



Research trends and applications of technology-supported peer assessment: a review of selected journal publications from 2007 to 2016

Qing-Ke Fu¹ · Chi-Jen Lin² · Gwo-Jen Hwang² 

Received: 3 October 2018 / Revised: 9 November 2018 / Accepted: 31 January 2019 /
Published online: 11 February 2019
© Beijing Normal University 2019

Abstract The study systematically reviewed 70 empirical studies on technology-supported peer assessment published in seven critical journals from 2007 to 2016. Several dimensions of the technology-supported peer-assessment studies were investigated, including the adopted technologies, learning environments, application domains, peer-assessment mode, and the research issues. It was found that, in the 10 years, there was a slight change in peer-assessment studies in terms of the adopted technologies, which were mostly traditional computers. In terms of learning environments, in the first 5 years, most activities were conducted online after class, while in the second 5 years, more activities were conducted in the classroom during school hours. Moreover, several researchers have started to consider peer assessment as a frequently adopted teaching strategy and have tried to integrate other learning strategies into peer-assessment activities to strengthen their effectiveness. In the meantime, it was found that little research engaged students in developing peer-assessment rubrics; that is, most of the studies employed rubrics developed by teachers. In terms of research issues, developing students' higher-order thinking received the most attention. For future studies, it is suggested that researchers can explore the value and effects of adopting emerging technologies (e.g., mobile devices) in peer assessment as well as engaging students in the development of peer-assessment

✉ Gwo-Jen Hwang
gjhwang.academic@gmail.com

Qing-Ke Fu
kekesoblue@gmail.com

Chi-Jen Lin
chi.jen1023@gmail.com

¹ School of Teacher Education, Huzhou University, Huzhou, People's Republic of China

² Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology, 43, Sec. 4, Keelung Rd., Taipei 106, Taiwan

rubrics, which might enable them to deeply experience the tacit knowledge underlying the standard rubrics provided by the teacher.

Keywords Applications in subject areas · Interactive learning environments · Pedagogical issues · Teaching/learning strategies

Introduction

Peer assessment aims to engage learners in scoring peers' work and providing constructive learning suggestions to peers based on the rubrics suggested by the teacher, enabling learners to have the opportunity to make reflections and improvements via playing the role of an assessor as well as an assessee (Hsia et al. 2016a; Li et al. 2012; Mulder et al. 2014; Tenorio et al. 2016; Topping 1998; van Popta et al. 2017). Peer assessment has been applied to diverse domains such as medicine, arts and humanities, science, computer science, and mechanics, indicating that its educational value is overwhelming (Falchikov 1995; Falchikov and Goldfinch 2000; Hsia et al. 2016a; Hsu 2016; Speyer et al. 2011; Yu and Liu 2009).

In the past few decades, the impacts of peer assessment have been widely investigated by researchers from different perspectives. For example, Topping (1998) discovered that peer assessment could have great effects on learning achievement, attitude, presentation skills, group work or projects, and some professional skills. Falchikov and Goldfinch (2000) stated that peer assessment was as effective as assessment by an instructor. Yu (2011) examined the directionality of peer assessment, and found that two-way peer assessment could bring about more interaction and greater assistance compared to one-way peer assessment. Yu and Sung (2016) announced that there was no significant difference in assessors' targeting behaviors in different identity revelation modes (e.g., Anonymity vs. Non-anonymity). Hsia et al. (2016a) asserted that it is very important to develop and clarify the assessment criteria for students' better understanding. Lai and Hwang (2015) further emphasized the positive impacts of engaging students in developing peer-assessment rubrics on students' learning motivation, meta-cognitive awareness, and learning achievements.

With the rapid development of computer technology, technology-supported peer assessment started to become a novel educational phenomenon in the 1990s. Compared to face-to-face peer assessment, technology-supported peer assessment overcomes the limitations of time and space, making the collection and processing of information more convenient (Li et al. 2008; Lu and Law 2012; Yu and Wu 2011). Several researchers have reported that the use of technological devices could relieve some negative effects such as anxiety, nervousness, and shyness regarding being an assessor or assessee (Carson and Nelson 1996; Chen 2016).

In light of the abundance of research on technology-assisted peer assessment in the past few decades, some researchers have attempted to conduct systematic reviews of peer assessment from different aspects. For example, Tenorio et al. (2016) systematically examined 44 target studies related to peer assessment in

online learning environments published from 2004 to 2014 by considering the aspects of participants, research goals, learning effects, benefits to teachers, and difficulties encountered. Chen (2016) investigated 20 studies of computer-supported peer assessment in English writing courses for English as second or foreign language (ESL/EFL) students by searching for papers published from 1990 to 2010, and Li et al. (2016) compared 69 papers on computer-supported peer assessment published since 1999 to explore the relationship between the ratings given by peers and by the teacher. That is, most of the systematic reviews on technology-assisted peer assessment have mainly emphasized specified learning environments, courses, or issues. To the best of our knowledge, no systematic review has been conducted for technology-supported peer assessment from diverse aspects, such as the use of new technologies and the application domains, peer-assessment modes, and research issues investigated. As suggested by several researchers, it is important to conduct review studies so as to identify the trends and potential research issues of technology-enhanced learning based on a theoretical model (Hung et al. 2018). Therefore, by referring to the theoretical model adopted by several literature review studies for technology-enhanced learning (Lin and Hwang 2018; Hung et al. 2018), the present study aims to explore the development and trends of technology-supported peer assessment in terms of the aspects of peer-assessment mode, participants and application domains, learning environments, and research issues, in order to provide some inspiration, guidance, and suggestions for future educators and researchers. Accordingly, the following research questions were investigated:

- (1) What was the overall development situation and what were the trends of technology-supported peer assessment from 2007 to 2016?
- (2) What was the development situation of technology-supported peer assessment in terms of learning devices and learning environments from 2007 to 2016? What were the development trends in the second (2012–2016) period compared to the first (2007–2011) period?
- (3) What was the development situation of technology-supported peer assessment in terms of participants and application domains from 2007 to 2016? What were the development trends in the second (2012–2016) period compared to the first (2007–2011) period?
- (4) What was the development situation of technology-supported peer assessment in terms of peer-assessment mode (identity revelation mode, peer feedback mode, and the source of the rubrics) from 2007 to 2016? What were the development trends in the second (2012–2016) period compared to the first (2007–2011) period?
- (5) What was the development situation of technology-supported peer assessment in terms of the research issues from 2007 to 2016? What were the development trends in the second (2012–2016) period compared to the first (2007–2011) period?

Method

In this section, the data collecting and coding procedures of the present study are presented. Those procedures were conducted following the suggestions of several previous studies, such as Lin and Hwang (2018), Hwang and Tsai (2011), and Hung et al. (2018).

Data collection

According to the definition of peer assessment in the present study, the search phrases (“peer assessment” OR “peer feedback” OR “peer comment”) were identified. Moreover, in order to acquire high-quality research related to technology-assisted learning, target studies were searched from seven SSCI (Social Science Citation Index) journals of technology-supported learning: *Computers & Education* (C&E), *British Journal of Educational Technology* (BJET), *Educational Technology & Society* (ETS), *Educational Technology Research and Development* (ETR&D), *Interactive Learning Environments* (ILE), *Journal of Computer Assisted Learning* (JCAL), and *Innovations in Education and Teaching International* (IETI), as suggested by Hwang and Tsai (2011). The types of studies were restricted to “Article,” and the time span was from 2007 to 2016. A total of 98 papers were selected in this stage.

In order to obtain those target studies closely related to the purpose of the study, we further constructed inclusion criteria to narrow down the literature; that is, target studies had to include experimental treatment using peer-assessment learning activities, and the studies must have consciously made use of peer-assessment learning activities to promote students’ development. To be more specific, the learning activities in the studies must have consisted of at least one cycle of learning: produce questions or artifacts, evaluate peers’ questions or artifacts, and reflect on or revise according to peers’ assessment. Thus, studies on system development, survey research, literature reviews, position papers, and so on were excluded. After further selection based on the inclusion criteria under the negotiation of two coders, 70 studies were finally included in the present study.

Data contribution

We classified the 70 studies based on the published journals. As shown in Table 1, the impact factors of most of the papers were high, which implied that the quality of the target studies was guaranteed. To a certain extent, they represent the development level of research on technology-assisted peer assessment in these 10 years, which was conducive to the generalization and accuracy of the study’s conclusions.

Coding scheme

Based on the research questions, for this study, we constructed the subcoding systems on the aspects of learning devices, learning environment, participants and

Table 1 The distribution of the literature on technology-supported peer assessment published from 2007 to 2016 in the 7 selected journals

Journal	Number of papers	Impact factor (2016)
1. Computers & Education	28	3.819
2. British Journal of Educational Technology	12	2.410
3. Educational Technology & Society	9	1.584
4. Interactive Learning Environments	8	1.674
5. Journal of Computer Assisted Learning	6	1.253
6. Educational Technology Research and Development	4	0.725
7. Innovations in Education and Teaching International	3	0.667

application domains, peer-assessment mode, and research issues. The coding scheme of each investigated dimension was determined by referring to the literature, and has been examined by two experienced researchers in the field, as shown in the following:

- (1) Learning devices were categorized according to four coding items: Mobile device, Traditional computer, Other, and Nonspecified.
- (2) “Learning environments” contained five coding items: Classroom (including computer room), In-field (Specified real-world environments), After-school settings (learning via the Internet without being limited by location), Other, and Nonspecified.
- (3) The coding items of “participants” proposed by Hwang and Tsai (2011) were adopted in this study, and included Elementary schools, Junior and senior high schools, Higher education, Teachers, Working adults, Others, and Nonspecified.
- (4) The coding items of application domains were adapted from the study of Hwang and Tsai (2011), including Engineering or computers, Science, Health Medical or Nursing, Social science, Arts or design, Languages, Mathematics, Business and Management, Other, and Nonspecified.
- (5) The mode of peer assessment contained three subcategories: “Privacy policy,” “Peer feedback mode,” and “the source of the rubrics,” as suggested by several researchers (Lai and Hwang 2015; Li et al. 2016; Liu and Carless 2006; van Genip et al. 2009). The coding items of “privacy policy” were determined based on the items proposed by Topping (1998), including Anonymous, Non-anonymous, and Nonspecified. The items of “Peer feedback mode” were determined according to the research of Hsia et al. (2016b), including Peer comments, Peer ratings, Mixed mode (peer ratings plus peer comments), and Nonspecified. The items of “source of the rubrics” included Constructed by the teacher, Constructed with the participation of the students, and Nonspecified, as suggested by Lai and Hwang (2015).
- (6) Research issues consisted of six categories: affection, cognition, skill, behavior, correlation, and other. Affection referred to Technology Acceptance Model/ Intention of use (including perceived ease of use and perceived usefulness of

the adopted learning system), Attitude/Motivation/Anticipation of effort, Self-efficacy/Confidence/Anticipation of performance, Satisfaction/interest, Opinions or perceptions of the learner (interview or open-ended question), and cognitive load; cognition contained Learning achievement (by tests), tendency/awareness of higher-order thinking (such as analysis, reasoning, evaluation, problem-solving, metacognition, critical thinking, or creativity), and tendency/awareness of collaboration or communication; skill referred to the accuracy and fluency of manipulation or demonstration; behavior referred to learning behavior or learning behavioral pattern analysis; correlation referred to the correlation or causal relationships between the factors related to peer assessment; other is an open coding category.

Coding procedure

After constructing the coding system, two coders first negotiated the coding scheme in order to reach consistent understanding of it; they then conducted the coding for the 70 studies independently. Finally, after finishing their own coding tasks, they compared two coding data sheets, discussed the different coding results, and negotiated until the coding results were identical.

Results

Overall development situation and trends

We classified the 70 studies according to the published years (see Fig. 1). The results indicated that there was an upward but slow movement in research on technology-assisted peer assessment from 2007 to 2016. However, in some specific years (such as 2009 and 2016), there was greater development compared to the previous year(s).

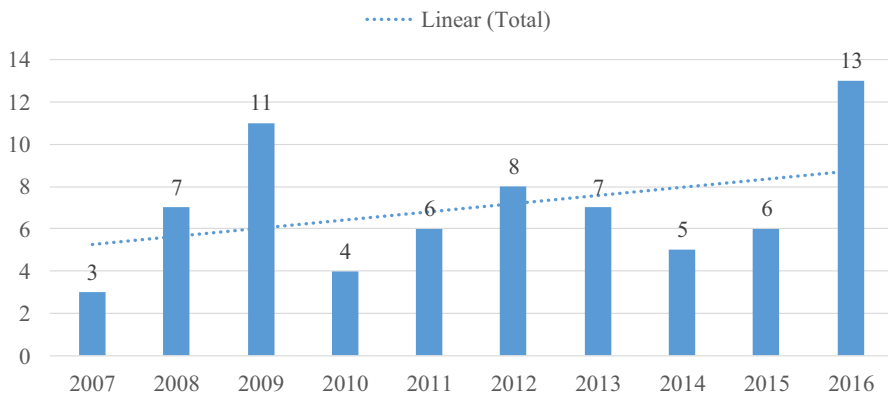


Fig. 1 Number of technology-assisted peer-assessment studies from 2007 to 2016

When the present study compared the literature published in 2009 with that in 2007 and 2008, we discovered that most studies in 2007 and 2008 put emphasis on the engineering/computer domain; nevertheless, after two studies on the curriculum of teachers' professional development in 2008, six out of 11 papers published in 2009 kept focusing on the curriculum of teachers' professional development, signifying that researchers attempted to explore the potential educational value from different angles (such as different domains).

The development situation and trends of learning devices and learning environments

As demonstrated in Fig. 2, most peer-assessment studies from 2007 to 2016 used traditional computers, while only four adopted mobile devices. In view of the fact that mobile technology and devices have begun to be largely involved in education and have shown tremendous educational value (e.g., Hwang and Tsai 2011; Wang et al. 2017), this reflects that the educational value of peer assessment supported by mobile devices is still in the stage which needs to be explored. Further investigation showed that one out of these four studies was published in 2010, which was in the first period (2007–2011), whereas the other three were published in the second period (2012–2016), that is, one in 2015 and two in 2016. This result implies that researchers started to focus more on the educational value of peer assessment supported by mobile devices in the second 5-year period compared to the first 5-year period.

As for the aspect of learning environments, most peer-assessment activities happened in the classroom (or computer room) and after class via online learning (see Fig. 3). Furthermore, there was almost no research conducted in authentic contexts outside of the classroom (such as ecological parks, zoos, or science museums). These findings indicate that most of the learning activities based on peer assessment still focused on traditional teaching contexts.

It was found in further analysis that the online peer-assessment activities conducted after class decreased in the second (2012–2016) period compared to the first (2007–2011) period, while peer-assessment activities in traditional classrooms

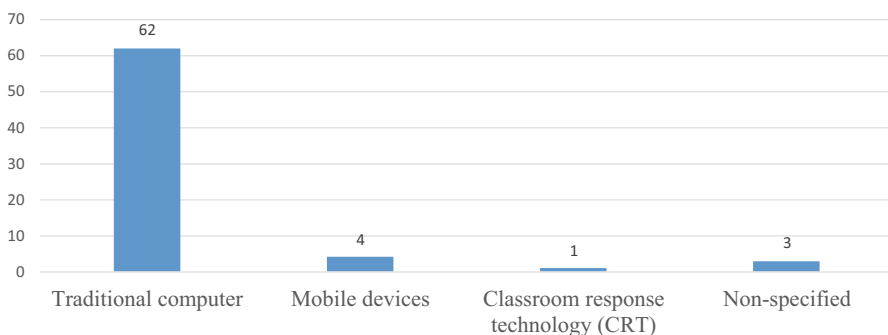


Fig. 2 Distribution of learning devices used in the technology-assisted peer-assessment literature

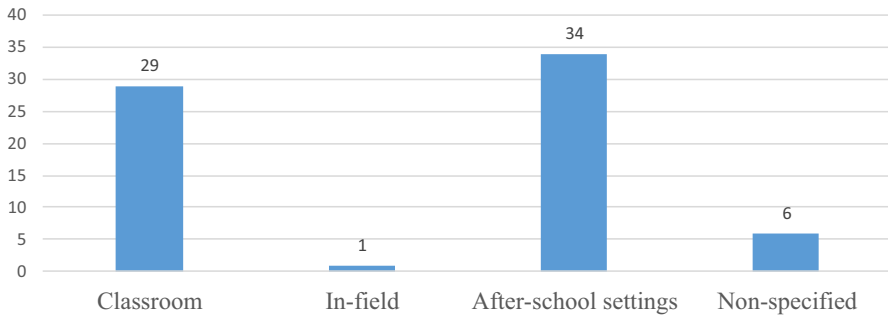


Fig. 3 Distribution of learning environments in the technology-assisted peer-assessment literature

increased significantly (see Fig. 4). This result indicates that the majority of technology-assisted peer-assessment activities have become semi-autonomous after being relatively autonomous.

The development situation and trends of participants and application domains

As shown in Fig. 5, with regard to participants, the main participants of research on technology-assisted peer assessment were higher education students, while there were not many studies on elementary schools or junior and senior high schools. Moreover, a small part of the literature focused on teachers; no research investigated working adults. These results signified that little attention has been paid to elementary schools, junior and senior high schools, and teachers. Future studies should further reinforce the research on these groups; additionally, introducing peer assessment into adult education should be focused on.

The present study also explored the development trends of participants in the first and second 5-year periods (see Fig. 6). There was no change in the quantity of research on higher education in the second (2012–2016) period compared to the first (2007–2011) period. Most participants of the peer-assessment

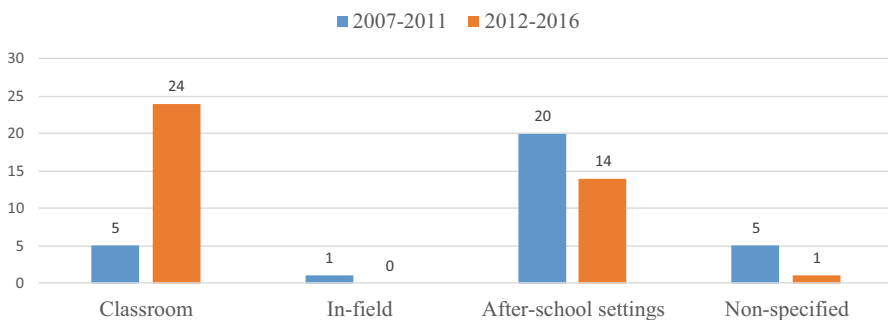


Fig. 4 Learning environments in the first (2007–2011) and second (2012–2016) 5-year periods

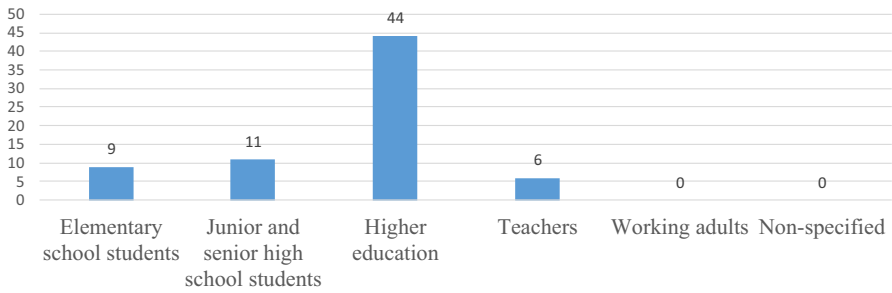


Fig. 5 Distribution of participants in the technology-assisted peer-assessment literature

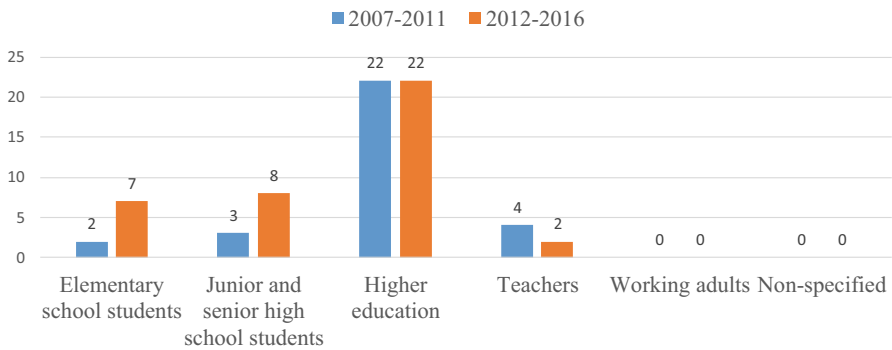


Fig. 6 Participants in the first (2007–2011) and second (2012–2016) 5-year periods

studies were also higher education students, specifying that peer-assessment studies seemed to encounter a development bottleneck in recent years, and that researchers have not yet made advances in examining the research gaps. However, there was substantial growth in the quantity of research on elementary school, and junior and senior high school students. This indicates that, in spite of the total quantity of peer-assessment studies on elementary schools and junior and senior high schools being small, more attention has been paid to these contexts in recent years. The possible reason might be that researchers discovered the specific educational value of peer assessment in primary and secondary schools, which is discussed in depth in the [Discussion](#) section.

As demonstrated in [Fig. 7](#), in terms of application domains, technology-assisted peer assessment was applied most in social science followed by engineering or computers, language, and science; there was also some application in art or design, health medical or nursing, and business/management. Further investigation indicated that out of 24 studies on social science, 18 were conducted in teacher development courses, while 7 out of 11 studies on language focused on writing courses. These findings denote that the educational value of technology-supported peer assessment was pervasive, but it had specific value for skill

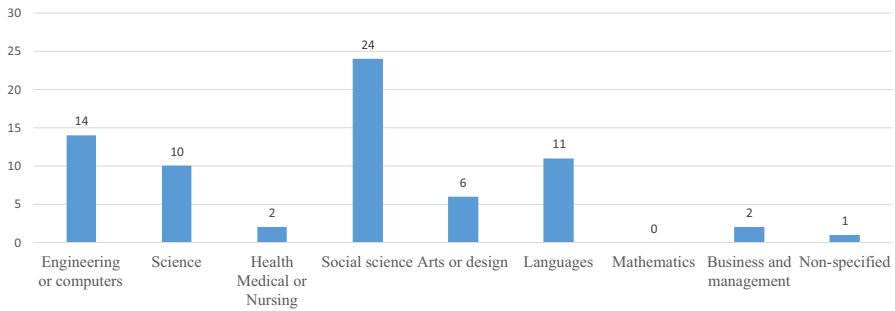


Fig. 7 Distribution of application domains in the technology-assisted peer-assessment literature

development courses which were closely related to daily life and work. Furthermore, there was no application of technology-assisted peer assessment in mathematics, suggesting that there is a need for educators and researchers to explore how to develop and apply peer-assessment activities in abstract courses in the future.

Besides, the present study attempted to disclose the development trends of learning domains in the first and second 5-year periods. As shown in Fig. 8, compared with the first (2007–2011) period, there was almost no improvement in technology-assisted peer-assessment studies in social science, engineering or computers, languages, science, and health medical or nursing in the second (2012–2016) period. This indicates again that, for the current technology-supported peer-assessment studies, it is still the time to look for new points of growth. On the other hand, the rapid increase in arts or design to some extent verifies that searching for new research growth points is an effective way to promote the development of technology-supported peer assessment.

The development situation and trends of peer-assessment mode

In the current study, peer-assessment mode was composed of three aspects: Privacy policy, Peer feedback mode, and the source of the rubrics. With regard to Privacy

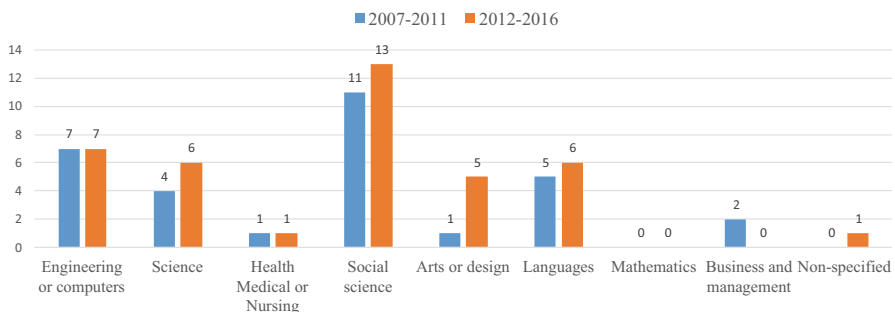


Fig. 8 Application domains in the first (2007–2011) and second (2012–2016) 5-year periods

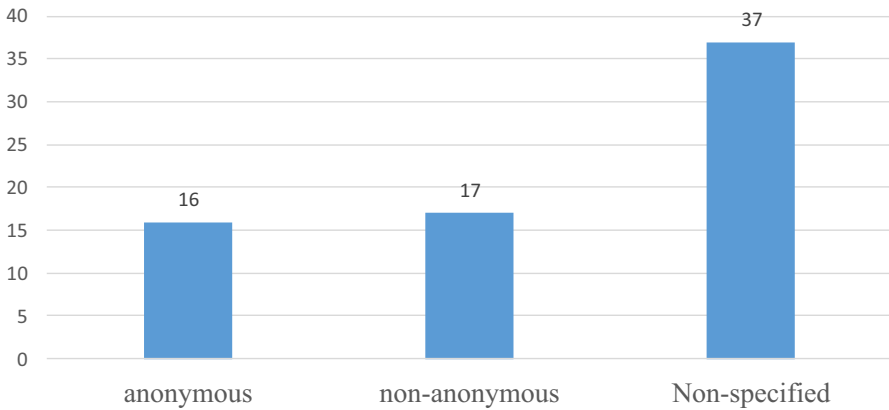


Fig. 9 Distribution of privacy policy in the technology-assisted peer-assessment literature

policy, over a half of the studies did not consider it in these 10 years (see Fig. 9). In the rest of the studies, the quantity of research which chose Anonymous or Non-anonymous peer assessment was nearly the same. It seems that researchers still hold opposing opinions on the effect of different identification revelation modes. When examining the development trends, we uncovered that, as illustrated in Fig. 10, there was a large increase in research on Non-anonymous assessment in the second (2012–2016) period compared with the first (2007–2011) period.

As demonstrated in Fig. 11, in terms of Peer feedback mode, mixed mode (peer ratings plus peer comments) was adopted most from 2007 to 2016, followed by peer comments and then peer ratings. In addition, an increasing number of studies tended to adopt mixed mode in the second (2012–2016) compared to the first (2007–2011) period (see Fig. 12).

With regard to the source of the rubrics, peer-assessment rubrics constructed by the teacher were adopted in most of the studies from 2007 to 2016 (see Fig. 13);

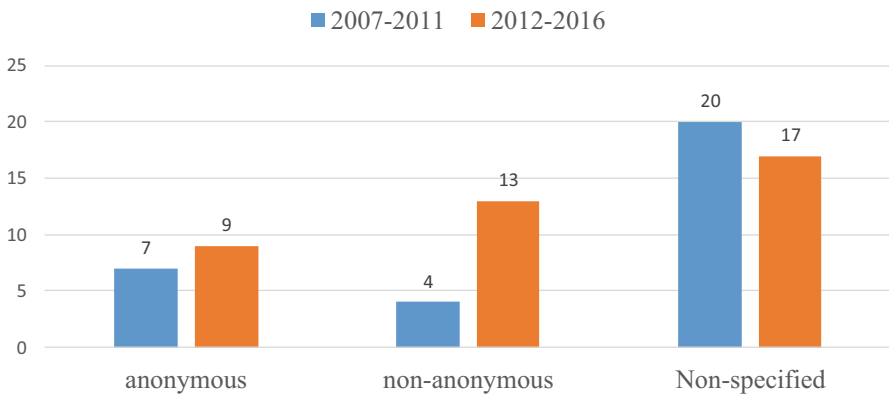


Fig. 10 Privacy policy in the first (2007–2011) and second (2012–2016) 5-year periods

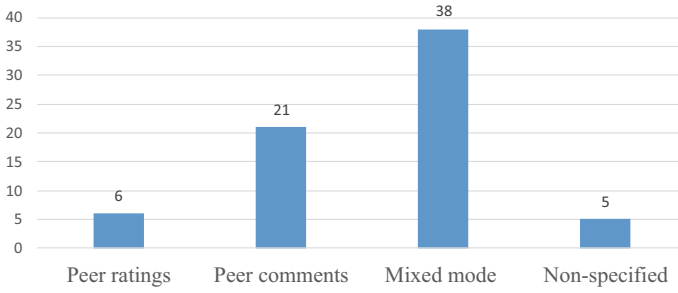


Fig. 11 Distribution of peer feedback mode in the technology-assisted peer-assessment literature

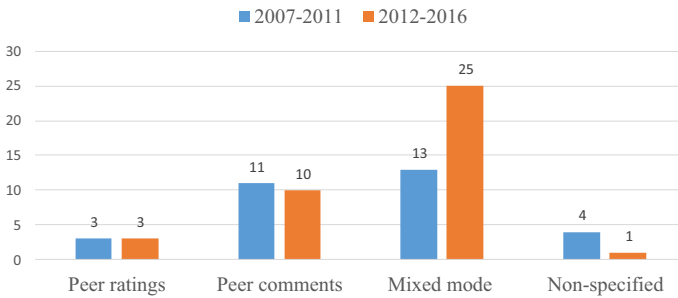


Fig. 12 Peer feedback mode in the first (2007–2011) and second (2012–2016) 5-year periods

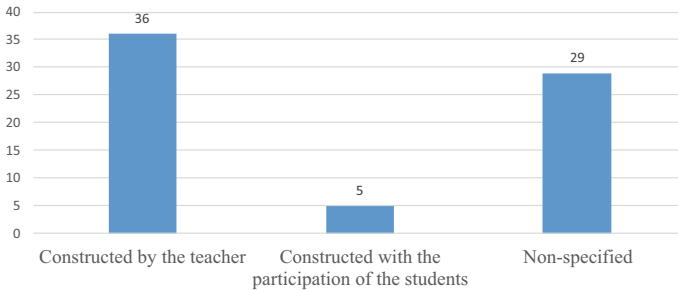


Fig. 13 The source of the rubrics in the technology-assisted peer-assessment literature

a significant amount of the literature did not identify peer-assessment rubrics, letting learners freely evaluate peers' learning. Very little research adopted the way of students' participation in constructing peer-assessment rubrics. Moreover, as illustrated in Fig. 14, more research did not identify peer-assessment rubrics in the first (2007–2011) period, while an increasing number of studies adopted peer-assessment rubrics constructed by the teacher in the second (2012–2016).

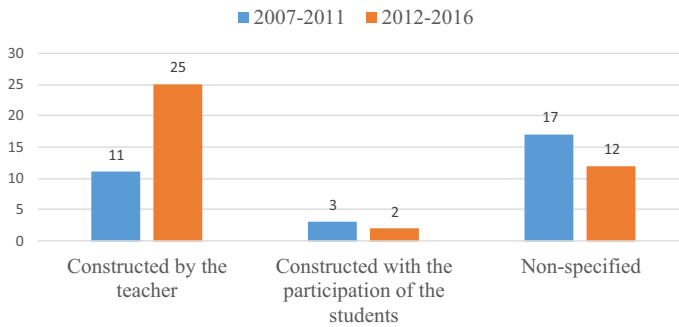


Fig. 14 The source of the rubrics in the first (2007–2011) and second (2012–2016) 5-year periods

The development situation and trends of research issues

As shown in Table 2, when using peer assessment to promote students’ development, most studies emphasized developing students’ cognition, followed by

Table 2 Distribution and growth rate of research issues from 2007 to 2016

	2007–2016	2007–2011	2012–2016	Growth rate
Cognition	79	35	44	25.7%
1. Learning achievement	22	12	10	– 16.7%
2. Tendency/awareness of higher-order thinking	42	20	22	10.0%
3. Tendency/awareness of collaboration or communication	15	3	12	300.0%
Affection	67	25	42	68.0%
1. Technology acceptance model/intention of use	10	4	6	50.0%
2. Attitude/motivation/anticipation of effort	23	8	15	87.5%
3. Self-efficacy/confidence/anticipation of performance	4	1	3	200.0%
4. Satisfaction/interest	2	1	1	0.0%
5. Cognitive load	2	0	2	NA
6. Opinions or perceptions of the learners	26	11	15	36.4%
Skill				
1. The accuracy and fluency of manipulation or demonstration	5	0	5	NA
Behavior				
1. Learning behavior or learning behavioral pattern analysis	23	10	13	30.0%
Correlation				
1. The correlation or causal relationships between the factors related to peer-assessment	23	8	15	87.5%
Other				
1. Workload of teachers	1	0	1	NA

affection, behavior, correlation, and skills. To be more precise, developing students' tendency or awareness of higher-order thinking received the most attention; such outcomes as learning achievement, learning attitude or motivation or anticipation of effort, opinions or perceptions of the learners, learning behavior or learning behavioral pattern analysis, and the correlation or causal relationships also obtained some attention. A little emphasis was put on the two issues: tendency or awareness of collaboration or communication, and technology acceptance model or intention of use. In contrast to the first (2007–2011) period, skill had the rapidest growth rate; correlation and affection also had faster development. Little development was found in cognition and behavior outcomes; among all of the issues, attitude/motivation/anticipation of effort and self-efficacy/confidence/anticipation of performance in affective outcomes, and collaboration or communication in cognition outcomes grew rapidly in the second (2012–2016) 5-year period.

Discussion

To further explore trends and the potential educational values of technology-supported peer assessment, several dimensions, including learning devices and environments, participants and application domains, the mode of peer assessment, and research issues, are discussed based on the findings, as follows.

Learning devices and environments

From 2007 to 2016, the devices used in technology-based peer assessment were mostly traditional computers, disclosing the distinctive educational value of computers in peer assessment; yet, researchers have pointed out several existing problems in conducting peer-assessment activities (e.g., Guardado and Shi 2007; Hwang et al. 2014). For instance, Hulsman and van der Vloodt (2015) found that the students generated more negative annotations which were directed to specific questions, whereas peer feedback was more positive but general. These results suggest that, in contrast to face-to-face peer assessment, computer technology-based peer assessment expanded the evaluation content (such as video learning content) and format (such as online annotation and feedback), and could also break through the time and space limitations. Nevertheless, there was less helpful peer feedback due to the participants' interpersonal relationships. It could be seen that technology-assisted peer assessment would not naturally ease the anxiety and nervousness which usually exist in face-to-face peer assessment. Li et al. (2016) uncovered that, compared to paper-based peer assessment, the consistency between computer-assisted peer assessment and instructor's assessment was even lower because of the insufficient instructional support mentioned by Suen (2014). Guardado and Shi (2007) suggested that even though peer assessment created more opportunities for students' learning, it did not ensure that students would have stronger learning motivation or more autonomy to learn outside the classroom. Hwang et al. (2014) implied that it was limited to improving students' performance if they just received peers' feedback rather than

conduct self-reflection. Cho et al. (2010) indicated that, in contrast to students who had no prior self-monitoring training, the writing performance of those who had self-monitoring training improved significantly in the same online peer-assessment environments. Moreover, Miller (2003) also suggested that high-quality peer assessment requires teachers to organize, monitor, and help learners; yet, teachers could seldom provide this help due to their workload in the online peer-assessment environments.

It seems that the researchers also recognized the existing problems of managing, guiding, and supporting the online peer-assessment activities conducted after class; therefore, in the second (2012–2016) 5-year period, peer assessment was directed into traditional teaching contexts (such as computer rooms) from extracurricular online environments. For example, in Luo's study in (2016), students made presentations while others used cell phones to submit feedback synchronously by Twitter in a face-to-face teaching activity; the teacher regularly checked the feedback posted by the students in class. The study denoted that students could listen to the presentation and point out the problems and errors at the same time. Lai and Hwang (2015) designed a classroom teaching activity to facilitate students' development of peer-assessment rubrics, and their learning enthusiasm and initiative were much stronger. Additionally, the support based on mobile devices made peer interaction more frequent and rapid.

Consequently, it is inferred that most teachers preferred to conduct peer-assessment activities in class, meaning that they could have more face-to-face interactions with students. On the other hand, fixing the time and location of learning activities in class might lead to some limitations, such as insufficient time for reviewing peers' work and providing feedback. From this perspective, it could be better to conduct peer-assessment activities after class. Moreover, although mobile devices were not frequently adopted in previous studies, several researchers have indicated the potential of using mobile technologies to help students interact with peers both in and out of class (Lai and Hwang 2014). Accordingly, the present study signifies the potential of conducting peer-assessment activities with the support of new technologies (such as mobile devices and wearable equipment) to promote peer interactions using different learning designs.

Participants and application domains

The findings indicated that higher education students were the main research participants of studies on technology-assisted peer assessment from 2007 to 2016, while little research has explored elementary schools or junior and senior high schools. Several studies also found that most of the peer-assessment activities took place in higher education environments (Falchikov and Goldfinch 2000; Li et al. 2016). The reason is that complicated cognitive processing (such as higher-order thinking) is usually involved in peer assessment. As can be seen from its underlying learning theories, peer assessment could require students to not only be learners in the traditional sense of the word, but to also assume such roles as observer, listener,

collaborator, thinker, and even creator. This is challenging for those younger than university students.

While studies on technology-supported peer assessment developed slowly in higher education, at the same time, researchers have started to put emphasis on research on elementary schools and on junior and senior high schools in recent years. This finding shows that the educational value of peer assessment is oriented toward students of all ages. For instance, Hwang et al. (2014) designed a peer-assessment-based game development approach to develop the learning performance, motivation, and problem solving of sixth-grade students. Lai and Hwang (2015) constructed the learning strategies of designing and developing peer-assessment rubrics by students, which improved fifth-grade students' learning motivation, metacognitive awareness, and ability to produce posters. Hsu (2016) pointed out that a peer-assessment activity based on the grid-based knowledge classification approach significantly increased students' learning performance in contrast to conventional peer-assessment activities.

Moreover, the value of technology-assisted peer assessment was fully embodied in social science, engineering or computers, language, and science from 2007 to 2016. Of the 24 studies on technology-supported peer-assessment in social science, 18 were to promote teacher professional development (including pre-service teacher training), while 7 out of 11 studies on language focused on developing writing skills. The possible reason may lie in the fact that students have more initiative and active attitudes in developing skills closely related to daily life or work, and more practical experience to support peer interaction. Falchikov and Goldfinch (2000) denoted that peer assessment was more effective than professional practice (e.g., teaching practice) in academic areas (e.g., essays, presentations). Li et al. (2016) proposed the possible reason that academic learning tasks have more specific right or wrong answers compared to professional training, which is easier for evaluators to evaluate. The current study indicates that peer assessment has a crucial value, no matter whether in academic domains or practical training domains; it is important to take measures to promote students' motivation in peer interaction. For example, there were no studies on technology-assisted peer assessment in mathematics in the recent 10 years; perhaps it was because students are normally not interested in mathematics, and the existing mathematics teaching content is dissociated from real life, making it difficult to integrate peer-assessment strategies into mathematics teaching activities. Arts or design are subject fields that are closer to our life, and thus it is easier to inspire learning interest in these subjects. In the recent 5 years, the development of technology-assisted peer assessment in arts or design has confirmed the perspectives of the present study to a certain degree.

The mode of peer assessment

Different ways of identity revelation have varying influences on learning behaviors and effects. Several researchers such as Vanderhoven et al. (2015) and Yu and Liu (2009) suggested that non-anonymity would bring more negative emotions for participants, while anonymity would result in a greater sense of psychological safety.

Yet, other researchers had different findings. Yu and Sung (2016) explored the online peer-assessment behaviors of fifth-grade students, and discovered that revealing the evaluator's identity did not influence evaluation behaviors. Students chose the participants to be evaluated mainly based on the questions rather than on who asked the questions. Ching and Hsu (2016) indicated that positive interpersonal relationships (e.g., psychological safety and trustworthiness) would affect peer-assessment performance; that is, if peers have not yet built positive interpersonal relationships under the anonymous mode, it would have some negative performance due to de-individualization.

Above all, a portion of the researchers did not consider the problems that might be caused by identity revelation because the studies took place in intact classes, where interpersonal interactions between peers occurred frequently. As a result, it was understandable that the studies did not try to use an anonymous approach. Additionally, even though it was regarded that identity revelation would have a great influence on learners' interactions (Topping 1998; Vanderhoven et al. 2015), there were plenty of researchers starting to actively look at students' interaction behaviors when their identities were uncovered. For instance, Li et al. (2016) implied that peer assessment was more effective when the assessment was non-anonymous rather than anonymous; this might be because assessors would evaluate more seriously when their identity was known. Thus, the present study discovered that an increasing number of researchers adopted non-anonymity in the recent 5 years, signifying that they started to put emphasis on the positive aspect of learners' social interaction.

In terms of peer feedback mode, most researchers believed that assessors would have more cognitive processing and conduct self-reflection as they need to produce constructive comments for their peers. Also, assesseees would benefit by receiving meaningful peer comments, and it was thereby more valuable to both assessors and assesseees than just proposing peer ratings (Falchikov 1995; van Gennip et al. 2009; van Popta et al. 2017). Lu and Law (2012) found that peer comments were the factor influencing learning performance, but peer ratings would not affect learning performance. They argued that commenters would be more involved in critical thinking activities when they raised comments. Furthermore, Hsia et al. (2016b) indicated that the mixed mode (peer ratings plus comments) was better than peer comments alone, and the behavior analysis denoted that the former could significantly increase students' willingness to participate in online learning. Falchikov (1995), however, discovered that students seemed to be unwilling to rate their peers.

The present study also found that the mixed mode was adopted the most from 2007 to 2016, followed by peer comments; peer ratings alone were adopted the least. Moreover, in contrast to the first (2007–2011) period, it seemed that studies adopting the mixed mode doubled over the second (2012–2016) period, whereas no growth was found in the research on peer comments and peer ratings. These results verified the findings or conclusions of the abovementioned studies.

Peer-assessment rubrics are an important component of peer assessment (Fraile et al. 2017; Stefani 1994). Topping (1998) believed that it was essential to clarify and demonstrate peer-assessment rubrics to students; furthermore, it would be much better if students could participate in the development and elaboration of peer assessment. Falchikov and Goldfinch (2000) found that it would increase the

consistency between peer assessment and teacher assessment if students took part in constructing peer-assessment rubrics rather than only teachers developing the rubrics. Li et al. (2016) also considered peer assessment as being more effective when evaluators participated in developing peer-assessment rubrics. However, some researchers had some different findings. Fraile et al. (2017) conducted research with an experimental group using co-created rubrics and a control group which just used rubrics. They believed that students would use language which can be understood mutually when they co-created rubrics, thereby having a bigger influence on their learning.

It could be seen from the above studies that most of them recognized the importance of peer-assessment rubrics. On the other hand, there was still plenty of research which did not create them. Jones and Alcock (2014) indicated that when there were no rubrics, it would encourage students to more freely express their opinions; moreover, they would pay attention to their own thinking without the restriction of rubrics, thereby allowing them to generate richer and more abundant perspectives, and developing the students' mutual knowledge. O'Donovan et al. (2004) also indicated that it was a great challenge to let students clarify and understand rubrics, and that it was even meaningless sometimes as different people or groups had different understandings of the rubrics; furthermore, the tacit knowledge expressed in rubrics is hard to describe and clarify. The current study further analyzed the literature which did not adopt peer-assessment rubrics, and discovered that these studies mostly encouraged students to discuss freely and evaluate each other, and to develop students' higher-order skills. According to O'Donovan et al. (2004), the best way to learn tacit knowledge is through the process of experiential learning.

The present study indicates that the benefits of peer-assessment rubrics constructed with the participation of students were greater than those of the approach by which students directly received the rubrics constructed by the teacher. It also surpassed the way of not giving students any rubrics. Consequently, future studies should be conducted to further explore the value of students participating in developing peer-assessment rubrics.

Research issues

The main purpose of researchers applying peer assessment was to develop students' cognition, especially their tendency or awareness of higher-order thinking. This reflected the critical and reflexive nature of peer assessment. Besides, researchers also took advantage of peer assessment in an attempt to encourage students to conduct higher-order cognitive processing in order to inspire them to actively and meaningfully deal with lower-order cognitive processing; that is, it assisted students in better memorizing, understanding, and applying the acquired knowledge, and finally developed the students' learning achievement. In a study conducted by Kao et al. (2008), students were divided into three cross-unit knowledge groups to learn function, class, and flow respectively in a C++ programming language course. When they created a mind map of the subject knowledge, they evaluated other groups' mind maps as well. They then acquired the cross-unit concepts, integrated their own

mind map into the cross-unit mind maps, and eventually obtained the understanding of the concepts in the three units. In this research, students kept changing between higher-order knowledge processing and lower-order knowledge processing, integrated their original concepts, peer comments, and their ongoing new knowledge, and finally constructed an active, meaningful, and contextual knowledge-processing procedure. The number of studies on collaboration and communication outcomes tripled, indicating that researchers have started to put emphasis on this issue. As collaboration and communication are crucial for students in their future life or work, future studies should keep fostering and developing students' tendency or awareness of collaboration or communication.

Researchers also paid attention to affective learning outcomes, which were related to peer assessment as a process of interacting with people. Learners' opinions or perceptions, attitude or motivation, or anticipation of effort were the affective learning outcomes which attracted most attention. In addition, compared to the first (2007–2011) period, these two outcomes had large growth in the second (2012–2016) period, indicating that researchers put emphasis on learners' affection closely related to learning, and that peer assessment is a complicated learning process in need of more qualitative analysis based on learners' perceptions. Researchers made efforts to examine this complex phenomenon from an open-minded perspective. One possible reason why little research has explored technology acceptance model/intention of use outcomes was that students are more familiar with conventional computers used in peer assessment. Such outcomes as self-efficacy, satisfaction, and cognitive load have seldom been investigated. Although there was a large growth in studies exploring self-efficacy and cognitive load in the second (2012–2016) period, the quantity was still low. Therefore, it is suggested that future studies strengthen the investigation on self-efficacy, cognitive load, and satisfaction.

As for the aspect of skill, in recent years, researchers have started to examine and use peer assessment to develop students' manipulation or demonstration skills. Hsia et al. (2016b) applied peer assessment to develop high school students' dancing skills. Students watched the dancing videos uploaded by peers, evaluated the technical and performance actions in their peers' dance, and gave suggestions. The assessee revised and practiced the dance actions according to their peers' comments, achieving good results. As a result, it is also necessary to strengthen the further investigation of the value of peer assessment in the development of manipulation or demonstration skills. Peers' learning behaviors or learning behavioral pattern analysis, and the correlation or causal relationships between factors involved in peer assessment were crucial outcomes as well.

Conclusions and suggestions

Despite researchers and school teachers possibly knowing the potential of peer assessment, the development level of technology-assisted peer assessment was stable but slow, which could be due to the fact that it takes time for teachers to get used to the technological tools and learning environment. Therefore, the potential value of technology-supported peer assessment needs to be further examined.

Based on the findings and discussion, the present study provides the following suggestions:

- (1) Focus on exploring the potential educational value of new technology-based peer assessment (e.g., mobile devices, wearable devices). For example, using mobile technology with effective learning designs has great potential in developing learners' self-management, self-learning, social interaction, awareness of reflection, and reflective strategies and abilities; however, there are few current studies with a focus on mobile technology-based peer assessment.
- (2) Keep intensifying the research on semi-autonomous peer assessment with the guidance of teachers, and explore the solutions to such questions as learning motivation, self-regulated learning, collaboration and communication, and teachers' support in student-centered autonomous peer assessment at the same time. Especially, pay attention to the application and promotion of research results in practical educational activities in order to realize the virtuous cycle of mutual development between research and teaching practice.
- (3) Investigate the value and practical strategies of mobile devices for facilitating a close combination of teaching and learning (including peer-assessment activities) so as to increase learning motivation, quality, and efficiency in peers' interactions. Especially, examine the value and practical strategies of peer-assessment activities based on mobile devices in real-life contexts.
- (4) Continue to promote the development of peer assessment in elementary schools, and junior and senior high school environments, and further investigate the forms and strategies of integrating and applying peer assessment with other teaching methods and modalities. In addition, apply this development idea to foster the development of peer assessment in higher education. Furthermore, start to explore the value of peer assessment for working adults. Keep examining the value of peer assessment in more domains or curricula, especially the potential value and implementation strategies in those domains or curricula which have abstract content and for which it is not easy to trigger learning motivation.
- (5) Develop the research on the mutual effects between interpersonal relationships and identity revelation in peer assessment; continue to intensify the studies on non-anonymity to explore the influence of the positive relationship in peer-assessment activities; pay attention to and reinforce the research on exploring the educational value of constructing peer-assessment rubrics with the participation of the students.
- (6) Intensify developing students' tendency or awareness of higher-order thinking by using peer assessment; take advantage of higher-order thinking to develop concepts and knowledge, and also focus on and reinforce the studies developing students' manipulation and demonstration skills at the same time. Carry out further investigations of learning behaviors or behavioral pattern analysis, and correlations or causal relationships based on mobile technology-assisted peer assessment. Reinforce the research on such outcomes as self-efficacy, satisfaction, cognitive load, and technology acceptance model/intention of use in technology-supported peer assessment. Keep strengthening the studies on col-

laboration or communication, which is important for students' future development.

Acknowledgements This study is supported in part by the Ministry of Science and Technology of the Republic of China under contract numbers MOST-105-2511-S-011 -008 -MY3 and MOST 106-2511-S-011 -005 -MY3.

References

- Carson, J., & Nelson, G. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing*, 5(1), 1–19.
- Chen, T. (2016). Technology-supported peer feedback in ESL/EFL writing classes: a research synthesis. *Computer Assisted Language Learning*, 29(2), 365–397. <https://doi.org/10.1080/09588221.2014.960942>.
- Ching, Y. H., & Hsu, Y. C. (2016). Learners' interpersonal beliefs and generated feedback in an online role-playing peer-feedback activity: An exploratory study. *International Review of Research in Open and Distributed Learning*, 17(2), 105–122.
- Cho, K., Cho, M. H., & Hacker, D. J. (2010). Self-monitoring support for learning to write. *Interactive Learning Environments*, 18(2), 101–113. <https://doi.org/10.1080/10494820802292386>.
- Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education & Training International*, 32, 175–187. <https://doi.org/10.1080/1355800950320212>.
- 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. <https://doi.org/10.3102/00346543070003287>.
- Fraile, J., Panadero, E., & Pardo, R. (2017). Co-creating rubrics: The effects on self-regulated learning, self-efficacy and performance of establishing rubrics with students. *Studies in Educational Evaluation*, 53, 69–76. <https://doi.org/10.1016/j.stueduc.2017.03.003>.
- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443–461.
- Hsia, L. H., Huang, I., & Hwang, G. J. (2016a). A web-based peer-assessment approach to improving junior high school students' performance, self-efficacy and motivation in performing arts courses. *British Journal of Educational Technology*, 47(4), 618–632. <https://doi.org/10.1111/bjet.12248>.
- Hsia, L. H., Huang, I., & Hwang, G. J. (2016b). Effects of different online peer-feedback approaches on students' performance skills, motivation and self-efficacy in a dance course. *Computers & Education*, 96, 55–71. <https://doi.org/10.1016/j.compedu.2016.02.004>.
- Hsu, T. C. (2016). Effects of a peer assessment system based on a grid-based knowledge classification approach on computer skills training. *Educational Technology & Society*, 19(4), 100–111.
- Hulsman, R. L., & van der Vloodt, J. (2015). Self-evaluation and peer-feedback of medical students' communication skills using a web-based video annotation system. Exploring content and specificity. *Patient Education and Counseling*, 98(3), 356–363. <https://doi.org/10.1016/j.pec.2014.11.007>.
- Hung, H. T., Yang, J. C., Hwang, G. J., Chu, H. C., & Wang, C. C. (2018). A scoping review of research on digital game-based language learning. *Computers & Education*, 126, 89–104.
- 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. *Etr&D-Educational Technology Research and Development*, 62(2), 129–145. <https://doi.org/10.1007/s11423-013-9320-7>.
- Hwang, G. J., & Tsai, C. C. (2011). Research trends in mobile and ubiquitous learning: A review of publications in selected journals from 2001 to 2010. *British Journal of Educational Technology*, 42(4), E65–E70. <https://doi.org/10.1111/j.1467-8535.2011.01183.x>.
- Jones, I., & Alcock, L. (2014). Peer assessment without rubrics. *Studies in Higher Education*, 39(10), 1774–1787. <https://doi.org/10.1080/03075079.2013.821974>.
- Kao, G. Y. M., Lin, S. S. J., & Sun, C. T. (2008). Beyond sharing: Engaging students in cooperative and competitive active. *Educational Technology & Society*, 11(3), 82–96.

- Lai, C. L., & Hwang, G. J. (2014). Effects of mobile learning time on students' conception of collaboration, communication, complex problem-solving, meta-cognitive awareness and creativity. *International Journal of Mobile Learning and Organisation*, 8(3), 276–291.
- Lai, C. L., & Hwang, G. J. (2015). An interactive peer-rubrics development approach to improving students' art design performance using handheld devices. *Computers & Education*, 85, 149–159. <https://doi.org/10.1016/j.compedu.2015.02.011>.
- Li, L., Liu, X. Y., & Zhou, Y. C. (2012). Give and take: A re-analysis of assessor and assessee's roles in technology-facilitated peer assessment. *British Journal of Educational Technology*, 43(3), 376–384. <https://doi.org/10.1111/j.1467-8535.2011.01180.x>.
- Li, L., Steckelberg, A. L., & Srinivasan, S. (2008). Utilizing peer interactions to promote learning through a computer-assisted peer assessment system. *Canadian Journal of Learning and Technology*, 34(2), 133–148.
- Li, H. L., Xiong, Y., Zang, X. J., Kornhaber, M. L., Lyu, Y. S., Chung, K. S., et al. (2016). Peer assessment in the digital age: A meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education*, 41(2), 245–264. <https://doi.org/10.1080/02602938.2014.999746>.
- Lin, H. C., & Hwang, G. J. (2018). Research trends of flipped classroom studies for medical courses: A review of journal publications from 2008 to 2017 based on the Technology-Enhanced Learning Model. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2018.1467462>.
- Liu, N. F., & Carless, D. (2006). Peer feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279–290. <https://doi.org/10.1080/13562510600680582>.
- Lu, J. Y., & Law, N. W. Y. (2012). Understanding collaborative learning behavior from Moodle log data. *Interactive Learning Environments*, 20(5), 451–466. <https://doi.org/10.1080/10494820.2010.529817>.
- Luo, T. (2016). Enabling microblogging-based peer feedback in face-to-face classrooms. *Innovations in Education and Teaching International*, 53(2), 156–166. <https://doi.org/10.1080/14703297.2014.995202>.
- Miller, P. J. (2003). The effect of scoring criteria specificity on peer and self-assessment. *Assessment & Evaluation in Higher Education*, 28(4), 383–394.
- Mulder, R., Baik, C., Naylor, R., & Pearce, J. (2014). How does student peer review influence perceptions, engagement and academic outcomes? A case study. *Assessment & Evaluation in Higher Education*, 39(6), 657–677. <https://doi.org/10.1080/02602938.2013.860421>.
- O'Donovan, B., Price, M., & Rust, C. (2004). Know what I mean? Enhancing student understanding of assessment standards and criteria. *Teaching in Higher Education*, 9(3), 325–335. <https://doi.org/10.1080/1356251042000216642>.
- Speyer, R., Pilz, W., Van Der Kruijs, J., & Brunings, J. W. (2011). Reliability and validity of student peer assessment in medical education: A systematic review. *Medical Teacher*, 33(11), E572–E585. <https://doi.org/10.3109/0142159x.2011.610835>.
- Stefani, L. A. J. (1994). Peer, self and tutor assessment—relative reliabilities. *Studies in Higher Education*, 19(1), 69–75. <https://doi.org/10.1080/03075079412331382153>.
- Suen, H. K. (2014). Peer assessment for massive open online courses (MOOCs). *International Review of Research in Open and Distance Learning*, 15(3), 312–327.
- Tenorio, T., Bittencourt, I. I., Isotani, S., & Silva, A. P. (2016). Does peer assessment in on-line learning environments work? A systematic review of the literature. *Computers in Human Behavior*, 64, 94–107. <https://doi.org/10.1016/j.chb.2016.06.020>.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249–276. <https://doi.org/10.2307/1170598>.
- van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H. (2009). Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. *Educational Research Review*, 4(1), 41–54. <https://doi.org/10.1016/j.edurev.2008.11.002>.
- van Popta, E., Kral, M., Camp, G., Martens, R. L., & Simons, P. R. J. (2017). Exploring the value of peer feedback in online learning for the provider. *Educational Research Review*, 20, 24–34. <https://doi.org/10.1016/j.edurev.2016.10.003>.
- Vanderhoven, E., Raes, A., Montrioux, H., Rotsaert, T., & Schellens, T. (2015). What if pupils can assess their peers anonymously? A quasi-experimental study. *Computers & Education*, 81, 123–132. <https://doi.org/10.1016/j.compedu.2014.10.001>.
- Wang, H. Y., Liu, G. Z., & Hwang, G. J. (2017). Integrating socio-cultural contexts and location-based systems for ubiquitous language learning in museums: A state of the art review of 2009–2014. *British Journal of Educational Technology*, 48(2), 653–671.

- Yu, F. Y. (2011). Multiple peer-assessment modes to augment online student question-generation processes. *Computers & Education*, 56(2), 484–494. <https://doi.org/10.1016/j.compedu.2010.08.025>.
- 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. <https://doi.org/10.1111/j.1467-8535.2008.00905.x>.
- Yu, F. Y., & Sung, S. (2016). A mixed methods approach to the assessor's targeting behavior during online peer assessment: Effects of anonymity and underlying reasons. *Interactive Learning Environments*, 24(7), 1674–1691. <https://doi.org/10.1080/10494820.2015.1041405>.
- Yu, F. Y., & Wu, C. P. (2011). Different identity revelation modes in an online peer-assessment learning environment: Effects on perceptions toward assessors, classroom climate and learning activities. *Computers & Education*, 57(3), 2167–2177. <https://doi.org/10.1016/j.compedu.2011.05.012>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Qing-Ke Fu is an Assistant Professor in the Department of Educational technology of the College of Teacher Education, Huzhou University. His research interests include mobile learning, digital game-based learning, and technology-assisted language learning.

Chi-Jen Lin is an Assistant Professor at the College of Liberal Arts and Social Sciences, National Taiwan University of Science and Technology. Her research interests include flipped learning, digital game-based learning, and language learning.

Gwo-Jen Hwang is a Chair professor at the Graduate Institute of Digital Learning and Education, National Taiwan University of Science and Technology. His research interests include mobile learning, digital game-based learning, flipped classroom, and AI in education.