the EGI and the EGII demonstrated better writing performance than the CG. A possible reason was that the EGI and EGII used the peer response so that students could exchange drafts and feedback and clarify meanings with each other. However, all of these were hard to achieve in a teacher-centred environment. On the other hand, some interesting findings are associated with the peer response approaches and the ability levels. The details are discussed in the following subsections.

Game context versus non-game context

The improvement of written expression and writing quality was related to the peer response approaches. More specifically, the IPR, a non-game context, was useful to improve written expression, while the G-IPR, the game context, was helpful to enhance writing quality. Written expression is concerned with the richness of the content, so students expected to be guided on 'what to write in the content' via peers' feedback. On the other hand, the IPR used a forum-based environment, where students could have free discussions, including what an author wants to express, what the audience's feelings about a peer's work are or what the interesting or unclear points of a work are. Thus, such discussions may guide students to reflect on 'what to write in the content'. This may be the reason why the non-game context is helpful to improve learners' written expression. Conversely, writing quality is related to the excellence of the work, so students expected to be guided on 'how to produce an excellent work' via peers' feedback, and the G-IPR provided a structured environment, where clear goals were set for achieving an excellent work. Accordingly, the game context is beneficial for learners to improve their writing quality.

Furthermore, the usefulness of the peer response approaches is related to the types of comments delivered. Directive feedback on surface features is concerned with the suggestions for punctuation, spelling and grammar, which may not be difficult for students but sufficient carefulness is needed. The G-IPR provided a structured checklist, so it was suitable for checking the surface features of their peers' works carefully. Additionally, giving critical comments, which are related to the criticism on logic, cohesion and coherence, is relatively difficult for students, so they may have low motivation to deliver such comments. Unlike the IPR, the G-IPR was implemented in a game-based environment, which might be able to enhance students' motivation. This may be the reason why the G-IPR was useful in delivering critical comments. Conversely, the IPR provided an unstructured forum, which could stimulate students' free discussion so that some interesting ideas could be generated to improve the content of their peers' works. This may be the reason why the IPR was beneficial in giving directive feedback on content features, which might, nevertheless, be very challenging to students, so such an impact has happened for the HAS only.

High ability versus low ability

The HAS and LAS shared some similar perceptions for the peer response approaches, but some differences also existed between them. Regarding similarities, students with the G-IPR, regardless of high ability or low ability, significantly performed better than those with the IPR, in terms of both feedback performance and writing performance. The former included directive feedback on surface features and criticism, while the latter contained coherence, title consistence and new and original ideas.

Regarding differences, the LAS with the G-IPR has significantly better achievement than those with the IPR in the aspect of clear paragraph, while the HAS with the IPR outperformed those with the G-IPR in the aspects of directive feedback on content features and the length of composition. These findings generally revealed that the G-IPR was more helpful for the LAS, while the IPR was more useful for the HAS. In general, the LAS lack learning motivation, but the G-IPR included some game elements, which could stimulate their motivation. For example, the G-IPR used virtual currency to reward students when they accomplished the tasks. Our findings are coherent with those of Brewer *et al.* (2013) and Eleftheria *et al.* (2013), which indicated that rewards could increase students' motivation. Conversely, the IPR employed an unstructured forum environment where students could have free discussion, which was not only useful to improve their peers' works but also helpful to reflect their own works. Hence, they were stimulated to create some interesting ideas so that they could obtain the sense of achievement, which was valuable to the HAS. This might be the reason why the HAS benefited from the IPR, instead of from the G-IPR.

In brief, the IPR, which employed the non-game context, and the G-IPR, which was implemented with the game context, were appreciated by the HAS and LAS, respectively. In other words, the game context and the non-game context may be suitable for students with different levels of abilities. These findings echo the claim made by Abramovich, Schunn and Higashi (2013), indicating that game elements may have various impacts on the motivation of students with different levels of abilities, which, in turn, affects their participation and engagement in learning activities. Therefore, there is a need to consider the levels of students' abilities when games are incorporated into peer response.

Possible generalization

The results presented in this study can be generalized for most GBL systems, especially game-based peer response, because we chose the most commonly used game elements. We also think that, to a large extent, our findings are likely to be valid for the majority of digital games as well. However, students' experience shown in this study cannot be guaranteed as outcomes for each learner (Seaborn & Fels, 2015). More specifically, the game used in this study (i.e., G-IPR) focuses on delivering peer feedback for elementary students, so the results cannot be generalized straightforwardly for games employed to support students who take courses in high schools or universities.

Furthermore, it is known that the usage of a single game cannot be easily generalized to all games (Hays, 2005) because diversity exists among such games. The G-IPR is a role-playing game, so our results can be most likely generalized for most role-playing games. Nevertheless, caution should be given when the results are generalized to other types of games, such as simulation games or problem-solving games, which may require different learner skills and abilities. Therefore, there is a call for follow-up studies to ascertain whether results presented in this study can be generalized to other types of games.

Conclusions

Three research questions were investigated in this study. Regarding the first research question, *whether students with different peer response approaches performed differently*, the results suggest that the G-IPR was more useful in improving all students' writing quality than the IPR. Additionally, the former was more helpful in giving direct feedback on surface feature and criticism than the latter. Regarding the second research question, *whether HAS with different peer response approaches performed differently*, we found that the IPR was more beneficial for the HAS than the G-IPR in the feedback performance of directive feedback on content features and in the written expression of length of composition. Regarding the third research question, *whether LAS with different peer response approaches performed differently*, the results reveal that the G-IPR was advantageous in enhancing the LAS's writing quality in the aspect of clear paragraph.

As shown in these answers, fruitful results are provided by this study, which, however, incorporates a small-scaled sample only. Therefore, further work needs to be undertaken with a larger-scaled sample to provide additional evidence. Furthermore, this study focused on students' ability levels, so future research can take into account other human factors, such as gender differences and culture background.

Acknowledgements

The authors would like to thank the Ministry of Science and Technology of the Republic of China, Taiwan, for financial support (MOST 101-2511-S-008-016-MY3, MOST 104-2511-S-008-008-MY3 and MOST 104-2811-S-008-007) and the Research Center for Science and Technology for Learning, National Central University, Taiwan.

References

Abramovich, S., Schunn, C., & Higashi, R. M. (2013). Are badges useful in education? It depends upon the type of badge and expertise of learner. *Educational Technology Research and Development*, *61*, 217–232.

Artemeva, N., & Logie, S. (2002). Introducing engineering students to intellectual teamwork: The teaching and practice of peer feedback in the professional communication classroom. *Language and Learning across the Disciplines*, *6*(1), 62–85.

Bagley, E. A., & Shaffer, D. W. (in press). Stop talking and type: Comparing virtual and face-to-face mentoring in an epistemic game. *Journal of Computer Assisted Learning*.

Bolzer, M., Strijbos, J. W., & Fischer, F. (2015). Inferring mindful cognitive-processing of peer-feedback via eye-tracking: Role of feedback-characteristics, fixationdurations and transitions. *Journal of Computer Assisted Learning*, *31*, 422–434.

Boscolo, P., & Ascorti, K. (2004). Effects of collaborative revision on children's ability to write understandable narrative texts. In G. Rijlaarsdam, L. Allal, L. Chanqouy & P. Largy (Eds.), *Revision: Cognitive and instructional processes. Studies in writing* (Vol. *13*, pp. 157–170). Boston, MA: Kluwer.

Brewer, R., Anthony, L., Brown, Q., Irwin, G., Nias, J., & Tate, B. (2013). Using gamification to motivate children to complete empirical studies in lab environments. In *the Proceedings of the 12th International Conference on Interaction Design and Children* (pp. 388–391). New York: ACM

Chang, K., Wu, L., Weng, S., & Sung, Y. (2012). Embedding game-based problem-solving phase into problem-posing system for mathematics learning. *Computers & Education*, *58*, 775–786.

Chanquoy, L. (2001). How to make it easier for children to revise their writing: A study of text revision from 3rd to 5th grades. *British Journal of Educational Psychology*, *71*(1), 15–41.

Cheng, W., Lam, S., & Chan, C. (2008). When high achievers and low achievers work in the same group: The roles of group heterogeneity and processes in project-based learning. *British Journal of Educational Psychology*, *78*, 205–221.

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

Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on writing: Typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication*, *23*, 260–294.

Ciftci, H., & Kocoglu, Z. (2012). Effects of peer e-feedback on Turkish EFL students' writing performance. *Journal of Educational Computing Research*, *46*(1), 61–84.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

DiPardo, A., & Freedman, S. W. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research*, *58*, 119–149.

Domínguez, A., Saenz-De-Navarrete, J., De-Marcos, L., Fernández-Sanz, L., Pagés, C., & Martínez-Herráiz, J. J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education*, *63*, 380–392.

Elbow, P. (1973). *Writing without teachers*. London, England: Oxford University Press.

Eleftheria, C.A., Charikleia, P., Iason, C.G., Athanasios, T., & Dimitrios, T. (2013). An innovative augmented reality educational platform using gamification to enhance lifelong learning and cultural education. In *the Proceedings of the 4th International Conference on Information, Intelligence, Systems and Applications* (pp. 1–5). Piraeus: IEEE.

Fuchs, D., Fuchs, L. S., Mathes, P. G., & Simmons, D. C. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*(1), 174–206.

Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave MacMillan.

Ghaith, G. M. (2001). Learners' perceptions of their STAD cooperative experience. *System*, *29*, 289–301.

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences*. Hawaii, USA (in press).

Hansen, J. G., & Liu, J. (2005). Guiding principles for effective peer response. *ELT Journal*, *59*(1), 31–38.

Hays, R. T. (2005). *The effectiveness of instructional games: A literature review and discussion*. Technical report 2005-004. Orlando, FL: Naval Air Warfare Center, Training Systems Division.

Ho, M., & Savignon, S. J. (2007). Face-to-face and computer-mediated peer review in EFL writing. *CALICO Journal*, *24*, 269–290.

Hwang, G. -J., Hung, C.-M., & Chen, N. -S. (2014). Improving learning achievements, motivations and problem-solving skills through a peer assessment-based game development approach. *Educational Technology Research and Development*, *62*, 129–145.

Ku, O., Chen, S. Y., Wu, D. H., Lao, A. C. C., & Chan, T. W. (2014). The effects of game-based learning on mathematical confidence and performance: High ability vs. low ability. *Educational Technology & Society*, *17*, 65–78.

Leki, I. (1990). Coaching from the margins: Issues in written response. In B. Kroll (Ed.), *Second Language Writing: Research Insights for the Classroom*, (pp. 57–68). New York, NY: Cambridge University Press.

Liang, M. Y. (2010). Using synchronous online peer response groups in EFL writing: Revision-related discourse. *Language, Learning and Technology*, *14*(1), 45–64.

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

Lockhart, C., & Ng, P. (1993). How useful is peer response. *Perspectives*, *5*(1), 17–29.

Maratou, V., Chatzidaki, E., & Xenos, M. (in press). Enhance learning on software project management through a role-play game in a virtual world. *Interactive Learning Environments*.

Marsh, T. (2011). Serious games continuum: Between games for purpose and experiential environments for purpose. *Entertainment Computing*, *2*, 61–68.

Mendonça, C. O., & Johnson, K. E. (1994). Peer review negotiations: Revision activities in ESL writing instruction. *TESOL Quarterly*, *28*, 745–769.

Min, H. T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, *15*, 118–141.

National Commission on Writing (2004). *Writing: A ticket to work…or a ticket out*. New York: The College Entrance Examination Board.

Palloff, R. M., & Pratt, K. (2007). *Building online learning communities: Effective strategies for the virtual classroom*. San Francisco, CA: Jossey-Bass.

Papastergiou, M. (2009). Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation. *Computers & Education*, *52*(1), 1–12.

Peckham, I. (1996). If it ain't broke, why fix it? Disruptive and constructive computer-mediated response group practices. *Computers and Composition*, *13*, 327–339.

Pivec, M. (2007). Editorial: Play and learn: Potentials of game-based learning. *British Journal of Educational Technology*, *38*, 387–393.

Prensky, M. (2001). *Digital game-based learning*. New York: McGraw Hill.

Sabitzer, B. (2013). Games for learning: A neurodidactical approach to computer science. *International Journal of Science, Mathematics and Technology Learning*, *19*, 45–55.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, *74*, 14–31.

Sims, D. (2001). *Improving elementary school students' writing using reading and writing integration strategies*. Illinois: ERIC Document Reproduction Service No. ED454502.

Stoddard, B., & MacArthur, C. A. (1993). A peer editor strategy: Guiding learning-disabled students in response and revision. *Research in the Teaching of English*, *27*(1), 76–103.

Sung, H.-Y., Hwang, G.-J., & Yen, Y. -F. (2015). Development of a contextual decision-making game for improving students' learning performance in a health education course. *Computers & Education*, *82*, 179–190.

Tiene, C. D. (2000). Online discussions: A survey of advantages and disadvantages compared to face-to-face discussions. *Journal of Educational Multimedia & Hypermedia*, *9*, 371–384.

Tsui, A. B. M., & Ng, M. (2000). Do secondary L2 writers benefit from peer comments? *Journal of Second Language Writing*, *9*, 147–170.

Tüzün, H., Yılmaz-Soylu, M., Karakuş, T., İnal, Y., & Kızılkaya, G. (2009). The effects of computer games on primary school students' achievement and motivation in geography learning. *Computers & Education*, *52*(1), 68–77.

Yang, Y. F. (2011). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*, *42*, 687–700.

Yang, J. C., Ko, H. W., & Chung, I. L. (2005). Web-based interactive writing environment: Development and evaluation. *Educational Technology & Society*, *8*, 214–229.