



Factors that Influence the Helpfulness of Peer Comment: a Perspective of Content and Sentiment

Haijia Liu*, and Yuqiang Feng

School of Management, Harbin Institute of Technology, Harbin, China

*juliawhites@foxmail.com, fengyq@hit.edu.cn

Abstract. This study starts from two perspectives: the content and emotion of peer review in the peer assessment. This study examines the influence of concrete and suggestive comment content on the helpfulness of peer assessment and the influence of positive and negative sentiments on peer comment helpfulness. Collecting the data of 10 courses from 2014 to 2020 generated in a learning information system, we use natural language processing (NLP) methods to process text data and combine multiple regression methods to analyze. The results suggest that peer assessment with less concrete comments could have been more helpful. Peer comments having suggestive guidance were more helpful than those only with general exposition. Moreover, positive sentiments improve the helpfulness of peer comments, while negative sentiments reduce it. The study provides guidance on how to improve the quality of reviews in peer assessment.

Keywords: peer assessment, comment helpfulness, comment concreteness, suggestive comment, sentiment.

1 Introduction

Peer assessment is a writing activity in teaching in which students exchange assignments or tests and propose suggestions for revision, also known as peer feedback. Peer assessment often combines quantitative ratings with qualitative comments and can be used in a wide range of courses, from primary to higher education, as well as online MOOC courses. Existing studies have proven that, in general, the effect of written comments may be better than that of numerical ratings^[3], and qualitative comments are more important in supporting students' learning^[5]. Over time, with technological advances, computer-based online peer assessment functions have been developed and are widely used^[14].

Peer assessment has significant advantages over traditional teacher assessment. Peer assessment among students can result in rich feedback without using teachers' resources, which enables teachers to spend more time on other aspects of teaching^[13]. Students' generation and acceptance of assessment facilitate their refinement of the final work^[11]. However, some things could be improved in applying peer assessment in practical scenarios. It can be influenced by the type of task, student proficiency,

and gender differences^[1]. On the premise of obvious advantages and disadvantages, the difficulty in the application and promotion of peer assessment lies in how to obtain highly specific and clear peer assessment and how to promote a sense of trust in the assessment for the participating students and obtain help from it, which requires content analysis of the written comment of the peer assessment to ensure the quality of peer assessment results.

2 Background

2.1 Comment on Peer Assessment

The literature comment of the peer assessment application in online learning environments found that peer assessment is mainly used in the university context, mostly in higher education environments^[10]. In the process of formulating a peer assessment, there are usually two aspects (text comments and digital ratings), but the advantages of the two could be more consistent. Assessors often use text comments to put forward improvement suggestions, which digital ratings cannot provide. The digital rating can more intuitively indicate the level of the work. Therefore, there are some differences between the two. The help and benefits to the assessee usually come from comments, and existing studies have proven this. In the environment of peer assessment, the effect of comments is better than that of ratings. With the comments' help, the assessee's revised works are significantly better^[3]. Therefore, this study mainly considers comments in peer assessment.

2.2 NLP in Peer Assessment

In the context of pedagogy, there have been many attempts to apply NLP. In the context of peer assessment, NLP can be used to analyze the comment of students and provide support for the intelligence of peer assessment systems. It is feasible to judge the performance of students' comments on problem location and solution^[12]. In recent years, more and more researchers have made use of NLP methods, combined with corresponding dictionaries to make statistics on sentiments and parts of words, extract and classify keywords, and digitize words for further analysis^[9]. Based on previous methods, this study further uses NLP methods for word segmentation and the statistics of text information to transform unstructured text information into structured information for regression analysis.

3 Hypotheses and Model

3.1 The Influence of Comment Content on Peer Assessment

Concreteness describes the extent to which language refers to entities that are more physical, tangible, or "real," describing objects and actions (imagined or vivid) in a way that seems more concrete, familiar, and more easily perceived. Some scholars

have studied the influence of concreteness on language guidance. For example, concrete language can improve customer satisfaction in the service industry by improving customers' perception of being listened to^[4]. In the context of crowdfunding, concrete language has a significant impact on the success rate of social enterprises^[8]. Similarly, in the peer assessment process, when the assessor makes more effort to make a comment more concrete, the assessee will more likely tend to be more favorable to the concrete comment. First, concrete comments tend to point out programming errors and suggest the details of the problems. Second, concrete comments may make the assessee feel that the assessor has made more effort and recognize the effort from the assessors. Therefore, we propose hypothesis 1.

H1: The concreteness of comments positively influences the helpfulness of peer comments.

The content classification of peer comments mainly includes sentimental and cognitive comments^[6]. The comments defined as guidance indicate that the assessment content provides a clear direction for improvement and measures. Many students rated the comments they received about their work as useless, expressed little gratitude, and most of them said they needed clarification, were critical of the comments, or disagreed with them. Their main suggestion was that comments should be more constructive^[7]. In the peer assessment process, if there are suggestions, it indicates that the assessor has found deficiencies and gaps in the work and points out the pre-conditions for content improvement. According to the content of the comments, this study can be divided into general exposition and suggestive guidance. Thus, we propose hypothesis 2.

H2: Suggestive guidance provides more helpfulness for peer assessment than general exposition.

3.2 The Influence of Comment Sentiment on Peer Assessment

Sentimental language usually involves two aspects: positive sentiments and negative sentiments. There have been many studies on sentiment analysis in texts. It is possible that people's negative bias leads them to believe that negative comments are more authentic and have more reference value. Some studies have shown that the more negative the comments, the higher the score is in helping others^[2]. In the context of this study, comments can be divided into positive (related to supportive comments) and negative (related to negative comments). According to the characteristics of peer assessment, negative comments are more likely to point out deficiencies and the need for improvements in some works. In contrast, positive comments show the excellence of work. Negative comments may be more meaningful for the assessee if making progress is their goal. So, we provide hypothesis 3a and hypothesis 3b.

H3a: Negative comments may improve the helpfulness of peer assessment.

H3b: Positive comments may decrease the helpfulness of peer assessment.

3.3 Research Model

This study aimed to explore the influence of the content and sentiment of comments on helpfulness. As depicted in Fig. 1, we tested these hypotheses.

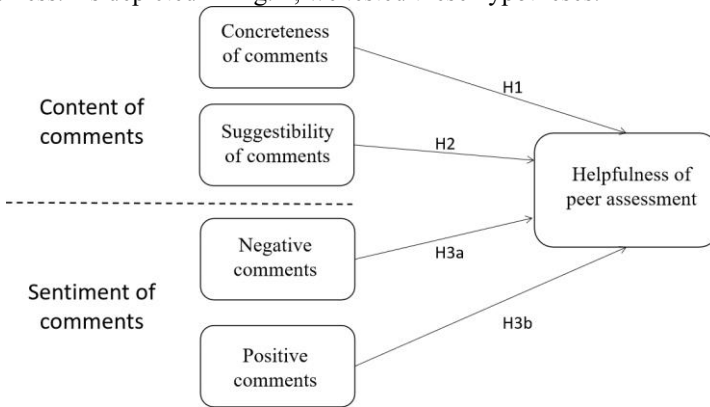


Fig. 1. The conceptual framework for the main relationship between the content and sentiment of comments with the helpfulness of peer assessment

4 Methods

4.1 Data Collection

We addressed our research questions by studying comments on online education platforms. The research of this paper relies on an educational peer code review system (EduPCR, <https://edupcr.net>) used in programming courses at a research-oriented national university in China. Because of ongoing system upgrades since 2004, the system has experienced many revisions, the data structure has constantly changed during upgrades, the data has been optimized, and the system was finally applied in studying the data from the ten courses since 2014. At the same time, to avoid causing privacy leakage problems, data is desensitized before analysis and use.

4.2 Analyzing Comments by NLP

In this study, two methods are used for analysis. The first method is to use the Python program to extract some comment data according to the statistical characteristics of keywords. The text information of the comments is processed by word segmentation, and then the number of different words is counted. According to our varying needs, the required keywords and the corresponding frequency of occurrence are screened out, and these data are used for further analysis. The second method is to use the existing functions of the Mandarin version of the LIWC (Linguistic Inquiry and Word Count) software package to form statistics on the words that need to be classified by the part of the comment. The LIWC software is the most commonly used language analysis tool for investigating the relationship between word use and psychological

variables. The scores obtained directly through software calculation are used as the source of variables for further analysis. The sentiment and content of peer comments are classified in this context by these two NLP analysis methods, and then the processed data are used for model analysis.

4.3 Measures

In this study, the LIWC software package was used to analyze the text information of peer comments. Existing research has identified three linguistic cues determining concreteness: articles, prepositions, and quantifiers^[8]. We use the sum of the scores of articles, prepositions, and quantifiers to calculate how concrete the comment is. We use $concrete_i$ to represent the degree of concreteness in comment i . In this study, suggestive guidance is an important component of comments. Our focus is on whether peer's suggestive guidance significantly impacts the helpfulness of comments, while general exposition does not. In the context of this study, the comment providing suggestions has typical iconic words, such as "suggest", at the beginning of the sentence. We use $suggestion_i$ to represent whether comment i is the style of suggestive guidance or general exposition (with 1 being suggestive guidance and 0 being general exposition). This study selected words of praise and appreciation to represent the comments' positive sentiments. In contrast, words of criticism and advice were selected to represent the negative sentiments in the comments. Through the processing of word segmentation results, the dictionary of positive and negative words is constructed by manual annotation. Then, the number of words appearing in the comment is counted, representing the comments' positive and negative sentiments. We use $positive_i$ to represent the positive sentiment in comment i and $negative_i$ to represent the negative sentiment in comment i .

In this study, feedback on peer comments from the assessee can be obtained on the platform. One of the feedback scores represents how helpful the comment is to the assessee (on a scale of 1 to 5, with 5 being the most helpful). Using this score as the level of helpfulness, $helpful_i$ represents the helpfulness of comment i .

Other studies have pointed out that the length of the comment's text may have a certain impact on the helpfulness of the comment. Therefore, the control variable is the length of the comment. The number of characters in each comment is counted, and $length_i$ represents the text length of comment i .

4.4 Data Analysis

We conducted the analysis using SPSS 22.0. We used multiple regression analysis to analyze the influence of comment content and emotion on comment helpfulness.

5 Results

5.1 The Influence of Comment Content

Our data suggest that the concreteness of comments negatively impacts peer helpfulness ($p < 0.01$), which is inconsistent with Hypothesis 1 (see Table 1). This may be due to the specific background of peer assessment. In most e-commerce or service situations, more concrete language will make the receiver feel that the assessor has exerted greater effort and listened to the receiver's needs more seriously, thus improving the perception of helpfulness. At this time, the receiver is often unfamiliar with the evaluation subject. However, in the context of peer assessment, the assessee completed their work, so they are more familiar with the work than the assessors. As the level of the assessor is often questioned, the assessee may not agree with it, causing them to disagree with the comment content. The more concrete the comments are, the more obvious the differences between the assessor and the assessee. This could lead to the consequence that the less helpful the assessee perceives the comments.

Our data indicate that suggestive guidance significantly impacts the helpfulness of peer comments ($p < 0.05$), consistent with Hypothesis 2 (see Table 1). This is because suggestive guidance is often highly targeted, and it presents a clear direction for improving deficiencies in the work, providing the method and operability of modification to the assessee. Therefore, this kind of comment can play an important role in peer assessment of the work in reference, pointing out that for evaluation to play a positive role, it is important to revise based on mistakes, and advice can serve as an important form of content by providing help.

5.2 The Influence of Comment Sentiment

Our data indicate that positive and negative sentimental language has different effects on the helpfulness of peer assessment: positive sentimental language has positive effects ($p < 0.01$), and negative sentimental language has negative effects ($p < 0.01$), which is inconsistent with Hypothesis 3a and Hypothesis 3b (see Table 1). The difference between the peer assessment and general product sales background may cause this. In the peer assessment system, the assessee object is the assessee's work rather than unknown products. In the online peer assessment system, positive sentimental language expresses the recognition and praise of the work, which makes the assessee feel supported and tends to produce positive feedback more easily. However, negative sentimental language represents different thoughts and ideas about the work. The assessee may be in love with their work, and it may be difficult for them to accept criticism from others fully, so in these cases, the assessors are less likely to receive positive feedback from the assessee. At the same time, other factors may contribute to this effect. In the research on peer assessment, we find that the peer relationship between the assessee and the assessor is particularly useful in a mutual peer assessment background. Because peer assessment scores are often used as part of course grading, there may be collusive behavior in the imperfect anonymous situation where assessors

give positive evaluations, and the assessee also gives high feedback scores. There may likewise be retaliatory low feedback scores for negative comments.

Table 1. Helpfulness multiple regression model (DV = helpful)

Explanatory variable	Coefficient	SE	t-value	[95% Conf Interval]	Sig
score _i	.019	.002	11.84	.016 .023	***
length _i	-.001	.001	-0.82	-.003 .001	
concrete _i	-.393	.121	-3.23	-.631 -.155	***
positive _i	.129	.012	10.87	.106 .152	***
negative _i	-.054	.02	-2.68	-.093 -.014	***
suggestion _i	.094	.047	1.98	.001 .186	**
Constant	2.104	.15	14.02	1.81 2.399	***
Mean dependent variables	3.967			SD dependent variables	1.415
R-squared	0.047			Number of observations	13140
F-test	87.153			Prob > F	0.000
Akaike crit. (AIC)	45792.917			Bayesian crit. (BIC)	45845.301

Note. *** $p < .001$, ** $p < .01$, * $p < .05$

6 Conclusion

We found the influence of comment content and emotion on comment helpfulness in peer assessment. Specifically, the more concrete the comments are, the less helpful the comments are, and the more suggestions they include, the more helpful the comments are, while positive comments are more helpful than negative comments. In the process of peer assessment, focusing on the helpfulness of comments can effectively enhance the advantages of this assessment method and provide more benefits for learners. This study provides the direction to guide learners to provide better comments.

References

1. Byl, E., & Topping, K., J. (2023). Student perceptions of feedback in reciprocal or nonreciprocal peer tutoring or mentoring. *Studies in Educational Evaluation*, 101304. <https://doi.org/10.1016/j.stueduc.2023.101304>.
2. Cao, Q., Duan, W. J., & Gan, Q. W. (2011). Exploring determinants of voting for the 'helpfulness' of online user reviews: A text mining approach. *Decision Support Systems*, 50(2), 511–521. <https://doi.org/10.1016/j.dss.2010.11.009>.
3. Dmoshinskaia, N., Gijlers, H., & Jong, T. D. (2021). Learning from reviewing peers' concept maps in an inquiry context: Commenting or grading, which is better? *Studies in Educational Evaluation*, 68, 100959. <https://doi.org/10.1016/j.stueduc.2020.100959>.

4. Grant, P., Jonah, B., Jeffrey, J., & Andrew, T. (2021). How concrete language shapes customer satisfaction. *Journal of Consumer Research*, 47(5), 787–806. <https://doi.org/10.1093/jcr/ucaa038>.
5. Huisman, B., Saab, N., van den Broek, P., & van Driel, J. (2019). The impact of formative peer feedback on higher education students' academic writing: A meta-analysis. *Assessment & Evaluation in Higher Education*, 44(6), 863–880. <https://doi.org/10.1080/02602938.2018.1545896>.
6. Lin, H., Hwang, G., Chang, S., & Hsu, Y. (2021). Facilitating critical thinking in decision making-based professional training: An online interactive peer-review approach in a flipped learning context. *Computers & Education*, 173, 104266. <https://doi.org/10.1016/j.cmpedu.2021.104266>.
7. Misiejuk, K., Wasson, B., & Egelanddal, K. (2021). Using learning analytics to understand student perceptions of peer feedback. *Computers in Human Behavior*, 117, 106658. <https://doi.org/10.1016/j.chb.2020.106658>.
8. Parhankangas, A., & Renko, M. (2017). Linguistic style and crowdfunding success among social and commercial entrepreneurs. *Journal of Business Venturing*, 32(2), 215–236. <https://doi.org/10.1016/j.jbusvent.2016.11.001>.
9. Stein, R. A., Jaques, P. A., & Valiati, J. F. (2019). An analysis of hierarchical text classification using word embeddings. *Information Sciences*, 471, 216–232. <https://doi.org/10.1016/j.ins.2018.09.001>.
10. Tenorio, T., Bittencourt, I. I., Isotani, S., & Silva, A. P. (2016). Does peer assessment in online 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>.
11. Wu, W. Y., Huang, J. Y., Han, C. W., & Zhang, J. (2022). Evaluating peer feedback as a reliable and valid complementary aid to teacher feedback in EFL writing classrooms: A feedback giver perspective. *Studies in Educational Evaluation*, 73, 101140. <https://doi.org/10.1016/j.stueduc.2022.101140>.
12. Xiong, W., Litman, D., & Schunn, C. (2012). Natural language processing techniques for researching and improving peer feedback. *Journal of Writing Research*, 4, 155–176. <https://doi.org/10.17239/jowr-2012.04.02.3>.
13. Yu, F., Liu, Y., & Liu, K. (2023). Online peer-assessment quality control: A proactive measure, validation study, and sensitivity analysis. *Studies in Educational Evaluation*, 78, 101279. <https://doi.org/10.1016/j.stueduc.2023.101279>.
14. Zheng, L., Zhang, X., & Cui, P. (2020). The role of technology-facilitated peer assessment and supporting strategies: A meta-analysis. *Assessment & Evaluation in Higher Education*, 45(3), 372–386. <https://doi.org/10.1080/02602938.2019.1644603>.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

