

Student attitudes that predict participation in peer assessment

Yan Zou, Christian Dieter Schunn, Yanqing Wang & Fuhui Zhang

To cite this article: Yan Zou, Christian Dieter Schunn, Yanqing Wang & Fuhui Zhang (2017): Student attitudes that predict participation in peer assessment, *Assessment & Evaluation in Higher Education*, DOI: [10.1080/02602938.2017.1409872](https://doi.org/10.1080/02602938.2017.1409872)

To link to this article: <https://doi.org/10.1080/02602938.2017.1409872>



Published online: 04 Dec 2017.



Submit your article to this journal [↗](#)




View related articles [↗](#)



View Crossmark data [↗](#)



Student attitudes that predict participation in peer assessment

Yan Zou^{a,b}, Christian Dieter Schunn^b, Yanqing Wang^{b,c}  and Fuhui Zhang^{b,d}

^aSchool of Foreign Languages, Beihang University, Beijing, China; ^bLearning Research & Development Center, University of Pittsburgh, Pittsburgh, USA; ^cSchool of Management, Harbin Institute of Technology, Harbin, China; ^dSchool of Foreign Languages, Northeast Normal University, Changchun, China

ABSTRACT

Peer assessment has been widely applied to actively engage students in learning to write. However, sometimes students resist peer assessment. This study explores reviewers' attitudes and other underlying factors that influence students' participation in online peer assessment. Participants were 234 Chinese undergraduates from two different academic backgrounds: engineering majors ($n = 168$) and English majors ($n = 66$). Gender, academic background and prior experience with peer assessment were all related to participation levels. Moreover, factor analyses revealed three attitudinal factors: (1) *positive attitude* (a general endorsement of the benefits of peer assessment), (2) *interpersonal negative* (concerns about the negative effects on interpersonal relationships), and (3) *procedural negative* (doubts about the procedural rationality of peer assessment). Among the attitudinal factors, *procedural negative* was negatively associated with participation, as expected. Interestingly, *interpersonal negative* was associated with greater participation, and *positive attitude* was associated with lower participation, in part because students worked hard on each review rather than doing many reviews superficially. Implications for instruction are discussed.

KEYWORDS

Peer assessment; student participation; student attitudes; EFL writing

Introduction

As modern education increasingly focuses on self-directed and collaborative learning (Voogt et al. 2013), teachers of English as a foreign language (EFL) are turning to more collaborative activities that actively engage students in learning to write (Reid 1993; Watanabe and Swain 2007; Yang 2014). One popular example is peer assessment, a collaborative learning activity that has been widely practiced in EFL writing classes over the last two decades, proving to be reliable and valid in many learning contexts (Cho, Schunn, and Wilson 2006; Cheng, Hou, and Wu 2014; Li et al. 2016). Peer assessment, also called peer evaluation, peer feedback or peer review, is defined in Topping's (1998) highly cited review as 'an arrangement in which individuals consider and specify the amount, level, value, quality or success of the products or outcomes of learning of peers of similar status' (250).

Research has consistently found that peer assessment is beneficial to learners as both an assessor and assessee (Berg 1999; Paulus 1999; Patchan, Schunn, and Clark 2011), leading to positive impacts on students' performance in subsequent examinations (Jhangiani 2016). However, student attitudes towards peer assessment are mixed. Some studies found that students held positive attitudes towards and were overall receptive to peer assessment (Gatfield 1999; Roskams 1999; Liu and Yuan 2003; Collimore, Paré,

and Joordens 2015; Schunn, Godley, and DeMartino 2016). However, other studies found a range of negative attitudes towards peer assessment implementation (Roskams 1999; Lin, Liu, and Yuan 2002; Wen and Tsai 2006; Kaufman and Schunn 2011; Praver, Rouault, and Eidswick 2011; McGarr and Clifford 2013; Collimore, Paré, and Joordens 2015). Negative attitudes towards peer assessment include the perceived reliability of student ratings (Liu and Carless 2006), perceived expertise of assessors (Cheng and Warren 2005; Liu and Carless 2006; Kaufman and Schunn 2011), power relations among students (Brown and Knight 1994; Liu and Carless 2006), time available to do the activity (Liu and Carless 2006) and the appropriateness of grades based on peer assessment (Wen and Tsai 2006; Kaufman and Schunn 2011).

Attitudes towards peer assessment may vary by the context or characteristics of the learners engaged in peer assessment; such variation may be particularly large given the wide range of contexts in which peer assessment is implemented. One regularly studied factor influencing attitudes is gender. Researchers sometimes found more positive attitudes towards peer assessment in male students (Wen and Tsai 2006), and sometimes more positive attitudes in female students (Warrington, Younger, and Williams 2000), and sometimes no difference (Cheng, Hou, and Wu 2014; Collimore, Paré, and Joordens 2015). These inconsistent gender effects may be explained by some other underlying factor (e.g. major) that is correlated with gender to different degrees across learning contexts. For example, compared to English majors, Chinese engineering students are less motivated in English learning (Tang 2012), and engineering students generally think reading and listening are more important than writing and are less willing to spend effort in writing (Wang, Xuan, and Chen 2006; Gu 2007). Other researchers found that there was a significant association between students' major and perceived ease of peer assessment (e.g. students majoring in business, economics, and law were more likely to report that peer assessment was easy to implement than students in humanities) (Praver, Rouault, and Eidswick 2011). In addition, some majors may have more prior peer assessment experiences, which have been previously associated with having less negative attitudes towards peer assessment (Wen and Tsai 2006).

Despite the extensive examination of student attitudes towards peer assessment, only a few case studies examined the consequences of these attitudes, including the most obvious consequence, levels of participation in peer assessment. As with any kind of in or out-of-class activity, not all students complete their required peer review tasks (Patchan, Schunn, and Correnti 2016). If student's participation is especially low, some students may fail to receive any feedback, and others will fail to receive the learning benefits of doing peer assessment. To foreshadow our study context, in one class, fewer than 50% of the students completed their peer review activities although these same students completed other class assignments at much higher rates.

Prior studies have attempted to reveal the underlying factors that impact peer assessment participation. Yu and Lee (2015) framed the issue from an activity theory perspective and concluded that students' participation in peer feedback was 'driven and defined by their motives' (589). Cheng, Hou, and Wu (2014) found that students with more positive attitudes towards peer assessment tended to participate more in peer assessment, defined as the number of peer comments the students provided. However, their study did not examine what kind of attitudes drove participation. In general, little is known about which peer-related attitudes may lead to differences in students' participation in peer assessment. For example, are perceptions about the benefits of peer assessment more important or are concerns about peer assessment more important?

Based on these critical gaps in the literature on peer assessment, this study explores the underlying factors that influence students' participation in peer assessment in writing instruction:

(RQ1) What kinds of attitudes about peer assessment drive students' participation in peer assessment? To address issues of generalizability, data is collected from course contexts with very different attitudes towards writing: English majors who are likely to place considerable importance on writing instruction and engineering majors who are likely to have less positive assessments about the importance of writing instruction. We examine three additional demographic factors that were implicated in prior research on peer assessment and may be correlated with attitude: gender, academic background and prior experience with peer assessment. (RQ2) We then explore which of these demographic variables is directly associated with peer assessment participation, and (RQ3) whether attitudinal differences explain the effects of these variables on peer assessment participation.

Methods

Participants

In total, 234 students from two strong, research-oriented universities in two large cities in northeast China, one in Beijing (168 students), the other in Changchun (66 students), participated in the peer assessment activities. The Beijing students were engineering or engineering science majors (henceforth called engineering majors for simplicity), enrolled in an introductory English course called *English Reading and Writing*. All Changchun students were English majors, enrolled in a more advanced English course called *Thesis Writing for EFL Students*. These courses were selected because the instructors used a common online peer review system, SWoRD, and we did not wish to confound participation rates with differences in the ease of use of peer assessment system.

The final sample includes 105 students (59% engineering majors; 54% females) who completed the survey measures. Similar to other participation metrics, the survey response rate was higher in the English majors (59%) than in the engineering majors (37%); both rates are higher than typical response rates (30%) for anonymous student surveys (Nulty 2008). Although gender and major were correlated, it was not perfect: 34% of engineering majors were female, and 87% of English majors were female.

Materials

Scaffolded writing and reviewing in the discipline

SWoRD (Scaffolded writing and reviewing in the discipline) is an online peer assessment platform (Cho and Schunn 2007; Schunn 2016), with many features that are common to online peer assessment systems, such as Calibrated Peer Review (Russell 2004), Expertiza (Gehring 2010) and EduPCR (Wang et al. 2016). Drafts are distributed randomly and anonymously to (usually five) students for review. Reviewers are required to grade each essay on seven-point rating scales (with anchors that are specific to each reviewing scale) and write specific feedback for each writing dimension as specified by their instructor to match the focus of the given writing assignment.

After the reviewers have submitted the ratings and comments, the authors receive their peers' feedback and a mean rating on each writing dimension. With the received peer comments and ratings, students produce a second draft, which can then be peer reviewed again. For our study, students in both classes engaged in this writing-reviewing-revising-reviewing process once for each paper they wrote.

With SWoRD, after receiving each round of reviews and ratings, the authors also rate the helpfulness of the reviews they have received on a five-point rating scale: this step is called back-evaluation. Reviewers receive a helpfulness grade based upon the mean helpfulness ratings they receive. This step provides accountability for higher quality comments.

Assignments

Students in both courses were required to complete two essays, initially submitted for peer assessment via SWoRD and then to the instructor for grading (a few students failed to submit essays on SWoRD and the instructors graded those papers manually). The details of the writing assignments were matched to the performance level of the students. The engineering students wrote one autobiography and one summary of a novel read in class. The English students wrote a literature review and methodology section for their senior thesis paper. Because engineering students had high levels of resistance towards English writing (Wang, Xuan, and Chen 2006; Gu 2007; Tang 2012), they ended up only writing one draft for each assignment with peer review and back-evaluation; by contrast, the English students wrote two drafts for both assignments with peer review and back-evaluation. We therefore focus on the first draft.

Measures

Primary measures were derived from a survey that was completed anonymously to encourage honest responses about the sensitive topic of work non-completion. Note that all questions were asked

in Chinese to ensure understanding, but we present English translations (<https://wenku.baidu.com/view/96dd4287b1717fd5360cba1aa8114431b90d8e1b>).

Participation in peer assessment

The first section of the survey asked about students' participation in peer assessment. We investigated the degree of students' participation in peer assessment using the survey data instead of the data recorded in SWoRD in order to link it with the anonymous attitudinal data. Participation involved the following two variables:

- *reviews (how many papers did you review?)*: the number of papers reviewed by each reviewer. Although each student was required to complete five reviews, some students reviewed fewer, with the observed number of reviews ranging from 0 to 5. In this study, 20 out of 105 respondents did not do any reviews. We treated this variable as a count variable;
- *back-evaluations (did you do back-evaluation?)*: whether the author completed back-evaluation. This measure was implemented as a binary completed-or-not variable because this step was relatively simple to complete when started and thus students tended to do this in an all-or-none fashion.

Attitudes towards peer assessment

To understand what students thought about peer assessment, we developed a survey form based upon prior research on student attitudes towards peer assessment (Cheng and Warren 2005; Wen, Tsai, and Chang 2006; Kaufman and Schunn 2011; McGarr and Clifford 2013; Planas Lladó et al. 2014). Students indicated their degree of agreement on a Likert scale of 1 (strongly disagree) to 5 (strongly agree) with statements about the usefulness of peers' and one's own feedback, the positive and negative nature of peers' and one's own feedback, and the fairness of peer grades etc.

An Exploratory Factor Analysis (EFA) using varimax rotation was applied to the attitude data. We want to find out the underlying attitudinal factors through the 25 survey responses. EFA is a family of statistical methods used to discover latent factors underlying a relatively large set of individual response variables, especially in survey-based research when researchers have no strong a priori hypothesis about how many separable factors there will be. Since individual response items may involve a combination of underlying factors, a rotation method is typically used in the EFA to find the factors. The result of EFA will present a factor loading (varying from -1 to 1) for each survey question, which reflects the strength of the association of the item with each factor. A loading of 0.3 is a common threshold for deciding whether an item has a meaningful contribution to the factor.

The EFA on the attitude data revealed that the statements could be clustered into three underlying factors: (1) *positive attitude* (a general endorsement of benefits of peer assessment), (2) *interpersonal negative* (concerns about the negative effects on interpersonal relationships), and (3) *procedural negative* (doubts about the procedural rationality of peer assessment). The *procedural negative* involved a number of different doubts about whether peer assessment could sensibly be done by students, e.g. cognitive problems (not knowing what to do or feeling the assessment is the responsibility of teachers), lack of training in peer assessment, lack of confidence in their own and their peers' ability as assessors.

Table 1 lists the keywords for specific survey statements (translated from Chinese) associated with each factor and factor loadings (above 0.3) for each survey statement from the EFA. A few statements have significant loading on two factors. We clustered them into the factor with the larger value, with one exception that had similarly large loadings on two factors: 'It's not students' duty to review the writing'. We manually clustered it to *procedural negative* instead of *interpersonal negative* based on theoretical match.

A score for each factor was created by computing a mean across included items. The *positive attitude* factor was not significantly correlated with the *interpersonal negative* ($r = 0.05$), but negatively correlated with *procedural negative* ($r = -0.29$, $p < 0.01$). The two negative factors had a relatively high correlation ($r = 0.52$, $p < 0.001$), but were still sufficiently separated to allow for treatment as different factors.

Table 1. Survey questions on attitudes towards peer assessment along with factor analysis loadings.

Attitudinal factors	Attitudes towards peer assessment	Factor loadings		
		1	2	3
Positive attitude	Identify the weaknesses	0.84		
	Understand assignment	0.77		
	Improve my linguistic communication	0.81		
	Enhance my motivation	0.86		
	More effective than teacher feedback	0.69		
	Take peer assessment seriously	0.70		
	I'm competent for peer review	0.62		
	Clear rubrics essential	0.52	0.39	
	Student making rubrics	0.56	0.34	
	Quality of peer assessment will be improved if graded	0.44		
Interpersonal negative	Doubt fairness		0.65	
	Friendship is impaired		0.68	
	Reviews be fairer if anonymous		0.51	
	I have the right to complain		0.62	
	A lower helpfulness grade for a lower document grade		0.68	
Procedural negative	I give a higher grade to friends		0.60	
	Makes me nervous			0.44
	Takes too much time		0.52	0.54
	It is not my duty		0.60	0.55
	I'm not clear what to do		0.35	0.68
	Competence influence quality of peer assessment			0.34
	Prefer teacher reviews			0.72
	Prefer face-to-face reviews			0.64
	Five essays to review is too much			0.61
Not fair to grade reviews			0.64	

Demographics and prior experiences

To increase the sense of anonymity, the survey did not ask about the university students attended. Instead, major was inferred from the IP address of the completed online survey based on proximity to Beijing or Changchun; since the cities were more than 1000 km apart and in different provinces, which insured highly distinct IP addresses, this approach allowed for easy separation.

The survey also asked about students' prior experience with peer review and gender in order to see whether major, gender or prior experiences were the underlying causes of differential attitudes and behaviours. For prior experience, students were provided with four choices: (A) *Yes, online anonymous peer assessment* (45%); (B) *Yes, online real identity peer assessment* (7%); (C) *Yes, face-to-face in-class peer assessment* (24%); (D) *Never* (25%). For analysis, responses were converted into a binary yes/no (25% no peer assessment experience at all, 75% with some peer assessment experience).

Procedure

Upon completion of the online class writing and reviewing assignments, students received a link to the online survey on their attitude towards SWoRD and peer assessment.

Analysis approach

To predict binary outcomes (e.g. doing back-evaluation or not), we use chi-square (χ^2) for single binary predictors (e.g. gender, major and prior experience), and logistic regression for continuous predictors and for analyses of combining predictors (e.g. attitudinal factors). To predict continuous variables which were approximately normally distributed (e.g. attitudes), we use *t*-tests (with corrections for unequal variance, as appropriate) and multi-linear regression. To predict continuous variables that were not normally distributed or count variables (e.g. number of reviews completed), we used Spearman correlation and Poisson analysis (which is used to predict a dependent variable that consists of count data).

Results

Attitudinal differences by background

Table 2 presents the relationship between demographic variables and attitudes towards peer assessment. We first treated the three background factors as independent predictors of attitude and did several independent *t*-tests to check the relations separately. We found that males had significantly lower concerns about *interpersonal negative* ($p = 0.047$). Major showed similar patterns, with engineering majors having significantly lower concerns about interpersonal issues ($p < 0.01$). Prior peer assessment experience was not significantly associated with any of the three attitudinal factors.

However, only the effects of major remained statistically significant on *interpersonal negative* ($p < 0.05$) when major, gender and prior experience were included into an overall model (see Table 3): English majors tended to worry more about whether peer assessment activities would have some negative impact on the friendship with peers. In other words, as predicted, gender was only associated with attitudinal differences by virtue of being correlated with major.

Peer assessment participation by background

Table 2 also presents the simple associations between each demographic variable and participation in peer assessment. Female students reviewed more papers ($p < 0.001$) and were more likely to complete the back-evaluation task ($p < 0.001$). Similarly, students with prior peer assessment experience reviewed significantly more papers ($p = 0.002$) and were more likely to do back-evaluation ($p < 0.01$). Likewise, English majors reviewed more papers ($p < 0.001$) and were also more likely to complete the back-evaluation task ($p < 0.001$). Rather than explore their unique contributions in a multiple regression, we instead focus on the underlying attitudinal explanations.

Attitude and peer assessment participation

We first considered the three attitudinal factors as independent predictors of peer assessment participation. Based upon a Spearman correlation, none of the factors was significant predictors of number of papers reviewed, with *procedural negative* marginally correlated. Then we considered the three factors in an overall model and did a Poisson analysis. The result revealed that all three factors were statistically significant predictors of number of papers reviewed. Interestingly, *positive attitude* and *procedural negative* were negative predictors ($\text{Exp}(B) < 1$), and *interpersonal negative* was a positive predictor ($\text{Exp}(B) > 1$) (see Table 4).

Similarly, when treated independently, none of the three attitudinal factors was a significant predictor of back-evaluation, with *interpersonal negative* marginally correlated. However, based upon a multiple

Table 2. Association between demographic variables and participation and attitudes.

Predictor	Positive attitude	Interpersonal negative	Procedural negative	Reviews	Back-evaluations
	(mean and SE)	(mean and SE)	(mean and SE)	(mean and SE)	
Gender	$t = 1.7, p = 0.09$	$t = -1.98, p = 0.047$	$t = 0.04, p = 0.97$	$r = 0.52, p < 0.001$	$\chi^2 = 18.8, p < 0.001$
Female	3.4 (0.1)	3.2 (0.1)	3.2 (0.1)	4.1 (0.2)	75%
Male	3.6 (0.1)	3.0 (0.1)	3.2 (0.1)	1.9 (0.3)	33%
Major	$t = 1.75, p = 0.08$	$t = -3.14, p < 0.01$	$t = 1.0, p = 0.3$	$r = 0.53, p < 0.001$	$\chi^2 = 36.4, p < 0.001$
English	3.4 (0.1)	3.3 (0.1)	3.1 (0.1)	4.6 (0.2)	95%
Engineering	3.6 (0.1)	2.9 (0.1)	3.2 (0.1)	2.2 (0.3)	34%
Prior peer assessment experience	$t = -1.0, p = 0.3$	$t = -0.6, p = 0.5$	$t = -0.1, p = 0.9$	$r = 0.3, p = 0.002$	$\chi^2 = 6.5, p < 0.01$
No	3.4 (0.2)	3.0 (0.1)	3.2 (0.1)	1.9 (0.5)	35%
Yes	3.5 (0.1)	3.1 (0.1)	3.2 (0.1)	3.5 (0.2)	63%

Table 3. Partial correlations of background and prior experience with attitudinal factors.

	Positive attitude	Interpersonal negative	Procedural negative
Gender (male = 1, female = 2)	-0.17	0.05	0.06
Major (engineer = 1, English = 2)	-0.09	0.27*	-0.13
Prior peer assessment experience (no = 0, yes = 1)	0.08	0.05	0.02

* $p < 0.05$.**Table 4.** Attitudinal predictors of number of papers reviewed and back-evaluation completion.

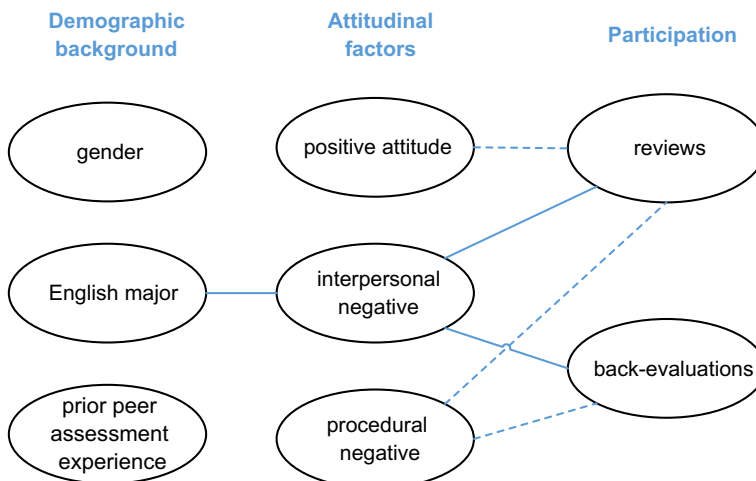
Predictor	Reviews (Beta)		Back-evaluations (Odds-ratio, Exp(B))	
	Simple regression (Spearman)	Multiple regression (Exp(B))	Simple regression	Multiple regression
Positive attitude	-0.11	0.76**	1.1	0.7
Interpersonal negative	0.06	1.37**	1.7 [†]	3.47**
Procedural negative	-0.19 [†]	0.66**	0.6	0.28**

[†] $p < 0.1$.** $p < 0.01$.

logistic regression, both *interpersonal negative* and *procedural negative* were significant predictors of back-evaluation. *Interpersonal negative* had a positive effect ($\text{Exp}(B) > 1$) on back-evaluation: the more students are worried about the negative impact of peer assessment on interpersonal relationship, the more they are likely to do back-evaluation, which agrees with the result of *reviews*. *Procedural negative* had a negative effect ($\text{Exp}(B) < 1$): the more students doubt the rationality of peer assessment, the less likely they are to do back-evaluation.

Together, the significant connections between the variables across layers can be summarised in Figure 1, with solid lines for positive effects and dashed lines for negative relationships.

From Figure 1, we can see that English majors worried more about interpersonal relationship than engineering majors, which may lead them to participate more in peer assessment. They might think that more effort in peer assessment is a representation of their attachment to friendship. As a separate relationship not related to demographics or prior experience, the more students doubted the procedural rationality of peer assessment, the fewer papers they would review and the less likely they were to submit back-evaluation. More surprisingly, *positive attitude* appeared to have a significant negative effect on completing *reviews* but no effect on completing *back-evaluations*.

**Figure 1.** Connections among three layers: demographic background, attitudinal factors and participation.

General discussion

The aim of this study was to investigate the impact of attitudes upon participation in peer assessment with English and engineering majors. There were several important findings.

Positive attitude has a significant negative impact on one predictor of peer assessment participation: the number of papers reviewed. People generally perform better when they have more positive conceptions regarding a task because their positive conceptions make them feel more competent at the task (Brown et al. 2009). This would suggest that students who have a more positive conception of peer assessment would tend to participate more in peer assessment. However, we found the opposite pattern.

To find a possible explanation, we considered another factor of participation which was not an initial focus of our research (i.e. not included in the survey): comment length. We hypothesised that the reviewers with more positive attitudes might tend to write more comments for each paper, and then get tired (or run out of time). We obtained reviewing data from SWoRD and computed the mean comment length (in characters) reviewers wrote for each paper: engineering students ($M_{\text{assignment1}} = 349$, $SD_{\text{assignment1}} = 254$, $M_{\text{assignment2}} = 285$, $SD_{\text{assignment2}} = 205$), English students ($M_{\text{assignment1}} = 856$, $SD_{\text{assignment1}} = 644$, $M_{\text{assignment2}} = 323$, $SD_{\text{assignment2}} = 245$). Then we computed a correlation between the average comment length reviewers wrote for each paper with the number of papers reviewed. We focused on the case of the engineering majors who were especially reticent to do reviewing and provided a large enough N to conduct this correlation analysis with strong statistical power. There was a strong negative correlation for the first assignment ($r = -0.50^{**}$) and a weak negative correlation for the second assignment ($r = -0.09$). The negative correlation, especially for the first assignment, provides support for a 'getting tired' hypothesis that could explain the negative relationship between *positive attitude* and *reviews*. This explanation is also consistent with the lack of a negative effect on back-evaluation: those are much shorter/faster to complete and therefore less likely to involve students getting tired from giving in-depth feedback.

Concerns about the negative effects of peer assessment on interpersonal relationships may encourage students' peer assessment participation. Peer assessment is an inherently social and collaborative learning process in which students, by assessing each other, learn with and from each other as peers. Although reviewers' perception and attitudinal factors have been extensively studied in peer assessment research and some students have reported that peer assessment had hindered their relationships with peers (Planas Lladó et al. 2014), few studies focus on the relationship between reviewers' interpersonal concern and peer assessment participation. In the context of team research, Edmondson (1999) argued that psychological safety, 'a sense of confidence that the team will not embarrass, reject, or punish someone for speaking up' (354), facilitated learning behaviours. Similarly, in peer assessment, if learners have fewer concerns about others' reactions to feedback, they will contribute more to the learning process, i.e. have higher levels of participation in peer assessment. However, we found the opposite relationship. We hypothesised that the students who were more worried about hurting interpersonal relationships were the ones generally placing more value on supporting others and therefore were more likely to help others, such as providing peer feedback to other students. Future research should examine whether conscience or altruism is a critical underlying factor of peer assessment participation, rather than having this interpersonal concern about peer assessment in particular.

Concerns about the procedural rationality of peer assessment may discourage students' peer assessment participation. It is a commonly held belief that teachers are the authority in class and the proper assessors of students' learning performance, especially in the context of Chinese culture. In our survey, for the statement 'I prefer to get feedback only from my teacher', more than 31% respondents agreed or strongly agreed, with another 32% undecided. Consequently, students tend to be reluctant to assume the responsibility to assess and thus participate less often in peer assessment. Furthermore, doubting the validity of peer assessment may be due to a lack of confidence in both their own grading and peers' grading, and reflect their concern of the fairness of peer assessment (Kaufman and Schunn 2011). In our survey, for the statement 'I doubt the fairness of peer review', more than 34% respondents agreed or strongly agreed, with another 27% undecided. For statement 'I think I am competent enough to review

my peer's paper', more than 26% respondents disagreed or strongly disagreed, with 38% undecided. With these concerns, students may lack the motivation to engage fully in peer assessment process.

Gender was indirectly associated with differences in participation. In addition to examining attitudinal factors, our study also investigated the impact of demographic factors on peer assessment participation. Although the literature abounds with studies on gender and peer assessment, most studies focus on the impact of gender on rating, reliability, group work process, achievement and attitude (Wen and Tsai 2006; Sasmaz Oren 2012; Takeda and Homberg 2014). Few studies explored the effect of gender on peer assessment participation. Although our study found that female students participated significantly more than male students did, the discipline effect was much larger than the gender effect, and much of the gender effect was likely due to discipline.

Academic background showed large peer assessment participation differences. As expected, English students participated significantly more than did engineering students in our study. English students are generally more motivated to study English and hence are likely to devote more effort to all aspects of English learning, including English writing. Lower motivation to write in engineering students appeared to translate into a low number of papers reviewed and a lower likelihood of doing even the simple back-evaluation step. Our study thus reinforced the importance of motivation towards language learning. Indeed, motivation generally plays an important self-regulatory role in learning processes (Pintrich 1999) and appears to encourage learners to participate more in the learning activities like peer assessment. However, our analyses also suggested that differences in beliefs about the benefits of peer assessment was only a small part of the academic background difference; instead differences in concerns about interpersonal issues was the largest difference between the majors.

Prior peer assessment experience is also a critical factor in peer assessment participation. Research has often shown that prior knowledge is an important factor in learning. For example, learners with prior experience/knowledge are able to better deal with strong working memory capacity limitations with the help of an organised knowledge base (Ericsson and Kintsch 1995). Furthermore, classroom routines have long been known to be important for learner participation (Leinhardt, Weidman, and Hammond 1987). Accordingly, resistance from lack of prior experience has been a common concern for instructors implementing peer assessment. Our study further supports the perspective: while many of the students already had prior experience with peer assessment, lack of such experiences did lead to different degrees of participation in peer assessment. If more instructors include peer assessment activities in their classes, students' acceptance and performance in peer assessment will gradually improve.

Implications for instruction

Gardner and Lambert (1972) maintained that attitude and motivation were central to successful language learning, and that motivation to learn a foreign language was affected by the attitude and readiness to be identified with the foreign language community. Accordingly, instructors may commonly focus on the benefits of peer assessment to boost learning motivation. However, our study does not support this approach. Instead, we found that the strongest predictors involved concerns about peer assessment. Instead of highlighting the benefits of peer assessment, teachers also need to focus more on students' normative beliefs about the interdependence of students upon each other and address the concerns about procedural rationality of peer assessment. This could be done by helping students develop confidence in their own and their peers' capability as reviewers through specific training programmes. How to move (at least a portion of) the responsibility for assessment from the teacher to the students should be the focus of future research.

Furthermore, prior research has highlighted that students' negative attitude could discourage use of peer assessment (Cohen 1988; McNeil 1988). However, our study finds that some negative attitudes, e.g. *interpersonal negative*, may also enhance the depth of participation in peer assessment. This finding suggests that there should be more discussion of how to enhance peer assessment participation by arousing students' attachment to friendship among peers, which may in turn increase voluntary participation in peer assessment.

Limitations and future work

The analyses presented here were fundamentally correlational in nature, and strong claims about causality cannot be made. However, through including a number of demographic variables that are likely to be relevant, we ruled out a number of important confounds. Furthermore, attitudes are sometimes difficult to manipulate experimentally. Nonetheless, it will be important to conduct experimental studies in the future to test the causal nature of the relationships observed here.

Most of the participants in the survey are those who submitted essays on SWoRD. Those who did not submit an essay were not likely to take the survey since they had no peer assessment experience, which made the sampling biased towards the volunteer sample. Future research should better include those who have no online peer review experience to get a more comprehensive picture. Furthermore, we examined only two indicators – number of papers reviewed and back-evaluation completion – to measure student participation in peer assessment. Future studies could take more aspects into consideration, such as the length of comments, the number of comments per paper, comment word density and the sentence variety of comments.

This study only examined two different majors, which differ in many aspects. In the future, the attitudes and participation in peer assessment of other majors should also be examined, particularly teasing apart the ability to write in English from the importance placed on English for career goals and the general interest in English per se. Furthermore, we examined students at relatively strong research-oriented institutions. Less academically oriented students may show different patterns.

The current research has highlighted that a number of demographic features of potential relevance to peer review practices naturally tend to be confounded. For example, academic major was strongly correlated with gender in our sample, and will tend to be confounded in general. Such natural confounds calls for research from larger samples that allow for more systematic decoupling of underlying factors. In addition, samples drawn at different developmental time points can help untangle different effects of structural factors like gender and experiential factors like academic majors.

Finally, it is important to replicate and further investigate the causes of the negative relationship between *positive attitude* and student participation in peer assessment. We provided some support for a getting-tired-effect, but this should be examined in a more direct fashion.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was partially supported by Fundamental Research Funds for the Central Universities [grant number YWF-17-SK-G-22]; National Natural Science Foundation of China [grant number 71573065]; and Online Education Research Foundation of China's MOE Research Center for Online Education [grant number 2016YB130].

Notes on contributors

Yan Zou is a lecturer in the School of Foreign Languages, Beihang University. She has been teaching English as a foreign language and doing research in this field for more than ten years. Her research covers second language acquisition and peer assessment in academic writing. She has also written and published some articles on learning performance, classroom assessment and teaching practice. She has visited LRDC for a year.

Christian Dieter Schunn is a professor and senior cognitive scientist working with Learning Research & Development Center (LRDC) at the University of Pittsburgh. His research interest extends to a wide range of cognitive studies involving STEM reasoning and learning, web-based peer interaction and instruction, neuroscience of complex learning and engagement and learning. He is the founder of an online peer review system (SWoRD), which is widely used in USA, China, and some European countries.

Yanqing Wang is an associate professor working with the School of Management, Harbin Institute of Technology, China. He is the founder of an information system (EduPCR), which is dedicated to peer code review in educational contexts. His recent research interests include peer assessment, peer code review, reviewer assignment problem, gamification of peer learning and innovation & entrepreneurship. He has visited LRDC for a year.

Fuhui Zhang is an associate professor working with the School of Foreign Languages, Northeast Normal University in China. She has been teaching and researching English writing and peer review over eight years. Her research interest lies in how writing and peer review help improve students' critical thinking skills. Her PhD thesis is on the interplay of peer review and writing goals in revision. She has visited LRDC for a year.

ORCID

Yanqing Wang  <http://orcid.org/0000-0002-5961-0482>

References

- Berg, E. C. 1999. "The Effects of Trained Peer Response on ESL Students' Revision Types and Writing Quality." *Journal of Second Language Writing* 8 (3): 215–241.
- Brown, G. T. L., S. E. Irving, E. R. Peterson, and G. H. F. Hirschfeld. 2009. "Use of Interactive–Informal Assessment Practices: New Zealand Secondary Students' Conceptions of Assessment." *Learning and Instruction* 19 (2): 97–111.
- Brown, S., and P. Knight. 1994. *Assessing Learners in Higher Education*. London: Kogan.
- Cheng, Kun-Hung, H. T. Hou, and S. Y. Wu. 2014. "Exploring Students' Emotional Responses and Participation in an Online Peer Assessment Activity: A Case Study." *Interactive Learning Environments* 22 (3): 271–287.
- Cheng, W., and M. Warren. 2005. "Peer Assessment of Language Proficiency." *Language Testing* 22 (1): 93–121.
- Cho, K., and C. D. Schunn. 2007. "Scaffolded Writing and Rewriting in the Discipline: A Web-based Reciprocal Peer Review System." *Computers & Education* 48 (3): 409–426.
- Cho, K., C. D. Schunn, and R. W. Wilson. 2006. "Validity and Reliability of Scaffolded Peer Assessment of Writing from Instructor and Student Perspectives." *Journal of Educational Psychology* 98 (4): 891–901.
- Cohen, D. K. 1988. "Educational Technology and School Organization." In *Technology in Education: Looking toward 2020*, edited by R. S. Nickerson and P. P. Zodhiates, 231–264. Mahwah, NJ: Lawrence Erlbaum Associates.
- Collimore, L., D. E. Paré, and S. Joordens. 2015. "SWDYT: So What Do You Think? Canadian Students' Attitudes about PeerScholar, an Online Peer-Assessment Tool." *Learning Environments Research* 18 (1): 33–45.
- Edmondson, A. C. 1999. "Psychological Safety and Learning Behavior in Work Teams." *Administrative Science Quarterly* 44: 350–383.
- Ericsson, K. A., and W. Kintsch. 1995. "Long-Term Working Memory." *Psychological Review* 102: 211–245.
- Gardner, R. C., and W. E. Lambert. 1972. *Attitudes and Motivation: Second Language Learning*. New York: Newbury House.
- Garfield, T. 1999. "Examining Student Satisfaction with Group Projects and Peer Assessment." *Assessment and Evaluation in Higher Education* 24 (4): 365–377.
- Gehring, E. F. 2010. "Expertiza: Managing Feedback in Collaborative Learning." In *Monitoring and Assessment in Online Collaborative Environments: Emergent Computational Technologies for E-Learning Support*, edited by A. A. Juan, T. Daradoumis, F. Xhafa, S. Caballe, and J. Faulin, 75–96. Hershey, Pennsylvania: IGI Global.
- Gu, Z. W. 2007. "An Analysis of Chinese Non-English Majors' English Learning Beliefs." *Foreign Languages Education* 28 (5): 49–54.
- Jhangiani, R. S. 2016. "The Impact of Participating in a Peer Assessment Activity on Subsequent Academic Performance." *Teaching of Psychology* 43 (3): 180–186.
- Kaufman, J. H., and C. D. Schunn. 2011. "Students Perceptions about Peer Assessment for Writing: Their Origin and Impact on Revision Work." *Instructional Science* 39 (3): 387–406.
- Leinhardt, G., C. Weidman, and K. M. Hammond. 1987. "Introduction and Integration of Classroom Routines by Expert Teachers." *Curriculum Inquiry* 17 (2): 135–176.
- Li, H., Y. Xiong, X. Zang, M. L. Kornhaber, Y. Lyu, K. S. Chung, and H. K. Suen. 2016. "Peer Assessment in the Digital Age: A Meta-Analysis Comparing Peer and Teacher Ratings." *Assessment & Evaluation in Higher Education* 41 (2): 245–264.
- Lin, S. S., E. Z. F. Liu, and S. M. Yuan. 2002. "Student Attitude toward Networked Peer Assessment: Case Studies of Undergraduate Students and Senior High School Students." *International Journal of Instructional Media* 29 (2): 241.
- Liu, N. F., and D. Carless. 2006. "Peer Feedback: The Learning Element of Peer Assessment." *Teaching in Higher Education* 11 (3): 279–290.
- Liu, E. Z. F., and S. M. Yuan. 2003. "A Study of Students' Attitudes toward and Desired System Requirement of Networked Peer Assessment System." *International Journal of Instructional Media* 30 (4): 349–354.
- McGarr, O., and A. M. Clifford. 2013. "'Just Enough to Make You Take It Seriously': Exploring Students' Attitudes towards Peer Assessment." *Higher Education* 65 (6): 677–693.
- McNeil, L. M. 1988. *Contradictions of Control: School Structure and School Knowledge*. New York: Routledge.

- Nulty, D. D. 2008. "The Adequacy of Response Rates to Online and Paper Surveys: What Can Be Done?" *Assessment & Evaluation in Higher Education* 33 (3): 301–314.
- Patchan, M. M., C. D. Schunn, and R. J. Clark. 2011. "Writing in Natural Sciences: Understanding the Effects of Different Types of Reviewers on the Writing." *Journal of Writing Research* 2 (3): 365–393.
- Patchan, M. M., C. D. Schunn, and R. J. Correnti. 2016. "The Nature of Feedback: How Peer Feedback Features Affect Students' Implementation Rate and Quality of Revisions." *Journal of Educational Psychology* 108 (8): 1098–1120.
- Paulus, T. M. 1999. "The Effect of Peer and Teacher Feedback on Student Writing." *Journal of Second Language Writing* 8 (3): 265–289.
- Pintrich, P. R. 1999. "The Role of Motivation in Promoting and Sustaining Self-Regulated Learning." *International Journal of Educational Research* 31 (6): 459–470.
- Planas Lladó, A., L. F. Soley, R. M. F. Sansbelló, G. A. Pujolras, J. P. Planella, N. Roura-Pascual, J. J. Suñol Martínez, and L. M. Moreno. 2014. "Student Perceptions of Peer Assessment: An Interdisciplinary Study." *Assessment & Evaluation in Higher Education* 39 (5): 592–610.
- Praver, M., G. Rouault, and J. Eidswick. 2011. "Attitudes and Affect toward Peer Evaluation in EFL Reading Circles." *The Reading Matrix* 11 (2): 89–101.
- Reid, J. M. 1993. *Teaching ESL Writing*. Englewood Cliffs, NJ: Prentice Hall Regents.
- Roskams, T. 1999. "Chinese Efl Students' Attitudes to Peer Feedback and Peer Assessment in an Extended Pairwork Setting." *RELC Journal* 30 (1): 79–123.
- Russell, A. A. 2004. "Calibrated Peer Review – A Writing and Critical-Thinking Instructional Tool." In *Teaching Tips: Innovations in Undergraduate Science Instruction*, edited by M. Druger, E. D. Siebert, and L. W. Crow, 54–55. Arlington: National Science Teachers Association.
- Sasmaz Oren, Fatma 2012. "The Effects of Gender and Previous Experience on the Approach of Self and Peer Assessment: A Case from Turkey." *Innovations in Education and Teaching International* 49 (2): 123–133.
- Schunn, C. D. 2016. "Writing to Learn and Learning to Write through SWoRD." In *Adaptive Educational Technologies for Literacy Instruction*, edited by S. A. Crossley and D. S. McNamara, 244–259. New York: Taylor & Francis, Routledge.
- Schunn, C. D., A. Godley, and S. DeMartino. 2016. "The Reliability and Validity of Peer Review of Writing in High School AP English Classes." *Journal of Adolescent & Adult Literacy* 60 (1): 13–23.
- Takeda, S., and F. Homberg. 2014. "The Effects of Gender on Group Work Process and Achievement: An Analysis through Self- and Peer-Assessment." *British Educational Research Journal* 40 (2): 373–396.
- Tang, W. L. 2012. "Study and Analysis on Chinese Non-English Major College Students' Demotivation in English Learning." *Foreign Language Education* 33 (1): 70–75.
- Topping, K. 1998. "Peer Assessment between Students in Colleges and Universities." *Review of Educational Research* 68 (3): 249–276.
- Voogt, J., P. Fisser, N. Pareja Roblin, J. Tondeur, and J. van Braak. 2013. "Technological Pedagogical Content Knowledge – A Review of the Literature." *Journal of Computer Assisted Learning* 29 (2): 109–121.
- Wang, Y., Y. Liang, L. Liu, and Y. Liu. 2016. "A Multi-peer Assessment Platform for Programming Language Learning: Considering Group Non-Consensus and Personal Radicalness." *Interactive Learning Environments* 24 (8): 2011–2031.
- Wang, Y., A. Xuan, and Y. J. Chen. 2006. "Analysis and Investigation of Chinese Non-English Majors' English Writing Instruction." *Foreign Language World* 115 (5): 22–27.
- Warrington, M., M. Younger, and J. Williams. 2000. "Student Attitudes, Image and the Gender Gap." *British Educational Research Journal* 26 (3): 393–407.
- Watanabe, Y., and M. Swain. 2007. "Effects of Proficiency Differences and Patterns of Pair Interaction on Second Language Learning: Collaborative Dialogue between Adult ESL Learners." *Language Teaching Research* 11 (2): 121–142.
- Wen, M. L., and C. C. Tsai. 2006. "University Students' Perceptions of and Attitudes toward (Online) Peer Assessment." *Higher Education* 51: 27–44.
- Wen, M. L., C. C. Tsai, and C. Y. Chang. 2006. "Attitude towards Peer Assessment: A Comparison of the Perspectives of Pre-Service and in-Service Teachers." *Innovations in Education and Teaching International* 43 (1): 83–92.
- Yang, L. 2014. "Examining the Mediational Means in Collaborative Writing: Case Studies of Undergraduate ESL Students in Business Courses." *Journal of Second Language Writing* 23: 74–89.
- Yu, S., and I. Lee. 2015. "Understanding EFL Students' Participation in Group Peer Feedback of L2 Writing: A Case Study from an Activity Theory Perspective." *Language Teaching Research* 19 (5): 572–593.