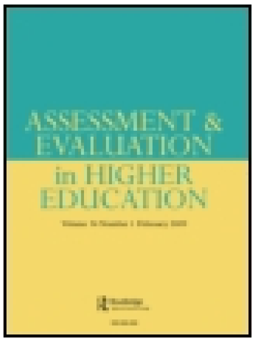

Assessment & Evaluation in Higher Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/caeh20

What makes students contribute more peer feedback? The role of within-course experience with peer feedback

Zheng Zong, Christian Schunn & Yanqing Wang

ASSESSMENT & EVALUATION IN HIGHER EDUCATION
https://doi.org/10.1080/02602938.2021.1968792

What makes students contribute more peer feedback? The role of within-course experience with peer feedback

Zheng Zong[a], Christian Schunn[b] and Yanqing Wang[a]

[a]School of Management, Harbin Institute of Technology, Harbin, China; [b]Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA, USA

ABSTRACT

The success of peer feedback approaches to instruction depends upon students contributing in-depth feedback to their peers. Prior researchers have examined the role of general attitudes towards peer feedback, but how experiences, especially the performance information during peer feedback, influence the subsequent amount of feedback that students provide to peers has received little attention. This study investigated what experience factors from one assignment predicted growth or declines in the amount of peer feedback provided on the next assignment in a course with many peer feedback assignments. Data on peer feedback experiences and behaviors across multiple assignments were taken from students across two programming courses ($N = 149$). Negative binomial regression analyses reveal three experiences in the prior assignment predicted growth in length of comments provided to peers: receiving more comments, doing well on the task, and receiving recognition for good reviewing (when not doing well). Implications for practice are presented.

KEYWORDS

Peer assessment; peer feedback; performance feedback; peer recognition

Introduction

After decades of intensive research in higher education, student peer review has proven to be reliable and effective (Bruce et al. 2016; Li et al. 2016; Huisman et al. 2019). As an interactive evaluation and learning method, peer review typically provides peers with valid ratings and comments that can be more effective than teacher comments for helping students improve their drafts (Cho and Schunn 2007; Bouzidi and Jaillet 2009). It is often called *peer assessment* when ratings are the focus and *peer feedback* when comments are the focus. Feedback is especially useful for students because it can be better understood (Nelson and Schunn 2009; Rotsaert, Panadero, and Schellens 2018; Wu and Schunn 2020). In the case of multi-peer feedback, several students provide more total comments (Cho and Schunn 2007, 2018). Further, both processes of providing feedback and receiving feedback offer learning opportunities for students (Lundstrom and Baker 2009; Crinon 2012; Adachi, Tai, and Dawson 2018; Deiglmayr 2018; Martin and Evans 2018). However, all of these positive outcomes depend upon students' willingness to participate actively in peer feedback and provide substantial comments to their peers. The more students offer, the more the feedback recipients revise (Popp and Goldman 2016) and the more the providers improve their writing (Lundstrom and Baker 2009; Wu and Schunn 2021; Zong, Schunn, and Wang 2021). Unfortunately, students are sometimes resistant to providing feedback (Liu

CONTACT Yanqing Wang yanqing@hit.edu.cn

and Carless 2006; Feng et al. 2019). Some researchers have suggested that peer review is unfamiliar to many students and that students become more willing with experience (Kankanhalli, Tan, and Wei 2005; Liu and Carless 2006). However, this claim is mainly untested yet, and here we take up the issue of how experience with peer feedback shapes the amount of feedback provided in later assignments.

Theoretical background

Over peer review cycles, students can regularly observe peer comments and reflect on them. Such feedback-practice loops can be conceptualized as a critical part of self-regulation. Self-regulation can involve three levels (Zimmerman and Risemberg 1997): *personal regulation*, which refers to cognitive and emotional regulation like self-efficacy and self-evaluation; *behavioral regulation*, which relates to modifying task behaviors; and *environmental regulation*, which involves seeking social supports and creating a more productive task environment. The three levels are thought to be interdependent. For example, changes in task behaviors can come about from changes in self-efficacy, such as being less willing to participate in peer feedback with low self-efficacy for the feedback topic.

In the context of peer feedback, student self-efficacy is likely to play an important role. In surveys and interviews, students often mention concern about their expertise as a reason to want to avoid participating in peer review (Liu and Carless 2006; Kaufman and Schunn 2011). This issue may be particularly problematic when students receive feedback that their expertise is relatively low compared with other students. Within the assignments involving peer review, students receive multiple sources of information that can shape their cognitive and emotional levels, particularly their self-efficacy. All the feedback students receive on their contributions provides a sense of relative strengths and relative weaknesses. For example, if students receive many negative comments on one assignment, their self-efficacy for the general topic (or even course overall) may decline. Alternatively, if the student is struggling to complete an assignment or notices that their contribution is noticeably weaker than those of other peers, they may have low self-efficacy for that specific topic. Past work on student peer review has found that when students review multiple papers by many strong peers, they are more likely to drop out of the course (Ekholm, Zumbrunn, and Conklin 2015).

At the behavioral regulation level, there can be norm-setting (Flower et al. 1986; Abubakar et al. 2019) and reinforcement feedback (Daniels 2016). Extensive organizational behavior management research has shown that feedback about performance significantly impacts participants' behavioral changes (Teig, Scherer, and Nilsen 2018; Abubakar et al. 2019). In general, organizations often rely on peer feedback to improve the performance of organizational members (VanStelle et al. 2012; Wang et al. 2017; Fernandes et al. 2019). For the current research focus, what influences the amount of feedback provided, students can learn how much feedback is expected to be provided through the amount of feedback they receive from others (i.e. as a kind of norm-setting) or from positive evaluations they receive from their peers for the feedback they provided (i.e. reinforcement feedback). Note, however, that norms tend to be set early on within a context (Kaufman and Schunn 2011), and thus likely play little role in changes in the amount of feedback provided from one assignment to the next. Instead, the amount of feedback received might predominantly signal strengths and weaknesses in the submitted documents.

In sum, we hypothesize that students learn to regulate the amount of feedback they provide through three sources: 1) the amount of feedback they receive, 2) success in initially completing the assignment being reviewed, and 3) reinforcement messages regarding the value of the prior feedback provided to their peers. To testify these hypotheses, we conduct analyses of gradual growth and decline in the amount of feedback provided from one assignment to the next as a function of the experiences in the prior assignment.

Methods

Course setting and participants

Peer feedback from 149 undergraduate students was examined (43% female; predominantly second and third years) in two different programming courses, *Object-Oriented Programming in Java* and *C Programming,* at a research-oriented national university in China. As introductory courses, the participating students are considered novices. These courses utilized the same online peer assessment system, EduPCR.

The courses involved 12 programming assignments, and students had to complete 10 of them. As a result, students rarely completed the 12th assignment, so we only focus on the first 11 assignments. The main task within each assignment was moderately difficult and scheduled to be completed by specific dates, typically five-to-seven days apart.

Materials

Students did all reviewing activities via a web-based, online peer assessment system. The courses used EduPCR, which is designed specifically for programming-based assignments. The system has many standard features: 1) *asynchronous reviewing* (i.e. via forms, not interactive), 2) *reviewing including numeric evaluation as well as open-response feedback comments*, 3) *double-blind reviewing* (i.e. reviewers and authors are anonymous), 4) *multiple reviews per document* (i.e. 4 to five reviews per document), and 5) *mechanisms in the system to encourage more constructive suggestions* (i.e. authors judge whether the feedback was helpful, which produces a grade for the reviewer; Patchan, Schunn, and Clark 2018). These standard features are found in many online peer assessment systems (e.g. Calibrated Peer Review, Chapman and Fiore 2000; ELI Peer Review, McCarthy et al. 2011). Most saliently, for the current study, the students had the option of submitting a variable amount of open-ended comments within each review, and the total length of comments provided is the critical outcome variable in the current study. Figure 1 presents the EduPCR student interface for reviewing a peer's task and being recognized for helpful reviews.

Measures

Based entirely upon data automatically collected within EduPCR, the measures were defined as follows (see Table 1 for a summary and Table 2 for means, standard deviations (SDs) and maximums of each variable). The data used in the analyses are organized by assignment (e.g. the total amount of feedback provided on one assignment across reviews). Thus, the dataset size is the number of students multiplied by the number of assignments in the course minus one (because the first reviewing assignment serves as the baseline). Because the courses were relatively homogeneous (similar assignment types, similar reviewing focus, similar student populations) and the courses showed identical patterns in pilot data analyses, the data across courses were combined into one dataset, producing 1,490 total data points.

The outcome variable was defined based on the reviewing behavior on the J^{th} assignment, where J is the assignment number. Predictor variables were defined in terms of the $(J\text{-}1)^{th}$ assignment's reviewing behaviors and the J^{th} assignment's submission performance.

Provided feedback

The primary outcome variable is the total length (in words) of comments provided to peers on their documents in a given round. There was no mandatory number of comments or word length within each comment, and thus students could substantially vary in how many comments and how much depth they provided in their reviews. For the time-series analysis, $LengthPrvd_J$ refers to the total length (in words) of comments provided on the J^{th} assignment, and $LengthPrvd_{J\text{-}1}$

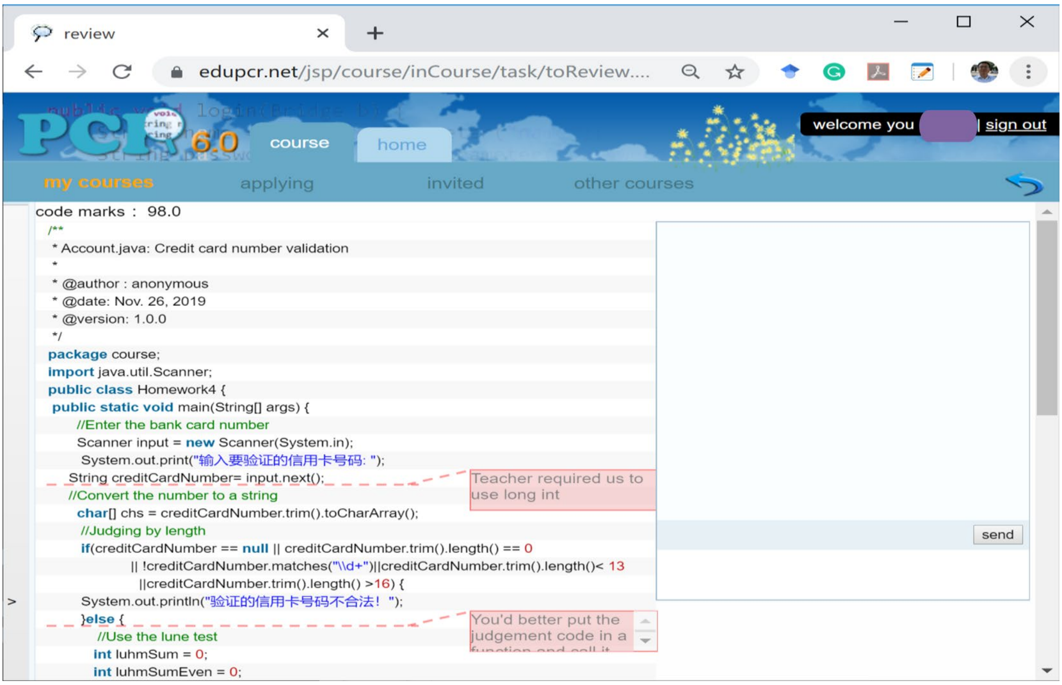

Figure 1. An example of the reviewer's interface within the *EduPCR* platform. Reviewers can submit and edit feedback suggestions through the textbox on the right. Red text boxes show feedback suggestions that the reviewer has already made.

Table 1. Definitions of outcome and predictor variables.

Construct	Variable	Definition
The length of feedback comments provided	$LengthPrvd_J$	The total length (in words) of peer feedback comments provided by a student across reviews on the J^{th} assignment
The length of feedback comments provided	$LengthRcvd_J$	The total length (in words) of peer feedback comments received from other students on the J^{th} assignment
Task Performance	$Score_J$	The standardized score of the student's task on the J^{th} assignment (i.e. score - mean assignment grade / SD of assignment grade)
Recognition received	$Recognition_J$	The total amount of praise feedback received by a student for their reviewing on the J^{th} assignment
Task Submitted	$Submitted_J$	1 if the student submitted a document on the J^{th} assignment, 0 otherwise
Assignment number	*Round*	The sequential number of the assignment requiring peer reviews in the course
Course	*Course*	Course dummy code with 2 = Elective, 1 = Required

Table 2. Pearson intercorrelations among predictors and the outcome variable across *J* (from 2 to 11).

	Mean	SD	Max	$LengthPrvd_{J-1}$	$LengthRcvd_{J-1}$	$Score_{J-1}$	$Recog\text{-}nition_{J-1}$	$Submitted_{J-1}$	Round
$LengthPrvd_{J-1}$	82.5	99.4	635						
$LengthRcvd_{J-1}$	68.2	66.6	495	.14***					
$Score_{J-1}$	0.0	1.0	1.99	.18***	-.07**				
$Recognition_{J-1}$	0.1	0.3	3.0	.15***	.04	.05*			
$Submitted_J$	0.9	0.2	1.0	.06**	.09***	.03	.04		
Round	6.5	2.9	11	-.01	-.001	-.02	−.09***	-.02	
$LengthPrvd_J$	80.0	96.7	635	.57***	.14***	.15***	.09***	.05*	-.10***

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

refers to the same measure on the $(J\text{-}1)^{th}$ assignment. $LengthPrvd_J$ is the primary dependent variable, and $LengthPrvd_{J\text{-}1}$ serves as a baseline measure for deriving growth or decline in the amount of feedback provided in response to the prior assignment's experiences with peer feedback.

Received feedback

In a complementary way, students also received a varying amount of feedback on a given assignment submission. Variations in the length and number of feedback comments received might reflect the conscientiousness of the reviewers and the quality of the submission. $LengthRcvd_{J\text{-}1}$ refers to the total length (in words) of comments received across reviews on the $(J\text{-}1)^{th}$ assignment. If no document was submitted, the value was treated as missing.

Recognition for helpful feedback

The system had a mechanism by which authors recognize the helpfulness of the comments received, which can be thought of as direct performance feedback on the provided peer comments. Specifically, authors could nominate in EduPCR some reviews as particularly helpful, and the sum of such nominations in the J^{th} round was calculated (range 0 to 3). This measure for the experience on the prior assignment was called $Recognition_{J\text{-}1}$.

Submitted assignments

Students may not provide much feedback if they have low self-efficacy for the current assignment. Although not directly assessed by the system, there are two indirect measures of students having low self-efficacy on the current assignment. The first indicator is that students did not complete the current assignment and thus would feel poorly positioned to comment on other students' submissions for that assignment. $Submitted_J$ refers to whether the student submitted for the J^{th} assignment (one if true, 0 if false).

Assignment score

The second self-efficacy indicator is that the students struggled with the performance of the prior assignment. Prior research has shown that the ratings generated by multiple peer reviews are reliable and valid scores (Li et al. 2016). The previous assignment undoubtedly affects the students' sense of self-efficacy during the current review. Taking into account differences in the difficulty of assignments, we normalized the ratings (subtract assignment mean score and divide by assignment SD) to produce the variable $Score_J$.

Assignment round

To account for general temporal trends in reviewing behaviors, the variable J was the reviewing assignment number (2 to 11 in both courses).

Course

It accounts for small differences in mean values in the variables across the two courses within each context. A *Course* indicator variable was created: set to 0 for the first course and one for the second course.

Analyses

The closest non-outlier value replaced outlier values in each continuous predictor variable. This situation denotes at most 0.4% of values on any given variable and less than 0.1% of values on most variables.

The overall analytic approach was to use multiple regression, with $LengthPrvd_J$ as the dependent variable, $LengthPrvd_{J\text{-}1}$ as the baseline control (Barnett, van der Pols, and Dobson 2005), *LengthRcvd*$_{J-1}$, *Recognition*$_{J-1}$, *Submitted* $_{J}$and *Score*$_{J-1}$ as core predictors, and *Round* and *Course* as additional control variables.

Since the dependent variable, the total amount of feedback comments provided by a student is a count variable, it has a non-normal distribution with a large right (positive) skew in the data. Thus, traditional linear regression is not the best modeling method. Instead, Poisson regression or negative binomial regression distribution is recommended for this kind of data, depending upon the relationship between the mean and the variance (Grogger and Carson 1991; Gardner, Mulvey, and Shaw 1995). Since the variance for the total length of provided comments was generally larger than its mean (see Table 2), it was likely that negative binomial regression would be the better choice. Based on these considerations and an examination of model fit statistics (Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC)), negative binomial regression was selected as the primary modeling approach. Pseudo-R^2 is reported to show model quality.

Results

The Pearson correlations among the predictors are also shown in Table 2. None of the core predictors were more than moderately correlated with one another, suggesting there would be no multicollinearity problems. As expected, an examination of the variance inflation factors (VIF) also suggested that multicollinearity is not a significant problem.

However, all core predictors significantly correlated with the baseline measure (*LengthPrvd*$_{J-1}$). Thus, multiple regression controlling for *LengthPrvd*$_{J-1}$ is needed to uncover the relationship of each predictor to relative growth or decline in the amount of feedback provided. Table 3 presents the results of the negative binomial regression models. The number of comments provided on the prior assignment was a good baseline predictor.

Turning to the core predictors, the total length of comments received and the score on the prior assignment were significant predictors in the base model. In particular, students who received more extensive comments and students who performed well on the prior assignment tended to provide extensive comments to their peers in the current assignment. Figure 2

Table 3. Estimated coefficients from the negative binomial regressions predicting *LengthPrvd*$_J$, along with *N* and fit statistics.

Predictor	Main effects only	Length received interaction	Score interaction	Round interaction
Baseline				
LengthPrvd$_{J-1}$	0.58***	0.58***	0.58***	0.58***
Core Predictors				
LengthRcvd$_{J-1}$	0.10*	0.26	0.11*	0.10*
Score$_{J-1}$	0.14***	0.10*	−0.05	0.14***
Recognition$_{J-1}$	0.02	0.10	0.03	−0.27
Submitted$_J$	−0.07	−0.11	−0.07	−1.03
Control variables				
Course (Required = 1; Elective = 2)	−0.18**	−0.11	−0.17*	−0.18
Round	−0.05***	−0.05***	−0.05***	−0.20*
Interaction predictors				
LengthRcvd$_{J-1}$* *Recognition*$_{J-1}$		−0.11		
LengthRcvd$_{J-1}$* *Score*$_{J-1}$		0.05		
LengthRcvd$_{J-1}$* *Course*		−0.11		
Score$_{J-1}$* *LengthRcvd*$_{J-1}$			0.08	
Score$_{J-1}$* *Recognition*$_{J-1}$			−0.16*	
Score$_{J-1}$* *Course*			0.10	
Round * *Submitted*$_J$				0.15
Round * *Recognition*$_{J-1}$				0.05
N	1,490	1,490	1,490	1,490
Pseudo R^2	0.02	0.02	0.02	0.02

Notes. ***=p<.001, **=p<.01, *=p<.05. Coefficients for length variables are multiplied by 100.

presents a visualization of these results. For the x-axis of this graph, the range of values of core predictors was grouped into bins based on visual inspection of the distributions. Only the relationship with the prior score was substantial.

Interestingly, although recognition for good reviewing did not generally relate to the total length of comments provided (see Table 3), the interaction predictor of the recognition for good reviewing and the score on the prior assignment was negatively associated with the total length of comments provided on the later assignment. Figure 3 presents a visualization of these results. For the x-axis of this graph, the range of values of $Recognition_{J\text{-}1}$ was grouped into bins based on visual inspection of the distribution. Given the relatively low number of recognitions for high quality reviewing (i.e. most received no recognition on a given assignment, very few received more than one recognition), only two bins are used. Note that this visualization presents the estimated means after controlling for the other predictors. For those who received low ratings on their prior assignment, recognition for high quality reviewing was associated with more extensive comments in the next assignment. By contrast, for those who had high average rates on the prior assignment, there were no relationships of commenting with recognition.

For the control variables (see Table 3), the required course produced more comments than did the elective course. Across assignments, students gradually produced less feedback, which may indicate decreased motivation towards the end of the semester.

Discussion

This study aimed to uncover the relationship between peer feedback given and provided. The study explored these findings in two different course contexts (required versus elective) to uncover patterns that are, therefore, more likely to be generally observed across contexts.

Two significant relationships were found to generalize across this context: 1) students provided more extensive comments when they had performed well on the prior assignment, and 2)

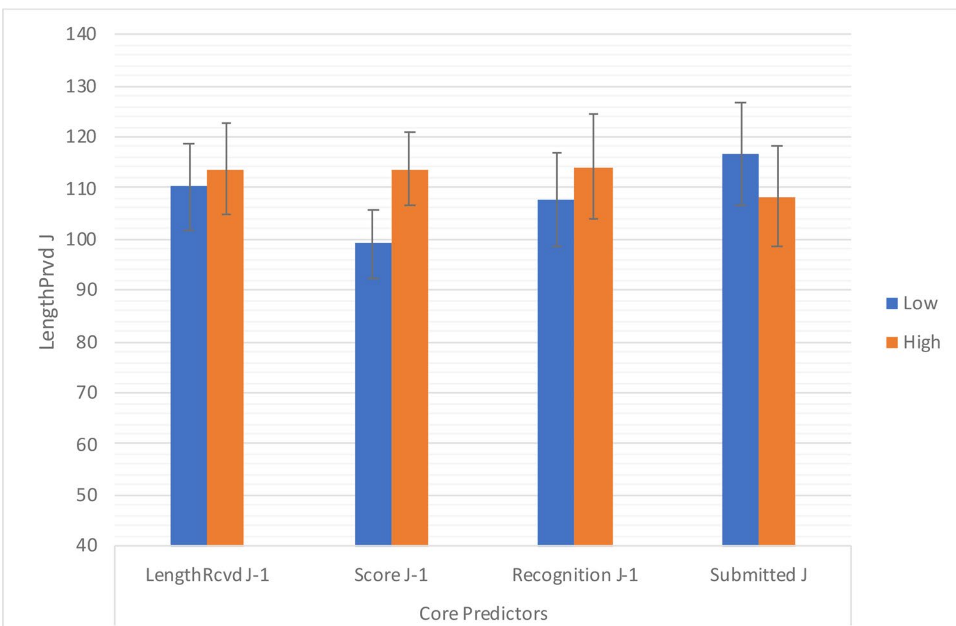

Figure 2. The estimated marginal mean of the total word length of comments provided on the J^{th} assignment (with standard error bars) as a function of length comments received on the prior assignment, whether the prior assignment submission received a high task score, whether the reviewing on the prior assignment received praise, and whether the current assignment had a submitted assignment or not.

students provide more extensive comments when they received more extensive comments. Pattern one points toward self-efficacy issues as a central factor in the decision not to provide any comments. Self-efficacy has been raised in interviews and survey studies of peer assessment (Liu and Carless 2006; Kaufman and Schunn 2011). The current study adds behavioral indicators to studies of this effect. Pattern two suggests that a norm-setting effect was supported (i.e. students did provide more feedback after receiving more feedback).

In addition, there was partial support for a positive relationship between recognition for good reviewing and the amount of provided feedback, but only for students who had performed poorly in the past assignment. This interesting finding shows that recognition is more effective for underperforming students. A possible reason is that both performance and recognition are factors that affect self-efficacy. Students with good grades are already likely to have higher self-efficacy. In contrast, students with low grades often need encouragement from their peers. Indeed, interviews with students about peer feedback suggest that lower-performing students often doubt that they have something useful to contribute to their peers (Kaufman and Schunn 2011).

Practical implications

Peer reviewing has proven not only to assess student performance effectively but also to provide a robust platform for students to learn from one another. Nevertheless, the willingness to participate at all (Liu and Carless 2006; Huisman et al. 2018) or with substantial levels of feedback (Cho and MacArthur 2011; Tsivitanidou et al. 2018) has proven to be a common bottleneck. With online peer assessment, interfaces can be adapted to increase the amount of feedback that students give (e.g. through minimum word counts in Calibrated Peer Review or the minimum number of comments in *Peerceptiv*), but still, other strategies for further improving participation are needed.

The current research, while not conclusive, does provide some suggestions for directions that are more likely to be productive. It is doubtful that rewards received for good reviewing will change the length of comments provided in later reviewing assignments. Allowing for the possibility of rewards for reviewing can increases the quality of comments/length of comments (Zou et al. 2018); however, this may not depend upon a student receiving a reward at a given

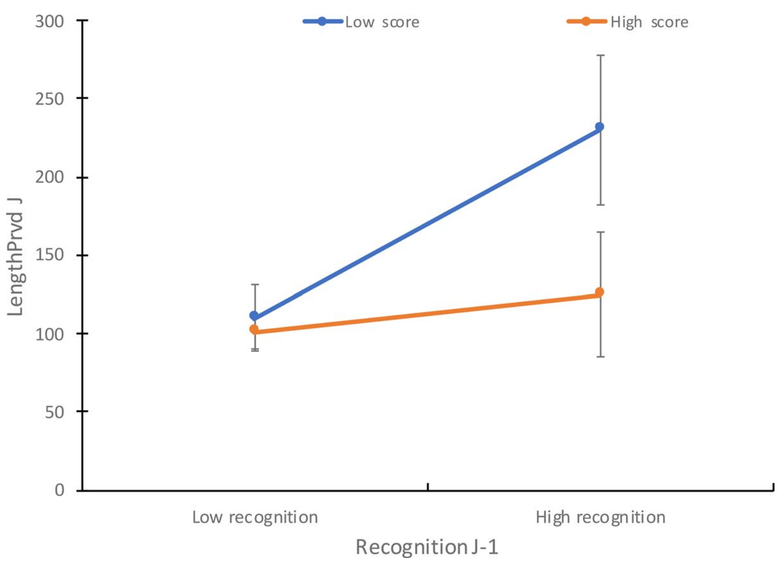

Figure 3. The estimated marginal mean (with standard error bars) of the total word length of comments provided on the J^{th} assignment (controlling for length provided on the prior assignment and other covariates) as a function of the interaction of task score and recognition on the prior assignment.

moment in time (i.e. the difference between an anticipated reward and an actual reward; Davies 2009). Instead, the current study suggests that student self-efficacy is a crucial factor to target. Thus, when students are lagging in producing substantial reviews or just a small amount of reviews, instructors might intervene with comments about the effectiveness of the comments they did produce in this round or prior rounds.

Other possible strategies might involve: (1) moderately increasing the time interval between assignments to improve students' negative emotions caused by poor performance in previous assignments, or (2) avoiding the broad effect of fatigue from too many peer assessment assignments, as in the contexts that had many assignments and a gradual increase in the amount of students providing no comments. Another strategy might involve small reductions in the difficulty of assignments to reduce students' self-efficacy concerns for these assignments, thereby increasing the feedback provided by students. A final strategy might encourage students to more commonly express gratitude for the feedback received.

Limitations and future research

The current study was fundamentally correlational in its examination of the role of experience in shaping the amount of feedback provided. Therefore, strong causal claims cannot be made based on the current data. However, as an initial exploration of a novel research topic, the study has provided evidence in an externally valid way (with a real interface and real classroom assignments) of potentially important empirical phenomena. Further, by examining change over time with various statistical controls, reverse causality and obvious third-variable confounds have been ruled out (e.g. general differences across reviewers in the amount of feedback provided), at least for the role of experience factors. However, future studies should be conducted that more carefully control student experiences (e.g. by experimentally manipulating recognition or relative performance feedback) to directly test the causal status of these factors.

Another open question relates to the underlying causes. We have posited that self-efficacy could explain the patterns that were observed in the current data, but self-efficacy was not directly measured. It would be challenging to measure self-efficacy across many assignments in an actual classroom regularly, but the role of feedback on self-efficacy could be tested on a smaller number of assignments or within an experimental study.

A third open question relates to more specific dimensions of feedback quality. Here we have focused on the amount of feedback provided in terms of the total number of comment words. Comments can vary substantially in terms of their depth in specific ways: with or without explicitly identifying the problem, explaining the situation, providing a constructive suggestion, explaining the constructive suggestion (Wu and Schunn 2020). These comment elements also shape the learning opportunities for feedback providers and receivers (Cho and Schunn 2007; Zou et al. 2018). Feedback experience could shape whether and how much reviewers include these specific elements in their comments.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributors

Zheng Zong, a doctor candidate in the School of Management, Harbin Institute of Technology, China. His research interests included peer assessment, text mining, education management and natural language processing.

Dr. Christian D. Schunn, 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.

Dr. Yanqing Wang, 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 integer linear programming.

ORCID

Zheng Zong http://orcid.org/0000-0002-2421-1545
Christian Schunn http://orcid.org/0000-0003-3589-297X
Yanqing Wang http://orcid.org/0000-0002-5961-0482

References

Abubakar, A. M., H. Elrehail, M. A. Alatailat, and A. Elçi. 2019. "Knowledge Management, Decision-Making Style and Organizational Performance." *Journal of Innovation & Knowledge* 4 (2): 104–114. doi:10.1016/j.jik.2017.07.003.

Adachi, C., J. H. M. Tai, and P. Dawson. 2018. "Academics' Perceptions of the Benefits and Challenges of Self and Peer Assessment in Higher Education." *Assessment & Evaluation in Higher Education* 43 (2): 294–306. doi:10.1080/02602938.2017.1339775.

Barnett, A. G., J. C. van der Pols, and A. J. Dobson. 2005. "Regression to the Mean: What It is and How to Deal with It." *International Journal of Epidemiology* 34 (1): 215–220. doi:10.1093/ije/dyh299.

Bouzidi, L., and A. Jaillet. 2009. "Can Online Peer Assessment Be Trusted?" *Educational Technology & Society* 12 (4): 257–268.

Bruce, R., A. Chauvin, L. Trinquart, P. Ravaud, and I. Boutron. 2016. "Impact of Interventions to Improve the Quality of Peer Review of Biomedical Journals: A Systematic Review and Meta-Analysis." *BMC Medicine* 14 (1): 1–16. doi:10.1186/s12916-016-0631-5.

Chapman, O. L., and M. A. Fiore. 2000. "Calibrated Peer Review™." *Journal of Interactive Instruction Development* 12 (3): 11–15.

Cho, K., and C. MacArthur. 2011. "Learning by Reviewing." *Journal of Educational Psychology* 103 (1): 73–84. doi:10.1037/a0021950.

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. doi:10.1016/j.compedu.2005.02.004.

Cho, K., and C. D. Schunn. 2018. "Finding an Optimal Balance between Agreement and Performance in an Online Reciprocal Peer Evaluation System." *Studies in Educational Evaluation* 56: 94–101. doi:10.1016/j.stueduc.2017.12.001.

Crinon, J. 2012. "The Dynamics of Writing and Peer Review at Primary School." *Journal of Writing Research* 4 (2): 121–154. doi:10.17239/jowr-2012.04.02.2.

Daniels, A. C. 2016. *Bringing out the Best in People: How to Apply the Astonishing Power of Positive Reinforcement*. New York: McGraw Hill Professional.

Davies, P. 2009. "Review and Reward Within the Computerised Peer-Assessment of Essays." *Assessment & Evaluation in Higher Education* 34 (3): 321–333. doi:10.1080/02602930802071072.

Deiglmayr, A. 2018. "Instructional Scaffolds for Learning from Formative Peer Assessment: Effects of Core Task, Peer Feedback, and Dialogue." *European Journal of Psychology of Education* 33 (1): 185–198. doi:10.1007/s10212-017-0355-8.

Ekholm, E.,. S. Zumbrunn, and S. Conklin. 2015. "The Relation of College Student Self-Efficacy toward Writing and Writing Self-Regulation Aptitude: Writing Feedback Perceptions as a Mediating Variable." *Teaching in Higher Education* 20 (2): 197–207. doi:10.1080/13562517.2014.974026.

Feng, S., Y. K. Wong, L. Y. Wong, and L. Hossain. 2019. "The Internet and Facebook Usage on Academic Distraction of College Students." *Computers & Education* 134: 41–49. doi:10.1016/j.compedu.2019.02.005.

Fernandes, E., M. Holanda, M. Victorino, V. Borges, R. Carvalho, and G. Van Erven. 2019. "Educational Data Mining: Predictive Analysis of Academic Performance of Public School Students in the Capital of Brazil." *Journal of Business Research* 94: 335–343. doi:10.1016/j.jbusres.2018.02.012.

Flower, L., J. R. Hayes, L. Carey, K. Schriver, and J. Stratman. 1986. “Detection, Diagnosis, and the Strategies of Revision.” *College Composition and Communication* 37 (1): 16–55. doi:10.2307/357381.
Gardner, W., E. P. Mulvey, and E. C. Shaw. 1995. “Regression Analyses of Counts and Rates: Poisson, Overdispersed Poisson, and Negative Binomial Models.” *Psychological Bulletin* 118 (3): 392–404. doi:10.1037/0033-2909.118.3.392.
Grogger, J. T., and R. T. Carson. 1991. “Models for Truncated Counts.” *Journal of Applied Econometrics* 6 (3): 225–238. doi:10.1002/jae.3950060302.
Huisman, B., N. Saab, J. van Driel, and P. van den Broek. 2018. “Peer Feedback on Academic Writing: Undergraduate Students’ Peer Feedback Role, Peer Feedback Perceptions and Essay Performance.” *Assessment & Evaluation in Higher Education* 43 (6): 955–968. doi:10.1080/02602938.2018.1424318.
Huisman, B., N. Saab, P. van den Broek, and J. van Driel. 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. doi:10.1080/02602938.2018.1545896
Kankanhalli, A., B. C. Tan, and K. K. Wei. 2005. “Contributing Knowledge to Electronic Knowledge Repositories: An Empirical Investigation.” *MIS Quarterly* 29 (1): 113–143.
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. doi:10.1007/s11251-010-9133-6.
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. doi:10.1080/02602938.2014.999746.
Liu, N. F., and D. Carless. 2006. “Peer Feedback: The Learning Element of Peer Assessment.” *Teaching in Higher Education* 11 (3): 279–290. doi:10.1080/13562510600680582.
Lundstrom, K., and W. Baker. 2009. “To Give is Better than to Receive: The Benefits of Peer Review to the Reviewer’s Own Writing.” *Journal of Second Language Writing* 18 (1): 30–43. doi:10.1016/j.jslw.2008.06.002.
Martin, A. J., and P. Evans. 2018. “Load Reduction Instruction: Exploring a Framework That Assesses Explicit Instruction through to Independent Learning.” *Teaching and Teacher Education* 73: 203–214. doi:10.1016/j.tate.2018.03.018.
McCarthy, J. E., J. T. Grabill, W. Hart-Davidson, and M. McLeod. 2011. “Content Management in the Workplace: Community, Context, and a New Way to Organize Writing.” *Journal of Business and Technical Communication* 25 (4): 367–395. doi:10.1177/1050651911410943.
Nelson, M. M., and C. D. Schunn. 2009. “The Nature of Feedback: How Different Types of Peer Feedback Affect Writing Performance.” *Instructional Science* 37 (4): 375–401. doi:10.1007/s11251-008-9053-x.
Patchan, M. M., C. D. Schunn, and R. J. Clark. 2018. “Accountability in Peer Assessment: Examining the Effects of Reviewing Grades on Peer Ratings and Peer Feedback.” *Studies in Higher Education* 43 (12): 2263–2278. doi:10.1080/03075079.2017.1320374.
Popp, J. S., and S. R. Goldman. 2016. “Knowledge Building in Teacher Professional Learning Communities: Focus of Meeting Matters.” *Teaching and Teacher Education* 59: 347–359. doi:10.1016/j.tate.2016.06.007.
Rotsaert, T., E. Panadero, and T. Schellens. 2018. “Anonymity as an Instructional Scaffold in Peer Assessment: Its Effects on Peer Feedback Quality and Evolution in Students’ Perceptions about Peer Assessment Skills.” *European Journal of Psychology of Education* 33 (1): 75–99. doi:10.1007/s10212-017-0339-8.
Teig, N., R. Scherer, and T. Nilsen. 2018. “More Isn’t Always Better: The Curvilinear Relationship between Inquiry-Based Teaching and Student Achievement in Science.” *Learning and Instruction* 56: 20–29. doi:10.1016/j.learninstruc.2018.02.006.
Tsivitanidou, O. E., C. P. Constantinou, P. Labudde, S. Rönnebeck, and M. Ropohl. 2018. “Reciprocal Peer Assessment as a Learning Tool for Secondary School Students in Modeling-Based Learning.” *European Journal of Psychology of Education* 33 (1): 51–73. doi:10.1007/s10212-017-0341-1.
VanStelle, S. E., S. M. Vicars, V. Harr, C. F. Miguel, J. L. Koerber, R. Kazbour, and J. Austin. 2012. “The Publication History of the Journal of Organizational Behavior Management: An Objective Review and Analysis: 1998–2009.” *Journal of Organizational Behavior Management* 32 (2): 93–123. doi:10.1080/01608061.2012.675864.
Wang, X. M., G. J. Hwang, Z. Y. Liang, and H. Y. Wang. 2017. “Enhancing Students’ Computer Programming Performances, Critical Thinking Awareness and Attitudes Towards Programming: An Online Peer-Assessment Attempt.” *Journal of Educational Technology & Society* 20 (4): 58–68.

Wu, Y., and C. D. Schunn. 2020. “From Feedback to Revisions: Effects of Feedback Features and Perceptions.” *Contemporary Educational Psychology* 60: 101826. doi:10.1016/j.cedpsych.2019.101826.

Wu, Y., and C. D. Schunn. 2021. “The Effects of Providing and Receiving Peer Feedback on Writing Performance and Learning of Secondary School Students.” *American Educational Research Journal* 58 (3): 492–526. doi:10.3102/0002831220945266.

Zimmerman, B. J., and R. Risemberg. 1997. “Caveats and Recommendations about Self-Regulation of Writing: A Social Cognitive Rejoinder.” *Contemporary Educational Psychology* 22 (1): 115–122. doi:10.1006/ceps.1997.0921.

Zong, Z.,. C. D. Schunn, and Y. Wang. 2021. “What Aspects of Online Peer Feedback Robustly Predict Growth in Students’ Task Performance?” *Computers in Human Behavior* 124: 106924. doi:10.1016/j.chb.2021.106924.

Zou, M., C. Schunn, Y. Wang, and F. Zhang. 2018. “Student Attitudes That Predict Participation in Peer Assessment.” *Assessment & Evaluation in Higher Education* 43 (5): 800–811. doi:10.1080/02602938.2017.1409872.