A Comparison of Anonymous Versus Identifiable e-Peer Review on College Student Writing Performance and the Extent of Critical Feedback

Ruiling Lu
Linda Bol
Old Dominion University

Abstract

Peer review has become commonplace in composition courses and is increasingly employed in the context of telecommunication technology. The purpose of this experiment was to compare the effects of anonymous and identifiable electronic peer (e-peer) review on college student writing performance and the extent of critical peer feedback. Participants were 92 undergraduate freshmen in four English composition classes enrolled in the fall semesters of 2003 and 2004. The same instructor taught all four classes, and in each semester, one class was assigned to the anonymous e-peer review group and the other to the identifiable e-peer review group. All other elements—course content, assignments, demands, and classroom instruction—were held constant. The results from both semesters showed that students participating in anonymous e-peer review performed better on the writing performance task and provided more critical feedback to their peers than did students participating in the identifiable e-peer review.

Introduction

Writing is recognized as a key factor in students' academic development, and it is the primary means by which students transform from passive recipients of knowledge to active participants in their own education (Harvard's Expository Writing Program, 2002). Unfortunately, in spite of the inarguable importance of writing in student life, a large number of American students present inadequate writing skills (Baker, Gersten, & Graham, 2003; National Center for Education Statistics, 1998; Richardson, 2001). In an attempt to discover more effective ways to help students become better writers, educators and researchers have been exploring and experimenting with various instructional strategies, and peer review—students critiquing one another's work with the intention of helping their peers revise and eventually improve their writing—is one such endeavor.

Merits and Drawbacks of Peer Review

Strong support for peer review has come from theories that emphasize the social nature of language, thought, writing, and learning (Zhu, 1995). According to social constructivism, learning "requires exchanging, sharing, and negotiation," and it "involves both personal inner process and social aspect" (Liu, Lin, Chiu, & Yuan, 2001, p. 247). The Vygotskyan (1962, 1978) perspective on learning a language strongly supports the use of peer review. For Vygotsky, learning is not an individual, secluded activity, but rather a cognitive activity that occurs in, and is mediated by, social interaction. Thus, on the theoretical level, peer interaction is vital to writing development because it allows students to construct knowledge through social sharing and interaction (Liu et al.). Peer review is also built on the notion of collaboration, which assumes that learning emerges through shared understandings of multiple learners, and that learning effectively occurs within interactive peer groups (Asberg & Nulden, n.d.; Leidner & Jarvenpaa, 1995).

Various reasons account for the popularity of using peer review in writing classes. First, many students like it. Student responses to a posttest questionnaire revealed that nearly 70 percent of the participants preferred using peer review for writing assignments (Liu et al., 2001), and 11 out of 12 students in Eschenbach's (2001) class recommended in the end-of-semester survey that peer review should be continued. The results of a post-assessment survey by Li and Steckelberg (2004) also indicated that students had generally positive perceptions of the peer review process. Second, peer review reduces the workload of instructors. Circumstances of worsening staff-student ratios make it increasingly difficult for instructors to provide sufficient, detailed, individualized, and timely feedback on student work. Some instructors use peer review as substitutional or supplementary to staff assessment (Topping, 1998). Third, students plan more extensively and write more carefully when they are communicating with an audience of peers than when they are evaluated solely by instructors because of peer pressure (Bagley & Hunter, 1992; Eschenbach, 2001; Kerr, Park, & Domazlicky, 1995; Lin, Lui, Chiu, & Yuan, 1999; Shaw, 2002).

Finally and most importantly, peer feedback seems to be associated with better student academic achievement. Richer (1992) compared the effects of peer feedback with teacher feedback on college students' writing proficiency. The pre/post measures of students' essays revealed that greater gains in writing proficiency were obtained by the peer feedback group. Ramsden (1992) found that students could often learn more from formal or informal assessment by their peers. A quasi-experimental study by Plutsky and Wilson (2004) also revealed that peer review helped students become proficient writers. Studies on e-peer review also support its academic benefits. Liu et al. (2001) surveyed a group of computer science students at a Taiwanese university and found that students performed better on writing tasks under a web-based peer review system. Along the same lines, the survey results from MacLeod (1999) indicated that students felt peer reviews were effective in improving their writing skills. Results from empirical studies on English as a Second Language (ESL) writing instruction also suggested that peer feedback was as good as or better than teacher feedback in helping revise and improve students' papers (Chaudron, 1983; Mangelsdorf, 1992; Paulus, 1999).

Despite the potential benefits of students' participation in peer review, there remain some concerns with the quality of peer feedback. The literature suggests that the biggest problem with peer review is that students are easily biased or not honest in giving feedback because of friendship, gender, race, interpersonal relationships, or personal preferences (Carson & Nelson, 1996; Ghorpade & Lackritz, 2001; MacLeod, 1999; Nilson, 2003; Zhao, 1998). Students conducting face-to-face reviews frequently express anxiety in sharing their writing for fear of being wrong or rejected by peers (Zhao). Students find it extremely difficult to give negative feedback to classmates, especially friends because they hate to hurt others' feelings or damage personal relationships (MacLeod; Schaffer, 1996; Topping, 1998). Another area of concern with peer review is the uneven quality of feedback. Kerr et al. (1995) and Robinson (1999) found that students with better writing ability were better at evaluating peers' writings. As a result, students often hesitate to take peers' feedback seriously when they know their peers are less capable writers than themselves, even if the comments are correct (Quible, 1997). Difficulties with group functioning occur as well. This is problematic as many peer review activities are organized in groups. For example, students often complain about the irresponsibility or procrastination of their group members (Dyrud, 2001), and in some other cases, students are more interested in maintaining positive group relations than in helping one another with writing during group interactions (Carson & Nelson).

To alleviate these concerns, many researchers advocate instructional approaches to increase the quality of peer feedback that include: 1) using electronic communication to avoid

the possible embarrassment students may experience in face-to-face interaction (Eisenberg, 1993; Mabrito, 1991; Siegel, Dubrovsky, Kiesler, & McGuire, 1986), 2) using multiple reviewers for any single piece of writing to balance the uneven quality of peer feedback (Kerr et al., 1995; Nilson, 2003; Quible, 1997; Robinson, 1999), and 3) using anonymous peer review to minimize opportunities for students to reward friends or otherwise cheat during peer review process (Haaga, 1993; Kerr et al.; Vinson, 1996; Zhao, 1998). The present study takes a comprehensive approach to all three methods with a primary focus on the effect of anonymity while controlling for the other two variables (electronic communication and multiple reviewers).

Anonymous Peer Review

Anonymous peer review refers to a kind of peer review condition in which both reviewers and reviewees are kept unknown to one another. Anonymity is an important component of deindividuation (Connally, Jessup, & Valacich, 1990; Zimbardo, 1969). According to social psychological research, deindividuation refers to situations in which individuals in groups "stop thinking of other members as individuals and feel that they cannot be singled out by others" (Jessup, Connolly, & Tansik, 1990, p. 338). This deindividuation results in a reduction of normal inner restraints and enables group members to engage in behavior that they would not ordinarily display (Jessup et al.). Anonymous peer review, therefore, is predicted to foster deindividuated behavior because the assessment system separates people from one another and detaches individuals from their contributions.

The literature suggests that the major advantage of anonymous peer review is that it provokes more critical feedback because reviewers are relieved from the social pressure and enabled to express themselves freely without considering interpersonal factors (Bornstein, 1993; Bostock, 2000; Liu et al., 2001; Valacich, Dennis, & Nunamaker, 1992; Zhao, 1998). Some argue that constructively critical feedback is more useful in helping students improve their work (Falchikov, 1996). Zhu (1995) argues that if students do not approach their peers' writing critically, they will fail to provide meaningful and useful feedback. A related benefit observed by Kerr et al. (1995) is that students who take a critical approach when reading and scoring peers' work are likely to be more critical of their own work, and thus create improved products.

Several studies support the contention that in anonymous situations, people are more honest and less anxious in expressing their opinions and that anonymity seems to encourage more critical feedback (MacLeod, 1999; Robinson, 1999; Stone, Spool, & Rabinowitz, 1977; Zhao, 1998). Zhao conducted two studies with college freshmen and sophomores, exploring the effects of anonymity on peer feedback. In each study, the participants were asked to review their peers' journal entries in two conditions: anonymous and identifiable. In the anonymous condition, the identification of both reviewers and authors were removed so that the reviewers were not aware of the authorship of the writings they were reviewing, and they were assured that their reviews would be anonymous to the authors. In the identifiable condition, the reviewers were aware of the authorship of the journals they were reviewing and knew that their reviews would be given to the authors with their names attached. Results from the three indicators—the grades assigned to each journal by peer reviewers, the degree of the overall critical nature of peer comments as rated by two experts, and the recipients' perceptions of the degree of negativeness or rudeness of the peer feedback they received—confirmed his predictions that the reviews provided in the anonymous condition were more critical than those made in the identifiable condition.

The work by Valacich, Dennis, and Nunamaker (1992) showed similar results. They tested the effects of group size (three and nine members) and group member anonymity on group idea generation using a computer-mediated idea-generation system. A significant main effect and

interaction were reported for the amount of critical feedback as measured by the number of "expressions of opposition to a proposal with, or without, evidence or arguments" (Valacich, Dennis, & Nunamaker, 1992, p. 59). Anonymous group members were more critical than identified group members; members of large groups were more critical than members of small groups; and small-identified groups were the least critical. The authors argued that the reason why larger groups provided more critical comments was due to the effects of anonymity—members of larger groups might have felt virtually anonymous. This conclusion was further supported by an empirical investigation by Connally et al. (1990), which showed that anonymous groups generated more unique critical ideas and more overall comments than non-anonymous groups.

Anonymous peer review is also believed to be related to improved writing skills. Pelaez (2002) compared the effects of two instructional approaches—problem-based writing with anonymous peer review versus traditional didactic lectures followed by group work—on student performance on physiology exams. He found that students learned better with anonymous peer review for problem-based writing. Similarly, Guilford's (2001) students commented that their course grades and the quality of their papers improved as a result of using anonymous peer review during the process of writing a term paper. The students in a pilot study which involved doing double-blind online peer review reported that the peer review process gave them insight into their own work and thus helped improve their writing skills (Tuautmann, Carlsen, Yalvac, Cakir, & Kohl, 2003). Haaga (1993) used anonymous peer review in his graduate courses, and his experiences suggested that anonymous peer review was an effective strategy in producing improved student papers.

Despite the many benefits of anonymity discussed in the literature, it is not without problems. Research on social loafing suggests that anonymity increases social loafing both physically and cognitively. Kerr and Bruun (1981) reported that identified group members generally exerted greater physical effort than those working anonymously. Similar results were obtained by William, Harkins, and Latané (1981). Their experiments confirmed that less physical effort was made by anonymous individuals, which they called "hide-in-the-crowd" effect. Weldon and Mustari found that anonymity reduced cognitive effort in a parallel "cognitive loafing" paradigm (as cited in Jessup et al., 1990, p. 339). Zhao's (1998) two empirical studies at the tertiary level also showed anonymity to be a double-edged sword: while it allowed participants to be more critical, it led them to work less. He concluded that accountability seems to be contrary to anonymity.

Feasibility of Anonymous Peer Review Made by Electronic Communication

The importance of using technology in writing instruction has been recognized by researchers and educators. In the past a few years, e-peer review—peer review via electronic communication—has become more prevalent in writing courses (Strever & Newman, 1997). Electronic communication presents several advantages over traditional face-to-face communication. It expands the boundaries of classrooms, and enables students to communicate anywhere, anytime. Electronic communication also discourages the emergence of dominant participants typical in face-to-face communications, and thus promotes more and equal participation among group members (Hartman et al., 1991; Hiltz, Johnson, & Turoff, 1986; Kelm, 1996; Mabrito, 1991; Siegel et al., 1986). Finally, some researchers contend that electronic communication helps students retain more information and results in better student learning (Hiltz & Turoff, 1978; Lin et al., 1999; Lin, Liu, & Yuan, 2001). For our purposes, its unique and important contribution is to make anonymity a reality in educational settings. In fact, electronic communication not only makes it feasible to conduct anonymous collaborative

activities but also raises the status of anonymity as an issue of inquiry in educational research and practice (Zhao, 1998). In the present study, it would have been extremely difficult, if not impossible, to ensure anonymity without this technological tool.

In sum, peer review can provide students with valuable learning experiences, which benefit reviewers and reviewees. However, there are potential problems regarding the quality of peer feedback. Even though previous studies illustrated how using anonymity can overcome such problems, they mainly focused on the critical nature of feedback and the release of social pressure due to the deindividuation effects associated with anonymity. Few studies examined the effect of anonymity on student writing proficiency or employed an experimental design to investigate its impact. Even fewer have adopted a comprehensive approach in the peer review process by incorporating a more complex treatment and multiple measures.

Purpose of the Present Study

The present study extends earlier research by examining the impact of anonymity on college student writing performance and the extent of critical feedback in a quasi-experimental design. An intervention package was introduced, which included the three most commonly advocated peer review methods—using electronic communication, multiple reviewers, and anonymous peer review. This study investigated whether the combined use of these methods could overcome some of the drawbacks associated with peer review and produce better student writing performance and more critical feedback. The research questions addressed were: 1) Does the use of anonymity in e-peer review process result in better student writing performance?, and 2) Are students engaged in anonymous e-peer review more critical in giving feedback than those engaged in identifiable e-peer review? Guided by the deindividuation theory and the bulk of studies identified in the literature, it was hypothesized that students participating in anonymous e-peer review would be more critical in their feedback than would students participating in identifiable e-peer review. Since there was not enough research to support whether anonymous e-peer review would also produce better writing performance than would identifiable e-peer review, a non-directional hypothesis was posited for this outcome.

Method

Overview

The present study was meant to determine whether an experimentally manipulated variable—anonymous peer review assisted by electronic communication—would result in better student writing performance and greater extent of critical peer feedback. The intervention was peer review under anonymous condition, via electronic communication, and with multiple reviewers (referred to as anonymous e-peer review in the remainder of the article). The control group also used multiple reviewers and electronic communication, but peer review was conducted under identifiable condition (referred to as identifiable e-peer review).

Participants

Participants were 92 college undergraduate students enrolled in four college entry level English composition classes in the fall semesters of 2003 and 2004 at an urban public university. All students must pass the Writing Sample Placement Test (WSPT) administered by the Writing Center of the University to be assigned to these classes. The purpose of the WSPT is to evaluate the writing abilities of all incoming degree-seeking students in order to place them in an appropriate composition class.

The original study was conducted in the fall of 2003. Two intact classes (n = 48, with 24 in each class) were randomly assigned to groups, with the treatment group giving and receiving

e-peer feedback anonymously, and the control group giving and receiving e-peer feedback among group members. The anonymous group met from 11:00 a.m. to 12:15 p.m. every Tuesday and Thursday, and the identifiable group met from 1:30 p.m. to 2:45 p.m. on the same days. The two groups were similar before the treatment in terms of prior academic achievement as measured by SAT verbal scores (M = 540 for the anonymous group; M = 561 for the identifiable group), and high school GPA (M = 3.06 for the anonymous group; M = 3.03 for the identifiable group). To strengthen the validity and reliability of the findings from the original study, it was replicated with another two English composition classes taught by the same instructor in the fall of 2004 (n = 44, with 21 in the anonymous group, and 23 in the identifiable group). These two groups were also very similar with respect to SAT verbal scores (M = 566 for the anonymous group; M = 553 for the identifiable group), and high school GPA (M = 3.09 for the anonymous group; M = 3.07 for the identifiable group). To avoid possible confounding elements caused by the timing of different class sessions, the 11:00 - 12:15 class was purposefully assigned to identifiable condition and the 1:30 - 2:45 class to anonymous condition in the replicated study.

Instruments

Student writing performance was measured via pre/posttest scores on two timed essays (400-500 words in 75 minutes). Two professional scorers from the university's Writing Center, who were totally blind to the experimental conditions, independently graded the pretest and posttest papers using Ransdell-Levy's (1996) writing evaluation rubric—Six-Subgroup Quality Scale (SSQS). SSQS is a set of holistic quality rating rubrics based on ratings of 13 dimensions of writing in six subgroups—Words: Choice and Arrangement; Technical Quality: Mechanics; Content of Essay; Purpose/Audience/Tone; Organization and Development; and Style. It was adapted from a university-level English placement exam and has proven both reliable and discriminating with college-level samples. In one source Ransdell and Levy report that "typical interrater r's are in the .80s and .90s." (Ransdell & Levy,1996, p. 95). However, in one published study, the reliabilities ranged from .74 to .88 (Ransdell, Levy, & Kellog, 2002). The latter reliability coefficients were more similar to those obtained in the present study. Measured by the Pearson-product-moment correlation, the inter-rater reliability of the two scorers in the present study was .64 in the original experiment, and .66 in the replicated experiment. The average of the two ratings of each paper served as the final score of the paper. As noted, the participants in this study were placed in these two classes based on their performance in one of such placement exams. Therefore, SSQS was considered an appropriate instrument to evaluate student writing performance for this study.

The extent of critical peer feedback was operationalized in this study as: 1) the frequency of peer negative comments, and 2) peer ratings. The frequency of peer negative comments was measured by the average number of negative comments per draft from each group. A negative comment referred to an independent meaning unit of a remark, a suggestion or a question that carried negative denotation or connotation. Examples of such comments are: "Wordy;" "Too many abbreviations;" "I cannot follow you;" "I think you should elaborate a little more here by giving some examples;" and "The comparisons here are irrelevant; would you try something more defining."

Peer ratings were measured by the average of the overall quality scores given by the two peer reviewers for each reviewed draft. Possible scores ranged from 1 (lowest quality) to 15 (highest quality).

The measures for the extent of critical peer feedback were adapted according to Zhao's (1998) model in which he designed multiple indicators to evaluate the critical nature of peer feedback: the grades assigned to each journal by peer reviewers, the degree of criticism of the

peer comments rated by the experts, and the degree of negativity of peer feedback perceived by the recipients. The data source for these measures was the peer feedback recorded on the *Editor Worksheets* and the edited drafts. The *Editor Worksheet* had been developed by the instructor, and had been used for peer review in this course for several years. It was modified by the instructor and researcher based on the SSQS for the purpose of this study. This evaluation rubric required students to provide a rating for the overall quality of the reviewed draft as well as specific comments to help with revision.

Procedure

The writing pretest was administered to the students during the second class meeting. The posttest was administered at the end of the semester. In order to encourage sincere and responsible writing in the pretest and posttest, students were told that their pre/posttest essays would be scored by the instructor to contribute 5% to their course grade. They were further informed that their papers would also be scored by the professional scorers for the purpose of the research, which would be independent of their course grade. Two class meetings in the second week were set aside for training in order to increase the effectiveness of the peer review process.

In both original and replicated studies, there were nine assignments throughout the semester, eight of which required e-peer review. For each assignment, each student reviewed two peers' drafts and received two reviews from peers on his/her draft. Several variables were controlled across the two groups: the same instructor, syllabus, textbooks, assignments, tests, number of reviewers for each draft, and peer review format—electronic communication (e-peer review). In addition, the number of words per written assignment was specified for both groups. This ruled out the possibility that the number of critical comments were higher because the length of the essays was significantly longer in one group versus another.

Students in the identifiable group were randomly divided into groups of three, and peer review was carried out among group members, which stayed stable throughout the semester; while in the anonymous group each student was given a 4-digit identification number (Class ID), which was used in all the paper work instead of his/her name. The instructor randomly assigned two reviewers to a reviewee for each assignment. A significant design dilemma arose in the grouping strategy for the two groups. If the peer review groups stayed stable for all reviewing tasks in both groups, anonymity would be lost in the anonymous group due to the small class size (21-24). If random assignment of reviewers for each reviewing task was employed in both groups, the identifiable group would become virtually anonymous, which would greatly diminish treatment effects. Since anonymous versus identifiable peer review was the primary interest of this study, protecting anonymity in the treatment group while maintaining identifiability in the control group was more critical. Therefore, random reviewer assignment was adopted for the anonymous group whereas stable member review was chosen for the identifiable group. The two treatments might be more thoroughly differentiated by describing the steps that comprised the peer review process.

Step 1: Filing draft. In the anonymous group, student authors anonymously posted their drafts for each assignment on *Blackboard.* On the same day, the instructor emailed each student two Class ID numbers, indicating whose drafts he/she should review for that assignment. In the identifiable group, student authors posted their drafts on *Blackboard* for group member peer review (students in this group used their names in all *Blackboard* postings). *Blackboard* allows the instructor or administrator to designate whether postings will be anonymous or not.

Step 2: Reviewing drafts. In the anonymous group, student reviewers went to *Blackboard*, identified the two drafts by their assigned authors' Class ID numbers, and downloaded them for review. In the identifiable group, student reviewers went to *Blackboard*, identified the two drafts

posted by the other two group members, and downloaded them for review. All reviewers were required to do two types of editing when reviewing their peers' drafts: 1) embedding editing—specific comments and suggestions about how to improve the paper were embedded in the draft by the reviewer, and 2) filing *Editor Worksheet* — each reviewer filed one *Editor Worksheet* on each draft he/she reviewed with suggested ratings (ranging from 1–15) as well as general comments on the overall quality of the paper. Then they posted both the edited draft and the *Editor Worksheet* on *Blackboard* on the due date for revision.

Step 3: Revising drafts. Student authors revised their drafts based on the peer feedback and their own reflections, and produced the final paper. The final paper was posted on *Blackboard* on the due date for instructor grading.

Anonymity Manipulation Checks

In order to check whether anonymity was protected, a rating scale questionnaire was administered to all the students in the two anonymous groups at the end of the semesters. The results were satisfactory. The majority of students indicated that neither they (77% in the original study, 79% in the replicated study) nor their peers (72% in the original study, 76% in the replicated study) were able to relate specific comments to a specific reviewer. The remaining students indicated that it was unlikely that reviewers could be identified.

Results

Quantitative Findings

The first research question addressed whether the use of anonymity in e-peer review process would result in better student writing performance. Using the pretest scores as the covariate, a one-way analysis of covariance (ANCOVA) was performed to assess differences in adjusted posttest mean scores between the anonymous group and the identifiable group. The difference between groups from both semesters was significant, $F(1, 42) = 9.54$, $p < .01$ in the original study, and $F(1, 41) = 6.39$, $p < .05$ in the replicated study (see Table 1).

Table 1: *Analysis of Covariance of Posttest Essay Scores as a Function of Treatment Condition with Pretest Essay Scores as Covariate*

Source	Variable	df	MS	F	p	η^2
Original	Covariate	1	2.55	6.59	.014	.14
	Posttest	1	3.69	9.54	.004	.19
Replicated	Covariate	1	.67	3.16	.083	.07
	Posttest	1	1.34	6.39	.015	.14

In the original study, the adjusted mean score on the posttest was 3.06 (out of 4) for the anonymous group, and 2.48 for the identifiable group. In the replicated study, the adjusted mean score on the posttest was 3.42 for the anonymous group, and 3.07 for the identifiable group (see Table 2).

Table 2: *Adjusted and Unadjusted Means of Pretest and Posttest Scores by Group*

Source	Group	Unadjusted Means	SD	Adjusted Means	SE
Original	Anonymous	3.09	.65	3.06	.13
	Identifiable	2.45	.67	2.48	.13
Replicated	Anonymous	3.43	.48	3.42	.10
	Identifiable	3.07	.46	3.07	.09

The descriptive statistics for the pretest and posttest essay scores are shown in Table 3. The findings revealed that in both studies students in the anonymous group had significantly higher posttest writing scores than those in the identifiable group. The results indicated that the use of anonymity in e-peer review process did result in better student writing performance.

Table 3: *Descriptive Statistics for Pretest and Posttest Scores by Group*

Source	Group	Pretest			Posttest		
		N	M	SD	N	M	SD
Original	Anonymous	24	2.43	.66	23	3.09	.65
	Identifiable	24	2.27	.75	22	2.45	.67
Replicated	Anonymous	21	2.69	.37	21	3.43	.48
	Identifiable	23	2.63	.73	23	3.06	.46

The second research question asked whether students were more critical in giving feedback when doing peer review anonymously. It was predicted that there would be more frequent negative peer feedback provided by students doing anonymous e-peer review than by those doing e-peer review identifiably, and that students doing anonymous e-peer review would tend to give lower ratings on their peers' writings than those doing e-peer review identifiably. A multivariate analysis of variance (MANOVA) was performed on the two dependent variables: negative peer comments and peer ratings, with group as the independent variable. In both studies, the overall MANOVA revealed a significant difference between the two groups. In the original study, Wilks' $\Lambda = .93$, $F(2, 625) = 22.49$, $p < .01$, and the follow-up ANOVA results were significant on the measures of both negative peer comments ($F(1, 626) = 37.15$, $p < .01$) and peer ratings ($F(1, 626) = 7.75$, $p < .01$). In the replicated study, Wilks' $\Lambda = .86$, $F(2, 645) = 53.60$, $p < .01$, and the follow-up ANOVA yielded significant differences on both dependent measures, $F(1, 646) = 101.04$, $p < .01$ on negative peer comments, and $F(1, 646) = 17.03$, $p < .01$ on peer ratings (see Table 4).

Table 4: *MANOVA Results for Peer Negative Comments and Peer Ratings*

Source	Dependent Variables	MANOVA			
		df	F	p	η^2
Original	Wilks's Lambda = .93	2, 625	22.49	.00	.07
	Peer Negative Comments	1	37.15	.00	.06
	Peer Rating	1	7.75	.01	.02
Replicated	Wilks's Lambda = .86	2, 645	53.60	.00	.14
	Peer Negative Comments	1	101.04	.00	.14
	Peer Rating	1	17.03	.00	.03

The descriptive statistics from both studies on these two measures indicated that the anonymous group did provide more frequent negative comments and lower ratings (Table 5). In the original study, the average number of peer negative comments per draft was 5.73 for the anonymous group contrasted with 3.94 for the identifiable group, and the average peer rating on each reviewed draft was 11.36 for the anonymous group but 11.82 for the identifiable group. In the replicated study, the average number of peer negative comments per draft was 4.38 for the anonymous group contrasted with 2.29 for the identifiable group, and the average peer rating on each reviewed draft was 11.75 for the anonymous group but 12.45 for the identifiable group. These findings confirmed that students working in anonymous condition were more critical than those working in identifiable condition.

Table 5: *Descriptive Statistics for Peer Negative Comments and Peer Ratings by Group*

Source	Group	Group Size	Negative Comments			Peer Rating		
		N	N	M	SD	N	M	SD
Original	Anonymous	24	332	5.73	4.15	332	11.36	2.32
	Identifiable	24	296	3.94	3.07	296	11.82	1.81
Replicated	Anonymous	21	312	4.38	3.21	312	11.75	1.46
	Identifiable	23	336	2.29	1.98	336	12.45	1.36

Qualitative Findings

To help explain the quantitative findings and understand in depth how anonymity worked in peer review, five open-ended questions eliciting students' perceptions on peer feedback and peer review system were embedded in the end-of-semester course evaluation survey. Students'

responses to these questions were analyzed using an inductive qualitative approach. A summary of major findings are presented below.

First, students in both groups agreed that peer review offered different and multiple ideas for revising their papers, which helped improve their writing. Furthermore, their ability to give feedback was improved through the peer review process. Nevertheless, there were more positive comments about the quality of peers' feedback and their feedback to peers in the anonymous group (N=54) than in the identifiable group (N=36).

Second, students in the anonymous group felt that it was easier for them to give and get honest and critical feedback under anonymous condition. Typical comments were: "Anonymous peer review is a good way to get critical opinions from everybody;" "With anonymous peer review, peers felt more comfortable giving their true feelings in their feedback." However, students in the identifiable group viewed the feedback they received from their group members as "less critical" or "less objective." Comments included: "Some people including myself weren't quite critical of others' papers;" "When you know each other, you tend to be less critical;" "I don't think people are objective in reviewing peers' papers—personal feelings often get involved." It was interesting to note that students tended to associate critical feedback with good feedback. They argued: "I attempted to be as critical as possible so that the author could make possible changes and improve their paper;" "I believe I gave good feedback, as I was straightforward and never hold my ideas back."

Third, students in the identifiable group were reluctant in sharing critical feedback with peers and worried about peers' reaction to their feedback. They sensed that sometimes peer authors would "get upset and take offense to editors' feedback," and at other times their comments "were not welcome" or "taken as criticism." One student revealed: "It was a little harder to be honest if you didn't like something because some people tended to take constructive criticism too personally." Another student commented on the same issue from a receiver's point of view: "They had problems letting me know what was wrong. They were just too nice." In contrast, students in the anonymous group did not express such concerns. They said: "One of the advantages of anonymous peer review is: it is all anonymous so you do not feel pressure to edit the paper the way the author would want you to;" "The real advantage of peer review was not knowing who the editors were so that you didn't have to worry about authors' feelings on your feedback."

Finally, students in both groups liked the flexible, convenient, and efficient characteristics of e-peer review system. They appreciated that the whole peer review procedure was administered through electronic communication so that they could "do it all in their comfortable homes, on their own schedules, and don't have to worry about printing out papers and finding time to meet group members."

The qualitative findings supported the quantitative findings that indicated students who conducted anonymous peer review provided more critical feedback than those who conducted peer review identifiably. These findings are also aligned with the literature suggesting that people were more honest and felt less stress in giving feedback under anonymous conditions.

Discussion

Peer review has been used in writing instruction for several decades, and it has been found to be beneficial to student learning. With the ever-increasing popularity of electronic communication, anonymous e-peer review has begun to draw attention from researchers and educators. This study examined the effects of anonymous e-peer review on student writing performance and the extent of critical peer feedback. The findings obtained from the original and

replicated studies support the benefits of anonymous e-peer review compared to identifiable e-peer review.

Student Writing Performance

The treatment in this study—using anonymous e-peer review in writing instruction—was targeted at improving students' writing performance. The findings from both original and replicated studies indicated that the students in the anonymous e-peer review group performed significantly better on the writing tasks than those in the identifiable e-peer review group. The results were consistent with previous studies on peer review and peer assessment reporting that learning gains were observed from student participating in peer review process (Chaudron, 1983; MacLeod, 1999; Mangelsdorf, 1992; Paulus, 1999; Plutsky & Wilson, 2004; Ramsden, 1992; Richer, 1992). The findings also seemed plausible for the research on e-peer review. A series of studies on Web-based peer review by Lin, Liu and their associates (Lin et al., 1999; Lin et al., 2001; Liu et al., 2001) indicated that Web-based peer review promoted student achievement. Similar results were obtained from Tannacito's (2001) qualitative study, which showed that both the quantity and quality of students' comments and revisions improved with peer review process. The results of this study further confirmed Topping's (1998) conclusion based on his meta-analysis of the studies on peer review at the tertiary level. He found that peer assessment of writing had positive formative effects on student achievement. The current findings expanded the related literature on anonymous peer review and student academic achievement (Guilford, 2001; Haaga, 1993; Pelaez, 2002; Tuautmann et al., 2003) by using an experimental design to illustrate that anonymous peer review did lead to greater improvement in student written work.

Extent of Critical Peer Feedback

This study also provided empirical evidence to support the theory and findings that anonymous peer review invites more critical peer feedback. Significant differences were found between the anonymous group and the identifiable group on both dependent measures. The anonymous group did provide more frequent negative comments and did give lower peer ratings. The results from the qualitative analyses complemented the quantitative results, which showed that students working in anonymous condition were more critical than those working in identifiable condition.

The findings of the present study further contributed to the related literature by linking critical feedback with student writing performance. Anonymous peer review appears to overcome the "halo error" phenomenon and leads to more critical peer feedback. The deindividuation caused by not being identified seems to minimize the restraining effects of fear of being wrong or creating conflict and thus elicit more candid opinions (Jessup et al., 1990), which in turn seems to be more useful in improving student papers. The current findings lend support to Falchikov's (1996) argument that constructively critical feedback is more useful in helping students improve their work.

However, the results of this study are not aligned with the "social loafing" theory that predicts reduced effort and diminished performance in anonymous rather than identifiable groups (Kerr & Bruun 1981; William et al., 1981). In contrast, the students in anonymous condition scored higher on their writing tests and provided more critical feedback than did the students in identifiable condition. The present results also conflict with those of Zhao (1998). He found that although anonymity led to more criticism, it also resulted in less work. In this study, the requirements and accountability structure intended to control potential confounding variables might have diminished potential "social loafing" effects linked to lower productivity.

Implications for Research and Practice

The insights gathered from this study, although informative, are in no way exhaustive. Limitations of the present study call for further research. The design changes that would most logically improve the study would include expanding the sample size and utilizing randomization of individuals to conditions rather than assigning intact classes. To gain a more in-depth understanding of how and why anonymous e-peer review was more effective, a more comprehensive qualitative study including interviews or analysis of feedback should be employed. Future research may also include mid-term measures in addition to pretest and posttest measures to investigate the patterns or trends in feedback and writing performance. In this study, critical feedback was simply operationalized as the frequency of negative comments and peer ratings on written drafts. Additional research is needed for a more comprehensive understanding of different types of critical feedback and their impact on student learning. For example, would mildly or strongly worded criticism lead to more revisions and academic gains?

The present findings suggest several benefits to educational practice. First, even though peer review has some drawbacks, it has become commonplace in writing environments. With its increasingly important role in college writing instruction, exploring a more effective peer review system to inform classroom practice is of great practical importance. The findings of this study support the use of anonymous peer review for promoting more critical feedback and improving writing performance. This may be particularly true when the course is designed to minimize the potential social loafing effects that might diminish accountability and productivity. Second, in the past, anonymity was rarely considered an important issue in educational research because most educational activities were conducted in face-to-face situations where anonymity was difficult if not impossible. Only in recent years has anonymity been more commonly incorporated into classrooms by computer and Internet technology (Zhao, 1998). Therefore, introducing anonymity in writing instruction through peer reviews conducted electronically is a relatively new approach, which has immediate and long-term values. Third, the present study examined the impact of a treatment package that combined the three most commonly advocated solutions to peer review problems—using electronic communication, using multiple reviewers, and using anonymous peer review. The findings obtained in an ecologically valid context of a real world classroom support the effectiveness of these solutions. Finally, this study laid a foundation for the feasibility of online writing instruction. As the world becomes more connected through electronic communication and more dependent on Web resources, exploring an effective strategy for writing instruction in online collaborative environments becomes imperative.

References

Asberg, P., & Nulden, U. (n.d.). Making peer review in large undergraduate courses an option. *The Viktoria Institute and Department of Informatics, School of Economics and Commercial Laws*, Goteborg University, Sweden.

Bagley, C., & Hunter, B. (1992, July). Restructuring, constructivism, and technology: Forging a new relationship. *Educational Technology*, 22-27.

Baker, S., Gersten, R., & Graham, S. (2003). Teaching expressive writing to students with learning disabilities: Research-based applications and examples. *Journal of Learning Disabilities, 36*(2), 109-124.

Bornstein, R. F. (1993). Costs and benefits of reviewer anonymity: A survey of journal editors and manuscript reviewers. *Journal of Social Behavior and Personality, 8*(3), 355-370.

Bostock, S. (2000). *Student peer assessment*. Retrieved August 19, 2003, from Keele University, Learning Technology Web site:

http://www.keele.ac.uk/depts/aa/landt/lt/docs/bostock_peer_assessment.htm

Carson, J. G., & Nelson, G. L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing, 5*(1), 1-19.

Chaudron, C. (1983). *Evaluating writing: Effects of feedback on revision*. Paper presented at the 17th Annual TESOL Convention, Toronto, Ontario, Canada.

Connolly, T., Jessup, L. M., & Valacich, J. S. (1990). Effects of anonymity and evaluative tone on idea generation in computer-mediated groups. *Management Science, 36*(6), 689-703.

Dyrud, M. A. (2001). Group projects and peer review. *Business Communication Quarterly, 64*(4), 106-113.

Eisenberg, M. B., & Ely, D. P. (1993). Plugging into the Net. *ERIC review, 2*(3), 2-10.

Eschenbach, E. A. (2001). Improving technical writing via web-based peer review of final reports. *31st ASEE/IEEE Frontiers in Education Conference*, F3A, 1-5.

Falchikov, N. (1996). Improving learning through critical peer feedback and reflection. *HERDSA conference*. Retrieved September 9, 2003, from http://www.herdsa.org.au/confs/1996/falchikov.html

Ghorpade, J., & Lackritz, J. R. (2001). Peer evaluation in the classroom: A check for sex and race/ethnicity effects. *Journal of Education for Business, 76*(5), 274-282.

Guilford, W. H. (2001). Teaching peer review and the process of scientific writing. *Advances in Physiology Education, 25*(3), 167-175.

Haaga, D. A. (1993). Peer review of term papers in graduate psychology courses. *Teaching of Psychology, 20*(1), 28-32.

Hartman, K., Neuwirth, C. M., Kiesler, S., Sproull, L., Cochran, C., Palmquist, M., & Zubrow, D. (1991). Patterns of social interaction and learning to write: Some effects of network technologies. *Written Communication, 8*, 79-113.

Harvard's Expository Writing Program (2002). *Longitudinal Study Highlights Importance of Writing*. Retrieved June 20, 2003, from http://www.fas.harvard.edu/~expos/index.cgi?section=archive&version=longitudinal

Hiltz, S. R., & Turoff, M. (1978). *The network nation: Human communication via computer*. Reading, MA: Addison-Wesley.

Hiltz, S. R., Johnson, K., & Turoff, M. (1986). Experiments in group-decision making. *Human Communication Research, 13*(2), 225-252.

Jessup, L. M., Connolly, T., & Tansik, D. A. (1990). Toward a theory of automated group work: The deindividuating effects of anonymity. *Small Group Research, 21*(3), 333-347.

Kelm, O. (1996). The application of computer networking in foreign language education: Focusing on principles of second language acquisition. In M. Warsehauer (Ed.), *Telecollaboration in Foreign Language Learning*. Hawaii: Second Language Teaching and Curriculum Center, 19-28.

Kerr, N. L., & Bruun, S. E. (1981). Ringelmann revisited: Alternative explanations for the social loafing effect. *Personality and Social Psychology Bulletin, 7*(2), 224-231.

Kerr, P. M., Park, K. H., & Domazlicky, B. R. (1995). Peer grading of essays in a principles of microeconomics course. *Journal of Education for Business*, July/August, 356-361.

Leidner, D. E., & Jarvenpaa, S. L. (1995). The use of information technology to enhance management school education: A theoretical view. *MIS Quarterly, 19*(3), 265-291.

Li, L., & Steckelberg, A. (2004). *Using peer feedback to enhance student meaningful learning*. Chicago, IL: Association for Educational Communications and Technology. (ERIC Document Reproduction Service No. ED485111).

Lin, S. J., Liu, Z. F., Chiu, C. H., & Yuan, S. M. (1999). Peer review: An effective web-learning strategy with the learner as both adapter and reviewer. *IEEE Transactions on Education.*

Lin, S. J., Liu, Z. F., & Yuan, S. M. (2001). Web based peer assessment: Attitude and achievement. 2001 *IEEE*. Retrieve June 28, 2003, from http://web.archive.org/web/20020511230446/http://www.ece.msstate.edu/~hagler/May2001/05/Begin.htm

Liu, E. Z., Lin, S. S., Chiu, C. H., & Yuan, S. M. (2001). Web-based peer review: The learner as both adapter and reviewer. *IEEE Transactions on Education, 44*(3), 246-251.

Mabrito, M. (1991). Electronic mail as a vehicle for peer response. *Written Communication, 8*(4), 509-532.

MacLeod, L. (1999). Computer-aided peer review of writing. *Business Communication Quarterly, 62*(3), 87-95.

Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? *ELT Journal, 46*(3), 275-285.

National Center For Education Statistics (1998). *Writing: State Report for Virginia*. U.S. Department of Education, Office of Educational Research and Improvement, NCES 1999-463.

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching, 51*(1), 34-39.

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing, 8*(3), 265-289.

Pelaez, N. J. (2002). Problem-based writing with peer review improves academic performance in physiology. *Advances in Physiology Education, 26*(3), 174-184.

Plutsky, S., & Wilson, B. A. (Winter 2004). Comparison of the three methods for teaching and evaluating writing: A quasi-experimental study. *The Delta Pi Epsilon Journal, 46*(1), 50-61.

Quible, Z. K. (1997). The efficacy of several writing feedback system. *Business Communication Quarterly, 60(2),* 109-124.

Ramsden, P. (1992). *Learning to teach in higher education.* London, Routledge.

Ransdell, S., & Levy, C. M. (1996). Working memory constraints on writing quality and fluency. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing: Theories, Methods, Individual Differences, and Applications*. Mahwah, N. J., Erlbaum.

Ransdell, S., Levy, C. M., & Kellog, R. T. (2002). The structure of writing processes as revealed by secondary task demands. *Educational Studies in Language and Literature, 2*, 141-163.

Richardson, V. (Ed.). (2001). *Handbook of research on teaching*. (4th ed.). American Educational Research Association, Washington, D. C.

Richer, D. L. (1992). The effects of two feedback systems on first year college students writing proficiency. *Dissertation Abstracts International, 53*(08A), 2722.

Robinson, J. M. (1999). Computer assisted peer review. *Computer Assisted Assessment*, 95-102.

Schaffer, J. (1996). Peer response that works. *Journal of Teaching Writing, 15*(1), 81-89.

Siegel, J., Dubrovsky, V., Kiesler, S., & McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes, 37*, 157-187.

Shaw, V. N. (2002). Peer review as a motivating device in the training of writing skills for college students. *Journal of College Reading and Learning, 33*(1), 68-77.

Stone, E. F., Spool, M. D., & Rabinowitz, S. (1997). Effects of anonymity and retaliatory potential on student evaluations of faculty performance. *Research in Higher Education, 6*, 313-325.

Strever, J., & Newman, K. (1997). Using electronic peer audience and summary writing with ESL students. (English as a second language). *Journal of College Reading and Learning, 28*(1), 24-34.

Tannacito, T. (2001). Teaching professional writing online with electronic peer response. *Kairos, 6*(2), 1-7.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*, 294-297.

Tuautmann, N., Carlsen, W., Yalvac, B., Cakir, M., & Kohl, C. (2003, March). *Learning nature of science concepts through online peer review of student research reports.* Presented at the annual meeting of the National Association for Research in Science Teaching, Philadelphia, PA.

Valacich, J., Dennis, A. R., & Nunamaker, J. F. (1992). Group size and anonymity effects on computer-mediated idea generation. *Small Group Research, 23*(1), 49-73.

Vinson, M. N. (1996). The pros and cons of 360-degree feedback: Making it work. *Training & Development, 50(4*), 11-14.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, M.I.T. Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Cambridge: Harvard University Press.

Williams, K., Harkins, S., & Latare, B. (1981). Identifiability as a deterrent to social loafing: Two cheering experiments. *Journal of Personality and Social Psychology, 40*(2), 303-311.

Zhao, Y. (1998). The effects of anonymity on computer-mediated peer review. *International Journal of Educational Telecommunication, 4*(4), 311-345.

Zhu, W. (1995). Effects of training for peer response on students' comments and interaction. *Written Communication, 12*(4), 492-528.

Zimbardo, P. G. (1969). The human choice: Individuation, reason, and order versus deindividuation, impulse, and chaos. In W. J. Arnold & D. LeVine (Eds.), *Nebraska Symposium on Motivation* (pp. 237-307). Lincoln: University of Nebraska Press.