Factors Influencing Unfair Behavior in Peer Assessment: Evidence from a Situational Experiment

Haijia Liu[1], Yuqiang Feng[2]

{1071606947@qq.com[1], fengyq@hit.edu.cn[2] }

Harbin Institute of Technology, Harbin, Heilongjiang China

Abstract. We explored the influence of different factors on unfair behavior in peer assessment and designed interventions aimed to reduce the occurrence of unfair behavior. Different peer information will affect students' attitudes and behavior toward peer assessment. We analyze the influence of different helpful reviews. Based on social value orientation theory, we developed novel information frameworks in combination with cooperation (e.g., your behaviors benefit others) and competition (e.g., your behaviors benefit yourself more than others). We invited real students who use peer assessment to a situational experiment, conducted a covariance analysis, . We tested peer reviews that provided different levels of helpfulness and found that high review helpfulness helped reduce the incidence of unfair behavior. Both the competition and cooperation frameworks can effectively reduce the occurrence of unfair behavior, with the competition framework more effective. Our work identifies the influence of peer information on unfair behavior in peer assessment, and the results suggest that use of an information framework could reduce this behavior. Our research extends studies of peer assessment and provides feasible practical methods.

Keywords: unfair behavior, reviews helpfulness, information framework, score bias

1 Introduction

Peer assessment has been increasingly applied in recent years. In higher education, it can effectively reduce the workload of teachers, enabling them to devote more time to teaching preparation. Additionally, it provides a way for students to communicate with each other and promote group progress. Although there are many advantages of peer assessment, this approach may be questioned and distrusted by some participants. Because some participants view the process as unfair, many studies have investigated the fairness of peer assessment [1-4].

In studies of peer assessment, fairness is often mentioned by participants as a concern, making fairness an important source of trust for a platform. Some studies have proposed the concept of conflict of interest (COI) to measure the cooperative or competitive relationships between authors and reviewers. Non-compliant COI values indicate the presence of cooperation and competition between authors and reviewers, so peer assessment should not be used [5]. However, in addition to the relationships between authors and reviewers, there may be other factors that can affect fairness, and few empirical studies have been performed to comprehensively assess the fairness of peer assessment.

NMDME 2023, October 13-15, Xi'an, People's Republic of China

DOI 10.4108/eai.13-10-2023.2341083

The use of peer assessment in higher education seems highly prone to complications of competition and cooperation between the people doing the assessment and the people being assessed. Most unfair behaviors can be attributed to reciprocal behavior. In the peer assessment literature, the term "reciprocity effect" is often used to refer to biases caused by interpersonal relationships in peer assessment [6], including giving higher or lower scorings than appropriate [2]. With interactive characteristics, the process of peer review can be divided into two stages. In the first stage, reviewers review and score works submitted by authors. In the second stage, authors give feedback on the reviews. There have been many studies of the first stage [7-9], but less attention to the second stage of the process [10]. Therefore, to more comprehensively study more influencing factors (including peer information) on the fairness of peer assessment, we focused on the authors' feedback process of reviewer assessment.

In the process of author feedback to reviewer assessment, reciprocity may lead to two different kinds of unfair behavior, collusion behavior and malicious grading behavior. Collusive behavior is when reviewers give high marks regardless of the quality of work and authors also give high feedback, due to the presence of friendship markers or social pressure between authors and reviewers. Malicious scoring behavior is when an author chooses to give a low score to the reviewer regardless of whether the review is correct after the reviewer provides a negative review to the author. Previous work found anonymity may reduce the harm of collusion, but may also increase malicious scoring behavior [11,12]. In this study, we explored ways to simultaneously reduce the occurrence of both behaviors. Thus, there are two main research questions:

RQ1: What factors can affect the occurrence of unfair behavior in the peer assessment process?

RQ2: What kind of countermeasures can reduce the occurrence of unfair behavior in the peer assessment process?

2 Theory and hypothesis

2.1 Fairness of peer assessment

Fairness is an important aspect of peer assessment, and many studies have explored this concept. Students regard fairness as a fundamental issue in the peer assessment process in the learning environment [13]. Learners' views on fairness can be used to judge whether the assessment method can support high-quality learning [4]. Learners may resist peer assessment due to perception bias [14,15]. Previous findings suggest that unfairness often comes from the interactions of peers [1], but different people may have different views on whether peer assessment is fair [16]. Although there have been many studies on the perception of fairness, few studies have explored the causes and consequences of fair behavior. Studies of the friendship markers among peers concluded that the lower the friendship level, the higher the possibility of fair peer assessment [2,3].

To effectively promote the application of peer assessment and address the concerns, studies have evaluated the factors influencing the fairness of peer review. The most studied factor is the relationship between author and reviewer, also known as conflict of interest (COI). A number of studies explored the impacts of COI and concluded that competitive or cooperative

relationships should be avoided for peer evaluation matching [17-19]. However, this may be impossible in some learning environments, so it remains a challenge to determine the best strategies to ensure fairness. Some studies have suggested the advantages of anonymous assessments, which can reduce the negative impact of interpersonal relationships on fairness [11], but this has not been explored in depth. Therefore, a goal of this work was to more comprehensively evaluate strategies to ensure the fairness of peer assessment.

2.2 Influence of peer information

In the context of higher education, peer assessment is a collaborative activity that occurs at least between two peers [20], so there is interaction. Peer interaction promotes learner-centered processes and collaborative learning [21], and peer networks around learners can affect these processes [22].

Previous studies were mostly carried out from the perspective of interpersonal relationships. The core of peer assessment is giving and receiving feedback from others to improve individual performance. As an interpersonal process, the variables associated with peer assessment include peer pressure due to friendship, hostility, or other power processes, and social discomfort with criticism [23,24]. Scholars have analyzed the influences of four interpersonal variables of psychological security, value diversity, interdependence, and trust on peer assessment. Psychological security will affect the friendship markers among peers and thus have an impact on fairness [25]. There are also studies suggesting that peer pressure may decrease evaluation accuracy [26].

Peer effects can also affect peer assessment, and studies have analyzed the effect of individual performance levels on peers [22]. Peer assessment is characterized by interactions between students at different levels, and screening useful information and providing critical judgment are parts of the communication between students [27]. Little work has explored this aspect of peer assessment. Therefore, from the perspective of peer information, we hypothesize that the helpfulness of peer information may affect the occurrence of unfair behavior. Therefore, hypothesis 1 is proposed:

H1: Peer reviews with high helpfulness differ less from a normal feedback score than peer reviews with low helpfulness.

2.3 Social value orientation theory and gender difference

Social value orientation (SVO) describes the relative "weights" that an individual will place on his or her own welfare and that of others [28,29]. SVO theory proposes that preferences are heterogeneous when individuals relate to their own and others' benefits. Individuals can be divided into three different types due to their idiosyncrasies: (a) "cooperative" (maximizing the interests of others and their own); (b) "individualistic" (maximizing their own interests, ignoring those of others); and (c) "competitive" (maximizing their own interests relative to those of others). **Figure 1** presents a graphical representation of SVO theory.

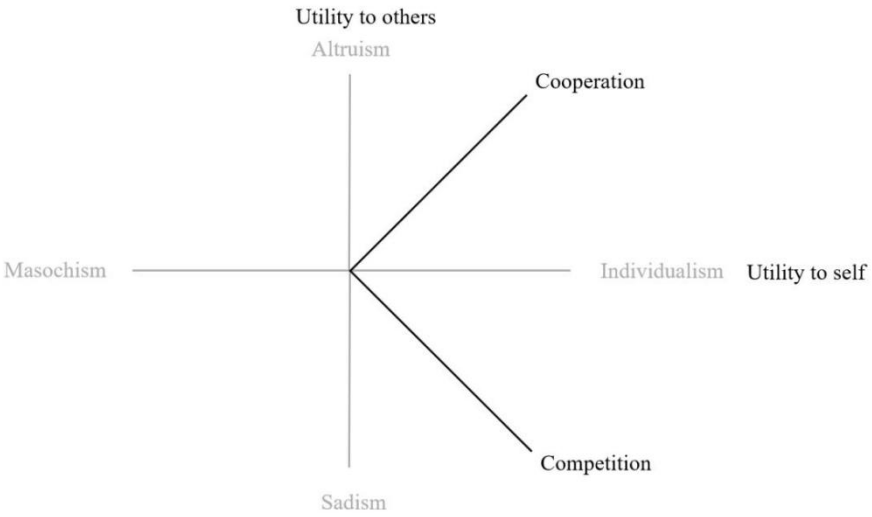

Fig. 1. Social Value Orientations [28,29]

Applying SVO theory, we focused our research on the effects of cooperation and competition tendencies affecting the interactive nature of peer assessment. Although the SVO of participants was not directly observed, previous work has suggested that gender is highly correlated with SVO [30,31]. Studies in other fields have found that women are more likely to be cooperative than men and men tend to be more competitive [32,33]. Women tend to be more inclined to oppose inequality when making decisions and choose a more equal (fair) allocation of capital than men [34,35].

Marketing assumes that consumers can vary in preferences and attitudes, resulting in differences in consumer behavior. However, much of the literature on peer assessment has not acknowledged the possible impacts of differences in student demographics, culture, and other factors [36]. However, different groups of students may use the peer assessment process in different ways and, if the process affects the formal grades awarded, these differences will produce inconsistent and potentially unfair results. Therefore, the goal of this study was to explore the differences in unfair behaviors of participants with different value orientations in peer assessment. We use different genders to replace different value orientations [37], and thus propose hypothesis 2:

H2: The score bias of male participants from the normal feedback scores is higher than that of female participants.

2.4 Information Frameworks

There have been many studies of behavioral intervention using different information frameworks. For the solicitation of charity fundraising, scholars tested altruistic and egoistic information frameworks and observed that altruistic frameworks are usually more likely to attract charitable donations than egoistic frameworks [38]. In studies of user-generated content, the performance feedback of cooperative frameworks was found to more effective in motivating users than competitive frameworks [37].

In the education environment, there have been few intervention studies on information frameworks. However, it is reasonable that technology and learning environment could affect the process of peer assessment [39]. Therefore, we considered the influence of the information framework on the learning environment to study unfair behavior in peer assessment. We used the SVO theory to construct a competitive information framework (egoism) and a cooperative information framework (altruism), and explored whether these different frameworks could affect participants' behaviors in the context of peer assessment, and tested for differences between different genders. Building on the different outcomes that different frameworks have produced for individuals with different value tendencies in previous research [37,38,40]. According to the conclusions of previous work, there are obvious differences in the influences of cooperation and competition frameworks on participant behavior. A cooperation framework will generate more incentives for participants to benefit each other, and competition framework will often strengthen the participants' hostility to each other [41-43]. However, at the same time, explicit intervention may make participants more careful in providing feedback and scores. Therefore, hypothesis 3 is proposed:

H3: Competitive and cooperative frameworks result in less bias compared with the absence of information frameworks.

3 Situational experiment

3.1 Experimental design

We designed and conducted a set of situational experiments using the Questionnaire Star website. As participants, we selected undergraduate students who had used an online peer assessment platform to evaluate assignments for programming courses. The selected participants were students in different types of programming courses and in different classes, but the same scoring system was used, for a similar experience of the peer evaluation process. As regular users of the system, the participants were very familiar with the different situations in the experimental design and were able to effectively accept the stimulus of different frames.

During our experiment, we set up a special scenarios and groupings to test the hypothesis. We used a pre-experiment to first divided the helpful reviews into high helpful reviews and low helpful reviews. We manipulated different frameworks for the stimulus, with control (no framework stimulus), competitive, and cooperative frameworks. A total of 6 (2*3) groups were finally set up for the test. In each group, we provided participants with the special scenarios (excellent homework completion scenarios) and a corresponding framework stimulus. For each situation and stimulus, we provided the participants with four different peer reviews, and then asked the participants to give feedback according to their own real situation. We then compared the obtained feedback scores with the objectively determined scores to assess the occurrence of unfair behavior and bias in the scoring and feedback.

3.2 Survey to verify review helpfulness

To determine the impact of review helpfulness, we conducted a questionnaire survey among 30 users of the peer assessment platform. We designed three scenarios for each participant with different levels of work completion and provided five different reviews for each scenario.

The reviews were based on real data from the courses taken by the participants. According to the premise of different situations, select the real reviews in the peer assessment platform. Participants were then asked to provide feedback scores and assign a helpfulness level for these reviews (a 5-point Likert scale). For scenarios 1 and 3, there were two significant differences in helpfulness level for reviews, and in scenario 2 there were no significant differences. Finally, we selected two reviews for scenario 1 and two reviews for scenario 3 as representative of the helpfulness levels in our formal experiment. Additionally, the average scores of 30 participants for these four reviews were calculated as the standard feedback scores for the reviews.

3.3 Variables and measurements

Dependent variables. We constructed a dependent variables to measure the occurrence of unfair behavior. The presence of a large bias between the feedback score and the standard score was assessed, with a larger bias corresponding to more likely unfair behavior. This is measured by the absolute value of the difference between participants' feedback scores and standard feedback scores (measured in 3.2). The larger the value, the more unfair the actual feedback grading behavior and the smaller the value, the less unfair.

Other variables. To study the influence of different personal traits (competition-cooperation tendency) on feedback scoring behavior, we obtained the gender information for each student and constructed a binary variable of Gender (1 representing male and 0 representing female) by replacing the personal trait variable with the gender variable. To study the influence of different peer information (helpfulness of peer reviews) on feedback scoring behavior, we obtained the helpfulness of review and constructed a binary variable of Helpfulness (1 representing high helpfulness and 0 representing low helpfulness). In addition, to reduce the occurrence of unfair behaviors, we set different background frameworks as variable Framework (0 represents the no information framework group, 1 represents the competition framework group, and 2 represents the cooperation framework group). The age and feedback time of participants were used as control variables.

4 Data Analysis

4.1 Descriptive Analysis

After screening the experimental data (excluding data with insufficient feedback time and abnormal scores), different grouping methods were used for statistical analysis. As seen from **Table 1,** compared with the reviews with low helpfulness, the reviews with high helpfulness obtained higher feedback scores with less bias from the normal scores. Female participants' feedback scores were higher than male participants, and the bias from the normal score was smaller, but there was no obvious gender difference. Among the three groups, the competitive framework gave the lowest feedback score, the cooperative framework gave the highest feedback score, and the competitive framework showed the smallest deviation, followed by the cooperative framework.

Table 1. Descriptive analysis

Items	Parameter	Helpfulness		Gender		Framework			Total
		High helpfulness	Low helpfulness	Female	Male	Competition framework	Cooperation framework	No information framework	
Score bias	n	68	77	76	69	47	50	48	145
	Mean	6.89	13.43	10.83	9.84	8.24	9.82	13.01	10.36
	SD	4.57	7.44	7.01	7.10	6.35	6.90	7.14	7.05

4.2 Covariance Analysis

Verification of the dependent variable To use covariance method for analysis, it is first necessary to verify whether the dependent variable follows a normal distribution. Since the sample size was over 50, the Kolmogorov-Smirnov test was used for verification (shown in Table 2). Although the results are significant and do not strictly conform to the normal distribution, the absolute value of kurtosis was less than 10 and the absolute value of skewness was less than 3, indicating that although the data is not absolutely normal, it can be basically accepted as a normal distribution. Therefore, covariance analysis can be used.

Table 2. Kolmogorov-Smirnov test

Items	n	Mean	Std.	Skewness	kurtosis	Kolmogorov-Smirnov test	
						Statistic *D*	*p*
Scoring bias	145	10.361	7.046	0.575	-0.935	0.201	0.000**

Parallelism test Before covariance analysis, we first tested the parallelism of the covariables. As shown in **Table 3**, the interaction terms showed no significance so pass the parallelism test.

Table 3. ANCOVA Parallelism Test

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	198.524	1	198.524	5.309	0.023*
Gender	1.455	1	1.455	0.039	0.844
Helpfulness	2.043	1	2.043	0.055	0.816
Framework	98.643	2	49.322	1.319	0.271
Age	26.611	1	26.611	0.712	0.400
Gender*Age	0.582	1	0.582	0.016	0.901
Helpfulness*Age	12.214	1	12.214	0.327	0.569
Framework*Age	111.059	2	55.529	1.485	0.230
Time	15.834	1	15.834	0.423	0.516
Gender*Time	7.771	1	7.771	0.208	0.649

Source	Type III Sum of Squares	df	Mean Square	F	p
Helpfulness*Time	37.632	1	37.632	1.006	0.318
Framework*Time	56.600	2	28.300	0.757	0.471
Residual	4861.514	130	37.396		

R^2: 0.320

* p<0.05 ** p<0.01

Covariance analysis The scoring bias was used as the dependent variable for covariance analysis. As shown in **Table 4**, the helpfulness of reviews and the information framework have a significant impact on scoring bias. The scoring bias of the group with high review helpfulness was significantly lower than that of the group with low review helpfulness, supporting hypothesis 1. From the perspective of gender, there was higher bias of females' scores than that of males, but this was not significant, which does not support hypothesis 2. For the information framework, the competitive framework group showed the smallest scoring bias, followed by the cooperative framework group. The group without information frame has the largest scoring bias, supporting hypothesis 3. Although both competitive information and cooperative information can effectively reduce the scoring bias, the competitive framework was more effective.

Table 4. ANCOVA

Source	Type III Sum of Squares	df	Mean Square	F	p
Intercept	163.193	1	163.193	4.416	0.037*
Gender	39.743	1	39.743	1.076	0.302
Helpfulness	1401.203	1	1401.203	37.919	0.000**
Framework	447.934	2	223.967	6.061	0.003**
Age	8.533	1	8.533	0.231	0.632
Time	13.031	1	13.031	0.353	0.554
Residual	5099.395	138	36.952		

R^2: 0.287

* p<0.05 ** p<0.01

5 Conclusion

With the continued development of information technology and the popularization of online teaching, online peer assessment is more widely used. With this increased use of peer assessment, there is significant interest in determining how to eliminate participants' concerns about unfair behavior. This study addressed unfair behavior for peer assessment by situational experiment and covariance analysis and there are several insights from our work.

First, our research shows that peer information affects unfair behavior. When peers offered more helpful and meaningful reviews, participants scored their feedback more fairly. This means that in the process of peer assessment, both sides of the process will have an impact on each other. When one partner participates in the process with a positive and serious attitude, the unfair behavior of the other side will be significantly reduced.

Second, the results of the situational experiment reveal that different information frameworks have different effects on unfair behavior. Of the two frameworks, the competitive framework is the better treatment. This may be because, in the context of a competitive framework, participants are able to recognize the value of peer review, so they give more serious reviews and mark more carefully. At the same time, because participants may believe that unfair behavior will be detected and punished in a competitive environment, it leads to more truthful and reliable assessments and fewer hostile reviews. However, although a cooperative framework builds an environment for participants to make progress together, making them more willing to help others seriously and provide quality reviews, they may have a more relaxed tendency in terms of scoring.

6 Contributions and limitations

This study effectively addresses the shortcomings of previous research on peer assessment of unfair behavior. This analysis finds significant influence of peer information and information framework on unfair behavior. Unlike previous work that focused on peer relationships, our study has stronger universality and suggests a new research direction for studying unfair behavior in higher education. Additionally, this is the first application of an information framework for study of peer assessment, thus expanding the research perspective of peer assessment and providing additional directions for future behavior research.

The results of this study are practically significant. First, we found that quality peer information can reduce unfair behavior, a finding that is consistent with previous speculation. To gain the trust of participants, appropriate behavior of participants is required in the peer assessment platform and effective review output is the key to maintain fairness. Second, different information frameworks will have different impacts on unfair behavior and treatments of cooperation and competition frameworks will significantly reduce the possibility of unfair behavior, with competition frameworks more effective. The results of this paper have strong practical significance, showing that the occurrence of unfair behavior can be effectively reduced by maintenance of the peer assessment environment or by relatively simple framework stimulation.

The limitation of this study lies in the data obtained from peer assessment participants in higher education. Peer assessment tends to exist in more broad backgrounds, and we can expand the research on peer assessment in the future to obtain more universal conclusions. At the same time, the study did not examine the individual characteristics of the participants. In future studies, heterogeneity analysis can be carried out according to the characteristics of participants, and more targeted methods can be proposed to reduce unfair peer assessment behavior.

Reference

[1] Carvalho, A.: Students' perceptions of fairness in peer assessment: Evidence from a problem-based learning course. Teaching in Higher Education. Vol. 18, No.5, pp. 491-505 (2013)

[2] Panadero, E., Romero, M., & Strijbos. J. W.: The impact of a rubric and friendship on peer assessment: Effects on construct validity, performance, and perceptions of fairness and comfort. Studies in Educational Evaluation. Vol. 39, pp. 195–203 (2013)

[3] Dominguez, C., Jaime, A., Sanchez, A., Blanco, J. M., & Heras, J.: A comparative analysis of the consistency and difference among online self-, peer-, external- and instructor-assessments: The competitive effect. Computers in Human Behavior. Vol. 60, pp. 112–120 (2016)

[4] Hailikari, T., Postareff, L., Tuononen, T., Ra€isa€nen, M., & Lindblom-Ylanne, S.: Students' and teachers' perceptions of fairness in assessment. Advances and innovations in university assessment and feedback. pp. 99-113 (2014)

[5] Zhao, X., & Zhang, Y.: Reviewer assignment algorithms for peer review automation: A survey. Information Processing and Management. Vol. 59, 103028 (2022)

[6] Strijbos, J. W., Ochoa, T. A., Sluijsmans, D. M. A., Segers, M. S. R., & Tillema, H. H.: . Fostering interactivity through formative peer assessment in (web-based) collaborative learning environments. Cognitive and emotional processes in web-based education: Integrating human factors and personalization. pp. 375–395 (2009)

[7] Dmoshinskaia, N., Gijlers, H., & Jong, T. D.: Learning from reviewing peers' concept maps in an inquiry context: reviewing or grading, which is better? Studies in Educational Evaluation. Vol. 68, 100959 (2021)

[8] Grant, P., Jonah, B., Jeffrey, J., & Andrew, T.: How Concrete Language Shapes Customer Satisfaction. Journal of Consumer Research. Vol. 47, No. 5, pp. 787–806 (2021)

[9] Jeffery, D., Krassimir, Y., Alison, C., & Kerry, R.: How to achieve accurate peer assessment for high value written assignments in a senior undergraduate course. Assessment & Evaluation in Higher Education. Vol. 41, No. 1, pp. 127–140 (2016)

[10] Huisman, B., Saab, N., van den Broek, P., & van Driel, J.: The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. Assessment & Evaluation in Higher Education. Vol. 44, No. 6, pp. 863–880 (2019)

[11] Lin, G.: Anonymous versus identified peer assessment via a Facebook-based learning application: Effects on quality of peer feedback, perceived learning, perceived fairness, and attitude toward the system. Computers & Education. Vol. 116, pp. 81–92 (2018)

[12] Vanderhoven, E., Raes, A., Schellens, T., & Montrieux, H.: Face-to-face peer assessment in secondary education: does anonymity matter? Procedia - Social and Behavioral Sciences. Vol. 69, pp. 1340–1347 (2012)

[13] Sambell, K., McDowell, L., & Brown, S.: "But is it fair?": an exploratory study of student perceptions of the consequential validity of assessment. Studies in Educational Evaluation. Vol. 23, No. 4, pp. 349-371 (1997)

[14] Kaufman, J. K., & Schunn, C. D.: Students' perceptions about peer assessment for writing: Their origin and impact on revision work. Instructional Science. Vol. 39, No. 3, pp. 387-406 (2011)

[15] Wen, M. L., & Tsai, C. C.: University students' perceptions of and attitudes toward (online) peer assessment. Higher Education. Vol. 27, No. 18, pp. 27-44 (2006)

[16] van Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H.: Peer assessment for learning from a social perspective: The influence of interpersonal variables and structural features. Educational Research Review. Vol. 4, pp. 41–54 (2009)

[17] Pradhan, D. K., Chakraborty, J., Choudhary, P., & Nandi, S.: An automated conflict of interest based greedy approach for conference paper assignment system. Journal of Informetrics. Vol. 14, No. 2, 101022 (2020)
[18] Wu, S., Leong Hou, U., Bhowmick, S. S., & Gatterbauer, W.: Conflict of interest declaration and detection system in heterogeneous networks. In 2017 ACM conference on information and knowledge management. pp. 2383–2386 (2017)
[19] Yan, S., Jin, J., Geng, Q., Zhao, Y., & Huang, X.: Utilizing academic-network-based conflict of interests for paper reviewer assignment. International Journal of Knowledge Engineering. Vol. 2, No. 3, pp. 65–73 (2017)
[20] Kollar, I., & Fischer, F.: Peer assessment as collaborative learning: A cognitive perspective. Learning and Instruction. Vol. 20, No. 4, pp. 344-348 (2010)
[21] Choi, E., Schallert, D. L., Jee, M. J. & Ko, J.: Transpacific telecollaboration and L2 writing: Influences of interpersonal dynamics on peer feedback and revision uptake. Journal of Second Language Writing. Vol. 54, 100855 (2021)
[22] Vignery, K., & Laurier, W.: Achievement in student peer networks: A study of the selection process, peer effects and student centrality. International Journal of Educational Research. Vol. 99, 101499 (2020)
[23] Cartney, P.: Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. Assessment & Evaluation in Higher Education. Vol. 35, No. 5, pp. 551-564 (2010)
[24] Topping, K. J.: Self and peer assessment in school and university: Reliability, validity and utility. Optimizing new modes of assessment: In search of qualities and standards. pp. 55-87 (2003)
[25] van Van Gennip, N. A. E., Segers, M. S. R., & Tillema, H. H.: Peer assessment as a collaborative learning activity: The role of interpersonal variables and conceptions. Learning and Instuction. Vol. 20, pp. 280-290 (2010)
[26] Falchikov, N.: Involving students in assessment. Psychology Learning and Teaching, Vol. 3, No. 2, pp. 102-108 (2003)
[27] Yao, Y., Wang, W., & Yang, X.: Perceptions of the inclusion of Automatic Writing Evaluation in peer assessment on EFL writers' language mindsets and motivation: A short-term longitudinal study. Assessing Writing. Vol. 50, 100568 (2021)
[28] Fiedler S., Glöckner A., Nicklisch A., Dickert S.: Social value orientation and information search in social dilemmas: An eyetracking analysis. Organizational Behavior and Human Decision Processes. Vol. 120, No. 2: pp. 272–284 (2013)
[29] Van Lange P. A.: The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. Journal of Personality and Social Psychology. Vol. 77, No. 2, pp. 337–349 (1999)
[30] Gupta D. N., Poulsen A., Villeval M.C.: Gender matching and competitiveness: Experimental evidence. Economic Inquiry. Vol. 51, No. 1, pp. 816–835 (2013)
[31] Croson R., Gneezy U.: Gender differences in preferences. Journal of Economic Literature. Vol. 47, No. 2, pp. 448–474 (2009)
[32] Gneezy, U., Niederle, M., and Rustichini, A.: Performance in Competitive Environments: Gender Differences. Quarterly Journal of Economics. Vol. 118, No. 3, pp. 1049–1074 (2003)
[33] Stocks, E. L., Lishner, D. A., and Decker, S. K.: Altruism or Psychological Escape: Why Does Empathy Promote Prosocial Behavior? European Journal of Social Psychology. Vol. 39, No. 5, pp. 649–665 (2009)
[34] Rand D. G., Brescoll V. L., Everett J. A., Capraro V., Barcelo H.: Social heuristics and social

roles: Intuition favors altruism for women but not for men. Journal of Experimental Psychology General. Vol. 145, No. 4, pp. 389–396 (2016)
[35] Dufwenberg M., Muren A.: Gender composition in teams. Journal of Economic Behavior and Organization. Vol. 61, No. 1, pp. 50–54 (2006)
[36] Mcleay, F., & Wesson, D.: Chinese versus UK marketing students' perceptions of peer feedback and peer assessment. The International Journal of Management Education. Vol. 12, pp. 142–150 (2014)
[37] Huang, N., Burtch, G., Gu, B., Hong, Y., Liang, C., Wang, K., Fu, D., & Yang, B.: Motivating User-Generated Content with performance feedback: Evidence from randomized field experiments. Management Science. Vol. 65, No. 1, pp. 327-345 (2019)
[38] Fisher R., Vandenbosch M., & Antia K.: An empathy-helping perspective on consumers' responses to fund-raising appeals. Journal of Consumer Research. Vol. 35, No. 1, pp. 519–531 (2008)
[39] Chang, C., Lee, D., Tang, K., & Hwang, G.: Effect sizes and research directions of peer assessments: From an integrated perspective of meta-analysis and co-citation network. Computers & Education. Vol. 164, 104123 (2021)
[40] Brunel F., Nelson M.: Explaining gendered responses to "help-self" and "help-others" charity ad appeals: The mediating role of world-views. Journal of Advertising. Vol. 29, No. 3, pp. 15-28 (2000)
[41] Andreoni J. & Vesterlund L.: Which is the fair sex? Gender differences in altruism. Quarterly Journal of Economics. Vol. 116, No. 1, pp. 293–312 (2001)
[42] Dickinson D. L. & Tiefenthaler J.: What is fair? Experimental evidence. Southern Economic Journal. Vol. 69, No. 2, pp. 414–428 (2002)
[43] Dufwenberg M., Muren A.: Gender composition in teams. Journal of Economic Behavior & Organization. Vol. 61, No. 1, pp. 50–54 (2006)