

THE IMPACT OF FACE-TO-FACE VERSUS ONLINE PEER REVIEW
ON STUDENTS' WRITING ACHIEVEMENT

DISSERTATION FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY IN READING EDUCATION

LISA ANNE HINE

OAKLAND UNIVERSITY

2017

THE IMPACT OF FACE-TO-FACE VERSUS ONLINE PEER REVIEW ON
STUDENTS' WRITING ACHIEVEMENT

by

LISA ANNE HINE

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY IN READING EDUCATION

2017

Oakland University
Rochester, Michigan

Doctoral Advisory Committee:

John E. McEneaney, Ph.D., Chair
S. Rebecca Leigh, Ph.D.
Stephanie Roach, Ph.D.

© Copyright by Lisa Anne Hine, 2017
All rights reserved

To Chad, my best friend and soulmate, I couldn't have done it without you.

To Garrett and Kendall, for growing up to be my most impressive accomplishments.

And to my parents, Ethel and Rollo, for always believing in me.

ACKNOWLEDGMENTS

I would like to thank my committee members who guided me on this path towards earning my PhD. First, I would like to thank my chair, Dr. John McEaney, for his wisdom and guidance during all stages of my dissertation. Not only was he conscientious and reliable with reading multiple drafts, he went above and beyond by assisting me in getting my IRB passed in two universities. In addition, I would like to recognize Dr. Stephanie Roach, who offered valuable feedback on my study and made herself readily available to me, despite being on sabbatical. Lastly, I'd like to thank Dr. Rebecca Leigh for her insightful comments on my content as well as helpful advice on sharpening my writing. I was so fortunate to have such a talented group of advisers who shared generously of their time and expertise.

I would also like to thank my fellow doctoral students from China: Xin Zhang, Yu Liu, Pei Li, Wen Wu, and Yue Zhuo. I learned so much from this group and admire their dedication as scholars. I'd also like to give special mention to Dr. Becky Riesterer for her encouragement and humor. And last but not least, I would like to acknowledge Dr. Laura Gabrion, who was with me every step of the way as we traveled this road together.

Finally, I would like to say thank you to my family, for their constant support. To my husband, Chad, thank you for encouraging me when I wanted to quit. Thank you to my children, for their understanding when I had to study rather than play. And to my mother, for picking up the slack so that I could achieve my dream. Thank you is not enough.

Lisa Anne Hine

ABSTRACT

THE IMPACT OF FACE-TO-FACE VERSUS ONLINE PEER REVIEW ON STUDENTS' WRITING ACHIEVEMENT

by

Lisa Anne Hine

Adviser: John E. McEneaney, Ph.D.

This study examines the impact of type of peer review first year college students participated in, either face-to-face (F2F) or online (OL) and its effect on students' writing achievement. Using a two-way repeated measures design, this research measured students' overall writing score as well as the five key areas of writing: focus and meaning; content and development; organization; language use, voice and style; and mechanics and conventions. Since this design allows for a pre- and post-test of the participants, it was chosen to measure the treatment effect, comparing the type of peer review between groups and its overall impact on students' final writing achievement. The results suggest that there was not a significant effect based on the format of the peer review (F2F vs. OL) in either students' overall score nor in the five key areas of writing. However, there was statistically significant growth in both overall writing achievement scores over time as well as in the five key areas of writing, indicating that regardless of peer review format, first-year writing students' writing achievement can improve over the course of one semester.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	iv
LIST OF TABLES	ix
LIST OF FIGURES	x
LIST OF ABBREVIATIONS	xi
CHAPTER ONE	
ONLINE VERSUS FACE-TO-FACE PEER REVIEW	1
Introduction	1
Background of the Problem	2
Reaction to the Problem	4
Instrumentation Employed in This Study	7
Potential Objections to Peer Review	8
Conclusion	9
CHAPTER TWO	
A REVIEW OF THE LITERATURE	11
Overview	11
Theoretical Framework for Peer Review	14
Writing Process Theory	14
Cognitive Theory	16
Sociocultural Theory	19
Benefits of Peer Review	22
Peer Review Implementation	23

TABLE OF CONTENTS—Continued

F2F Peer Review	25
Benefits of F2F Peer Review	25
Limitations of F2F Peer Review	25
OL Peer Review	26
Benefits of OL Peer Review	27
Drawbacks of OL Peer Review	31
Existing Scholarship on Peer Review	32
Scholarship Using ESL and EFL Student Populations	32
Scholarship Using American Student Populations	33
Writing Assessment	35
Automated Essay Scoring	37
Summary	39
CHAPTER THREE METHODOLOGY	41
Introduction	41
Research Questions	42
Research Design	42
Setting and Participants	45
Instrumentation	47
Intervention	52

TABLE OF CONTENTS—Continued

Independent Variables: Online Peer Review or Face-to-Face Peer Review	53
Dependent Variables: Student’s Writing Sample Scores	54
Data Analysis	54
Summary	55
CHAPTER FOUR RESULTS	57
RQ1	58
Descriptive Statistics	59
Statistical Analysis	59
RQ2ae	60
Descriptive Statistics	61
Statistical Analysis	61
Summary	66
CHAPTER FIVE DISCUSSION AND CONCLUSIONS	67
Improvement in Grammar	69
Improvement in Organization	70
Online Environment	71
Requires Students to Write	72
Lowers Cost	72
Encourages Reflective Writing	74

TABLE OF CONTENTS—Continued

Limitations	74
Write Experience	74
Essay Prompts	76
Sample Size	76
Implications for Teaching	77
Students Must Be Instructed on Giving Feedback	77
Novice Writers Can Serve as Helpful Peer Reviewers	78
Online Teaching Differs from Conventional Classroom Teaching	79
Recommendations for Future Research	80
Summary	82
APPENDICES	
A. IRB APPROVAL LETTER OAKLAND UNIVERSITY	84
B. IRB APPROVAL LETTER UNIVERSITY OF MICHIGAN-FLINT	86
C. PRE-TEST WRITING PROMPT FROM WRITE EXPERIENCE	88
D. POST-TEST WRITING PROMPT FROM WRITE EXPERIENCE	90
REFERENCES	92

LIST OF TABLES

Table 3.1	Diagram Illustrating the Two-way Repeated Measures Analysis of Variance (ANOVA)	44
Table 3.2	Treatment Schedule for OL and F2F Groups	45
Table 3.3	Comparison of Definitions Used to Identify Five Key Areas of Writing	50
Table 4.1	Change Over Time Between OL & F2F Students for the Total Score	61
Table 4.2	Pre and Posttest Scores for Five Key Areas of Writing in F2F and OL Groups	63
Table 4.3	Results of a 2X2 Repeated Measures ANOVA Testing Differences in 5 Areas of Writing	64

LIST OF FIGURES

Figure 4.1	Estimated Marginal Means	60
------------	--------------------------	----

LIST OF TERMS & ABBREVIATIONS

AES	Automated Essay Scoring, which is performed by a computer in order to evaluate student writing quality
Peer review	The in-class activity of students offering feedback on one another's writing for the sole purpose of improving writing
Online peer review	The activity of students posting their rough drafts online and peers leaving written feedback comments

CHAPTER ONE
ONLINE VERSUS FACE-TO-FACE PEER REVIEW

Introduction

Writing is a skill employers seek, educators prize, and students often dislike. Beginning with a brief overview of the current state of writing education in American schools, this chapter will discuss why writing is a challenging craft to both learn and teach, yet why it is a critical component in education as it has the potential to significantly affect a student's future success in college and career. In addition, the technique of peer review will be explored as a method to aid students in becoming better writers. Since computers and the Internet are becoming a standard part of today's classroom, the traditional, in-class method of peer review is now being moved to an online environment, which dramatically changes the way students interact with their peers. Traditionally, in-class peer review takes place in the classroom in real time, but online peer review can happen asynchronously, with days of time elapsing between a student posting his draft and his peer leaving him feedback. To measure the impact of each environment on students' writing ability, this study will use a computerized method of evaluation called Automated Essay Scoring (AES) to measure quality indicators in student writing. This paper will explore why examining these two types of peer review environments, face-to-face (F2F) and online (OL) matter, as this practice has the potential to influence students' writing achievement.

Background of the Problem

Unfortunately, the large majority of American students are not proficient writers. Based on a 2011 study by the National Center for Education Statistics (NCES), nicknamed the "Nation's Report Card," only 24% of twelfth graders scored at the proficient level, which is defined as performing at grade level. A staggering 74% performed below proficiency, earning either a basic or below basic rating (NCES, 2012). Clearly, students are not adequately prepared to write, which affects college performance and beyond. *Academically Adrift* includes the dismal statistic that only 50% of seniors felt that their writing had shown any improvement over four years of college (Arum & Roksa, 2011). This information is especially alarming considering employers report that 26.2% of recent college graduates are deficient in basic writing skills (Conference Board, 2006). Writing is one of the top skills employers seek, and those who don't write well end up costing their companies billions. For example, a published survey entitled "Writing: A ticket to work or a ticket out" included 120 major American corporations and shows these businesses spent roughly 3 billion dollars annually on remedial classes for current employees who could not write well (CollegeBoard, 2004). Beyond needing to know how to write to communicate in the workplace, writing is considered a "threshold skill" for employees to be hired or promoted, and those who cannot write well risk being passed over for advancement in their careers (CollegeBoard, 2004).

Writing is critical for career success, yet many American graduates enter the business world unprepared for the task. Perhaps the answer is that writing is a challenging task, even for those who are considered skilled writers. Kellogg (1993) equates the composing process to an arduous physical task: "writing anything but the

most routine and brief pieces is the mental equivalent of digging ditches" (p. 17). To further characterize the writing process, Donald Murray (1982) calls writing "a demanding, intellectual process" for all writers (p. 3).

Teaching students how to write can often be a frustrating task, for both teachers and students. Kiuvara, Graham, and Hawken (2009) report in their national survey that over one half of writing teachers feel ill-prepared for the job. In fact, many of them say they have never received direct instruction themselves on how to teach writing. And of those who have received formal training, much of that instruction focuses on writing as a product rather than a process because that is the method they themselves were taught. Murray (1982) argues that this view of writing as product is problematic and suggests that "instead of teaching finished writing, we should teach unfinished writing, and glory in its unfinishedness" (p. 4). He further suggests that by expecting their beginning students to write like experts, educators are hampering students' creativity rather than encouraging their writing endeavors.

Another problem facing compositionists is that students must draw on knowledge from related subjects to be successful writers. For example, a student cannot compose sophisticated sentence structures if he is not an avid reader who is learning through exposure to advanced texts. Likewise, his word choice will be limited if he does not possess an adequate vocabulary to draw upon when composing an essay. Writing is closely tied with reading; if a student is not a strong reader, he likely will not be a strong writer. These skills are intertwined, which requires authors to draw on "a variety of motor, cognitive, and affective skills" to be successful (Graham & Harris, 2016). In order to address the problem facing educators and make writing a priority in the elementary and

secondary curriculum, the states decided to take action by creating set standards that every public school must follow.

Reaction to the Problem

In 2010, the states formed the Common Core State Standards (CCSS), an initiative to standardize K-12 student performance among states. These uniform standards specify benchmarks for students to meet in order to pass a given grade level. Since this implementation, writing has received considerable emphasis in the curriculum, making it "a central part of the school reform movement" (Graham, Gillespie, & McKeown, 2013, p. 1). No longer strictly the purview of the English Language Arts teacher, this newfound emphasis on writing across subject areas is unique in that it requires all teachers, regardless of subject area being taught, to incorporate writing, reading, listening, and speaking skills into their classes, not just English classes. This widespread inclusion of writing into all subject areas highlights the importance writing plays in the creation of students who are well-prepared for college or careers when they graduate from high school. CCSS highlight the importance of collaboration to share their writing, which peer review groups facilitate.

One popular method of improving student writing has been by implementing peer review groups. While the term peer review has many aliases such as peer feedback and peer assessment, this study will use Breuch's (2004) definition which defines peer review as "responding to one another's writing for the purpose of improving writing" (p. 10). Just as the CCSS stress the importance of including writing across the curriculum, so too can peer review groups be used in a multitude of subjects, not strictly writing. Peer review groups have been successfully implemented in computer science courses (Figl, et

al., 2006), medicine (McGarr & Clifford, 2013), science (Bauer, Figl, Derntl, Beran & Kabicher, 2009), and education (Meyer, 2003).

Peer review has been shown to offer students a host of benefits, from elementary age through post-secondary level (Wang, et. al, 2014; Figl et al., 2006). Peer comments have been shown to be comparable to teacher comments in accuracy and helpfulness (Topping, 1998, 2009; Simpson & Clifton, 2016), to promote student authors to work together (Bruffee, 1984; Meyer, 2003), to improve a reviewer's own writing (Figl et al., 2006), and to scaffold writing development by offering formative feedback (Topping, 1998). Whether taking place F2F or OL, peer review in each of these two forms offers the participants specific advantages and disadvantages. The first method discussed will be face-to-face peer review (hereafter, F2FPR), followed by online peer review (hereafter, OLPR).

Typically, peer review has taken place in the classroom, where peers comment face-to-face on their peer's paper. The purpose of peer review groups is to give formative assessment, which "aims to improve learning while it is happening in order to maximize success rather than merely determine success or failure only after the event" (Topping, 1998, p. 249). Peer review aids the writer by fostering the exchange of ideas and offering peers the opportunity to critique one another's writing. Usually, the focus of in-class peer review is on dialogue, of students discussing each other's writing verbally. These groups serve to not only socialize students but offer them multiple perspectives rather than solely the teacher's.

F2FPR offers particular advantages that OLPR cannot replicate. F2FPR gives peers the ability to immediately clarify potential misunderstandings when reviewing their

work in class. Responses are immediate and interactions are many when students interact in a F2F setting. However, when moved OL, students have fewer interactions and the number of exchanges diminishes, which is less dialogic in nature (Figl et al., 2006). F2F does not come without disadvantages, though. One danger is that peers get off-task when conducting peer review (Yang, 2011). Related to this problem, teachers are unable to monitor what is happening in these groups, since many groups participate at once, with several conversations happening simultaneously in the classroom.

Although F2FPR has been used for many years, educators are incorporating technology into their curriculum by moving peer review groups to an online environment, where students can post drafts of a piece of writing and peers can leave typed comments in response. This shift has come about mainly to the development of the Web and the different applications that have allowed students to work collaboratively. While Web 1.0 allowed users to view information statically, Web 2.0 carries with it the implication that students become creators of web content, not just passive recipients of content delivered via the web (Handsfield, Dean, & Cielocha, 2009). Web 2.0 tools can be defined as the second stage in web content, which focuses on user-generated content. These tools allow students to make original contributions and publish their work for others to view, changing ownership from the individual to the collective. Educators prize these tools for their practicality and ease-of-use both in and out of the classroom (Hancock, 2010). While hundreds if not thousands of web-based applications exist to assist educators to make learning an active experience (Oliver, 2010), these tools emphasize collaboration over the individual.

OLPR offers distinct advantages to the peer review process, with one of the most notable being the advantage of time. In an online environment, student reviewers can read and reread their peer's writing and take their time composing thoughtful responses, which is an advantage to F2FPR, as it is usually a hurried activity (Meyer, 2003). Another benefit of OLPR is the degree of anonymity offered to students (Chong, Goff, & Dej, 2012); students feel better able to offer honest critique OL rather than in a F2F setting. Some OL forums are "blind," meaning that no identifiable information is provided to peer reviewers. However, even when peers are identifiable by names, the mere act of moving feedback OL provides a degree of anonymity that allows a less threatening environment than F2F because they are in a virtual world (Lu & Bol, 2007; Vanderhoven, Raes, Schellens, & Montrieux, 2012). One final benefit is that teachers are better able to monitor student feedback, as OLPR leaves tangible evidence of feedback students share with each other (DiGiovanni & Nagaswami, 2001).

Instrumentation Employed in This Study

One tool this study will use is online forums, which are made available through a Learning Management System (LMS) that many schools and universities use. An LMS is a software application that offers educators a way to share course information in an online space. They also allow students enrolled in a specific course to work collaboratively on projects. One such tool offered are online forums, where students can post their writing and leave written feedback. These forums are ideally suited for the practice of peer review, particularly for their ease-of-use.

Another tool this study will utilize is Automated Essay Scoring (AES) software, which uses mathematical models to grade student essays. AES was originally used to

grade large amounts of student writing typically used on standardized tests. Now, however, this software is regularly being offered as a supplement to online textbooks, so students can write more often and receive feedback quickly without the added burden of grading written assignments being added to the instructor's workload. The particular instrument this study used is Write Experience, and this software assigns students an overall grade and further assigns subscores based on five key content areas most important in quality writing.

Potential Objections to Peer Review

Even though peer review groups can be successful when implemented properly, both teachers and students sometimes reject this form of collaboration. Teachers who adhere to the teacher-centered model have trouble releasing classroom control to students, who miss out on opportunities to learn from each other rather than the teacher. Other educators may feel derelict in their duties if they allow classmates to do what they consider to be their job, which is to comment meticulously on every student paper. Golub (2005) believes there is hope for such "stubbornly held beliefs" and that teachers are "getting over the idea that they are the only ones who can provide helpful, pertinent, valuable feedback: (p. xi).

Likewise, students themselves may resist peer review groups because they believe peers incapable of giving accurate advice when critiquing a paper (Sengupta, 1998). They, too, may value solely the teacher's opinion to the exclusion of all others. This is especially true in ESL populations whose cultures favor more teacher-centered curriculum than a learner-centered one (Mangelsdorf, 1992).

Objections by both groups can be overcome through proper preparation prior to peer review, which includes discussion of expectations as well as practice in giving feedback with a teacher modeling proper procedure. Students must be given explicit instruction on how to carry out peer review (Bruffee, 1964; Simmons, 2003) as well as be trained on how to diplomatically critique a peer's draft (Spear, 1998).

Conclusion

Even though OLPR as well as F2FPR are regularly practiced in composition classrooms across the nation, very few studies using American students compare these differences, which can be "important, even fundamental...in practice" (Breuch, 2004, p. 8). Educators need to look more closely at how differences in F2F and OL environments alter the practice of peer review since they have the potential to change learning outcomes. The researcher's study encompasses three areas: it uses online forums to conduct peer review, it allows groups to participate exclusively in either F2F or OL peer review, and it implements a software program to measure student writing. Using prior research on peer review and AES software, the general question this study will explore is what impact either F2F or OL peer review has on students' writing achievement.

The following chapter will include a review of the literature regarding the benefits and drawbacks of peer review in general, and further examine F2FPR and OLPR separately, highlighting each method's strengths and weaknesses. Theoretical frameworks supporting the peer review process will be included, along with specific studies that examine F2FPR and OLPR used in different populations, particularly English as Second Language (ESL) and American students. Also of importance, the key elements to

evaluate student writing will be discussed along with the AES software used to measure such quality indicators in student writing.

CHAPTER TWO

A REVIEW OF THE LITERATURE

Overview

Since the late 1960's shift from product-centered to process-centered writing (Flower & Hayes, 1981; Emig, 1971; Elbow, 1981), compositionists in the United States have implemented student peer review as part of the writer's workshop. The in-class, face-to-face (hereafter F2F) collaboration has been touted for allowing students to view themselves as readers and encouraging each student to improve his own writing while helping his peer improve theirs, a reciprocal relationship that benefits both individuals. Since computers have gained popularity as a classroom tool, educators moving this once in-class, predominately oral exercise to an online (hereafter OL) environment have been criticized by some (Harris, 1992; Russell, 1999) and idolized by others (Bennett & Barp, 2008). OL learning is growing at a staggering rate: the number of OL learners has exploded from 411,000 in 2003 to now 7.1 million (Allen & Seaman, 2014). Educators who practice the traditional F2F format of peer review in writing are looking for ways to accommodate this large population of OL learners. In response, the Conference for College Composition and Communication (CCCC) developed Online Writing Instruction guidelines. Principle 4 states, "Appropriate onsite composition theories, pedagogies, and strategies should be migrated and adapted to the online instructional environment" (2013). Exploring how changing the way first year college students engage in discussion, from a F2F peer review (hereafter F2FPR) to an OL peer review (hereafter OLPR), is crucial to determine its effect on writing achievement.

One question that needs to be considered before adding new technology into the curriculum is, "How does this technology benefit students?" Cynthia Selfe (1989) answers this question in her book *Creating a Computer-Supported Writing Facility: A Blueprint for Action* wherein she argues that pedagogy must drive technology, that "in the rush to buy new equipment, to purchase new software...we have not often had time as professionals to take care in the planning of computer use...to support English composition programs" (p. 22). Selfe provides two suggestions for using technology in the classroom: tailor computer labs and classrooms to meet the needs of writers, not computers; and situate classroom work within current writing theory and pedagogy (Selfe, 1989). These suggestions, while simple, serve to remind educators that what is most important is the way educators use technology to further their pedagogical goals for their students' growth as writers. In other words, moving peer review OL is not justifiable unless using this technology actually improves students' writing, which is the pedagogical goal of a composition class.

Computers have had a significant influence on the way writing is taught today, and peer review is no exception. As Kathleen Blake Yancey (2004) in her Chair's address at the Conference on College Composition and Communication stated: "Literacy today is in the midst of a tectonic change. Even inside of school, never before have writing and composition generated such diversity in definition" (p. 298). But transferring the traditional form of peer review to cyberspace fundamentally changes the process by bringing with it benefits and challenges. Research suggests that online peer review offers students benefits such as richer feedback and more time for thoughtful, honest responses (Liu & Sadler, 1998; Chong, Goff, & Deg, 2012; Meyer, 2003) but little has been

researched about online peer review and its impact on writing outcomes. This research will investigate whether different forms of peer review, OL or F2F, result in different outcomes on the writing achievement of first year college students.

The term peer review has many aliases in the literature, such as “peer assessment” (Cheng, Liang, & Tsai, 2015; Falchikov & Goldfinch, 2000; Kaufman & Schunn, 2011), “peer response” (DiPardo & Freedman, 1988; Hacker, 1996; Spear, 1988), “peer feedback” (Guardado & Shi, 2007; Lan, Wang, Hsu, & Chan, 2011; Rollinson, 2005), and “peer revision” (Zhu, 1994) and “peer evaluation” (Sengupta, 1998). Since these terms can have slightly different meanings, it is important to define the terminology. As discussed in Chapter One and in accordance with Breuch, peer review refers to the activity of giving feedback to a peer as a process for improving his writing. Peer review is not tied to any type of summative grade; rather, it is used to give formative feedback, which Sadler (1998) defines as “assessment that is specifically intended to provide feedback on performance to improve and accelerate learning” (p. 77). Bruffee (1984) defines peer review as an activity where “students learn to describe the organizational structure of a peer’s paper, paraphrase it, and comment both on what seems well done and what the author might do to improve the work” (p. 637-638). While some researchers use the term to describe an activity of co-authorship to produce a single piece of writing (e.g., collaborative writing), or assessing a peer’s writing by assigning a letter grade (e.g., peer assessment), this research is focused on peer review as the act of sharing constructive criticism with a peer regarding his writing and can be considered formative feedback.

Not only can the terminology for peer review differ so, too, can the method itself (Brammer & Rees, 2007). Variations exist as to whether the teacher assigns peer review groups or whether the students are allowed to choose their own partners (Zhu, 1994). Also, the students may be required to follow a handout to answer specific questions regarding their partner's paper (DiGiovanni & Nagaswami, 2001) or may be left to respond to their peer's writing without any criteria to guide their conversation (Elbow, 1981). Lastly, the peer review process can vary with regard to the amount of time that is allotted to the activity (Rollinson, 2005), how often it is implemented in the classroom, and if the teacher assigns any grade to the process (Topping, 1998), which may influence the amount of effort students devote to the activity.

Theoretical Framework for Peer Review

Writing Process Theory

Proponents of Writing Process Theory (WPT) turned their focus from the current traditional approach that stresses the final written product and focused, instead, on the composing process itself. This process approach to writing can be found in Janet Emig's (1971) study examining the composing processes of twelfth graders; she identified steps she believed writers moved through when creating a piece of writing, including prewriting, drafting, revising, editing, and publishing. Later, Emig (1983) described this transition from viewing writing as "a silent and solitary activity" to a process "enhanced by working in, and with, a group of other writers... who gives vital response including advice" (p. 140-141). Peer review is a hallmark of Writing Process Theory (WPT), as it values the social over the individual; hence, collaboration is a necessary component.

Emig (1971) was not alone in her classification of writing as a recursive process. Donald Murray (1982) echoed her position by warning writing teachers against criticizing students' skills which he characterized as "irrelevant" and instead advised teachers to "be quiet, to listen, to respond" (p. 5). Murray (1982) advocated only three steps in the writing process: prewriting, which he estimated takes up roughly 85% of a writer's time; writing, which is the fastest step, taking as little as one percent; and rewriting, which accounts for the remaining 14%.

To further cement the notion that writing is a process, Maxine Hairston uses the term "paradigm shift" to mark the significance of this change in writing, which she saw as revolutionary. Hairston (1982) criticizes the traditional view of writing as a linear process and more accurately describes it as a "messy, recursive, convoluted, and uneven" (p. 85) one. She discusses twelve features of this "new paradigm," the first of which "focuses on the writing process" that calls for compositionists to intervene in students' writing when needed (Hairston, 1982, p. 86). The goal of peer review is to do just that: to foster a writer's sense of self and learning how language works, by engaging writers in a conversation about their own writing.

Writing Process Theory (WPT) aligns with peer review because one assumption of this activity is that an author will revise his work, that writing is not a "once and done" proposition but naturally goes through a series of steps. Since the author can use his peer's suggestions to make revisions, WPT is a good fit with the peer review process because it views composing as an activity that has different stages and audiences. This process approach differs from the stage process approach because the stages are not linear but instead "are activities that overlap and intertwine" (Hairston, 1982, p. 86).

While WPT focuses on the steps a writer goes through to create a written product, cognitivists focus on the thought processes writers use when composing.

Cognitive Theory

The cognitive revolution was born as a major area of research in the 1950s. Until this point, scientists emphasized what could be outwardly observed while a person's thoughts or cognitive processes did not meet these criteria. Miller (2003) calls this new emphasis on the mind as "free[ing] itself from behaviorism, thus restoring cognition to scientific respectability" (p. 141). A direct refutation to the reigning theory of behaviorism, cognitive research in composition has its roots in cognitive psychology, as its focus is how people process information. These cognitive researchers focus on learning to write as a "goal-directed process" and skilled writers have already established the major goals they want to achieve before they begin writing (MacArthur, Graham, & Fitzgerald, 2015, p. 26).

Cognitive processes are involved in taking in information, performing mental operations on it, and storing it. To explain this procedure, researchers have used the computer analogy to explain how cognitive theory operates. At the time cognitive science was developing as a major research field, the computer was considered a new and exciting technological tool for communication. Cognitive theory postulates that people go through stages when processing information similar to how a computer functions.

Cognitivists were interested in how a person's mind worked; to that end, Ericsson and Simon (1993) developed "protocol analysis," which proposed a methodology to reveal how a person went about solving a problem and the cognitive processes involved. What was unique about their research is that in order to investigate this process, they

asked the subject to speak aloud his thoughts, an introspective way of determining the steps a person used to think through a problem. Ericsson and Simon (1993) believe this strategy "reveal[s] in remarkable detail what information [participants] are attending to while performing their tasks, and by revealing this information...provide an orderly picture of the exact way in which the tasks are being performed" (p. 220).

As a continuation of this research and in order to examine how writers compose, Flower and Hayes (1981) applied this think-aloud protocol to how a writer solves a writing task, during which college-aged research participants spoke aloud as they composed. This technique offered further insight into the thought processes students used while trying to solve a writing problem, which extended the original protocol analysis research to focus on writing in particular.

One of the most respected and influential theories that grew from this cognitive perspective is Flower and Hayes' Social Cognitive Theory (SCT) of writing model, which investigated how the work in cognitive psychology could be applied to students' acquisition of writing skills. The SCT model disputed the long-standing belief that writing is linear and static; instead, they argued that an individual's writing process was recursive and iterative and the steps involved in composing could vary from person to person. They use as an example the term "revision," which they say "can occur at any time a writer chooses to evaluate or revise his text or plans" and can be interpreted as re-envisioning of his writing (Flower & Hayes, 1981, p. 376). This contrasts starkly with the stage process view of writing wherein writers go through separate stages in a linear fashion. Flower and Hayes criticize the stage process view because of its focus on the product rather than the internal process of the writer. In their view, the stage model does

not accurately reflect what writers actually do when they compose; the steps are not separate and distinct but rather "a constant process of ... re-seeing what goes on while they are composing" (Flower & Hayes, 1981, p. 367). The term "re-seeing" used by Flower and Hayes (1981) speaks to the idea of a writer's ability to look at his writing from different perspectives, such as through the audience's lens rather than solely through the author's own viewpoint.

In 1996, Hayes revised the original model by expanding it to include the latest in cognitive research. The original model included three components: task environment, cognitive processes, and long-term memory. Originally, the task environment included everything outside the writer, such as the audience, topic, and assignment itself; the newer version includes collaborators as significant players in the task environment. Hayes also updated the cognitive processes component, which focuses on a writer's problem-solving ability, to be more reciprocal with the task environment. Finally, he supplemented his long-term memory component to include working memory, which acts like short-term memory for keeping a writer's ideas and implementing cognitive processes. Hayes' revised model places more emphasis on recognizing the importance of social interaction, which leads directly to sociocultural theory. Although the measure of a student's writing ability is a cognitive one, the act itself of peer review is a social one. While researchers who view literacy through a cognitive lens see the act as the development of a skill, socioculturalists view literacy through social and cultural contexts; peer review can be considered a social act since its main premise is built upon social interaction to improve writing.

Sociocultural Theory

Sociocultural Theory (hereafter ST) most closely aligns with the social aspect of peer review, “from viewing language learning as an isolated individual phenomenon to viewing it as one inherently embedded in and shaped by situated social interactions” (Hawkins, 2004, p. 3). ST stresses the importance of the social context for learning to occur. Sociocultural theorists base their ideas on the work of Lev Vygotsky, a Russian developmental psychologist from the 1930’s. Vygotsky’s ideas on how children learn suit the nature of the classroom environment, particularly peer review, as it relates to learning as a consequence of social interaction. Vygotsky’s work discusses “the more knowledgeable other” that serves to aid a learner’s cognitive development. Oftentimes, teachers are not the most knowledgeable person in the room; in fact, some students have specialized interests that cause them to be more expert than the teacher. Collaborative peer review implements this sociocultural pedagogy in traditional oral and computerized communication by adhering to Vygotsky’s model of learning as a social act. Although “on the surface, cognitive and social practice theories are not competing...they have not been treated as complementary theories either” (Purcell-Gates, Jacobson, & Degener, 2004, p. 63). One area where they differ is how they situate literacy.

Sociocultural theorist Brian Street argues that two different social models of literacy exist: one mainstream that is taught in schools and emphasizes a cognitive perspective and another that is marginalized and stresses a sociocultural orientation. Street (1984) criticizes the primary cognitive perspective as “autonomous,” meaning that “literacy consists of technical skills that are learned independently from social or cultural influences” (Davidson, 2010, p. 247). He further accuses this autonomous model as

favoring the cognitive model above other cultures. The alternative view of literacy, which Street calls the “ideological model,” attempts to overcome this inequity by imbedding literacy within social practice. Street (2003) argues that literacy is “not simply a technical and neutral skill; that it is always embedded in social constructed epistemological principles” (p. 77). By including peer review as a staple in writing courses, instructors situate learning within the social, “in contrast to traditional approaches to literacy which construe it as a set of cognitive skills possessed (or lacked) by individuals” (Hamilton, 2000, p. 16).

Similar to Street’s view, linguist James Paul Gee (2001) highlights inequity between how some children learn literacy at home, their “primary Discourse,” with that learned in school, a “secondary Discourse.” Gee (1990) believes that “Discourses are intimately related to the distribution of social power and hierarchical structure in society” (p. 20). Gee distinguishes between Discourses with a capital “D” as a sort of identity kit which includes the way a person from that Discourse would dress, speak, and act; Gee uses lowercase discourse to mean simply “language in use” (2001, p. 719). Gee’s argument means that knowing a particular Discourse allows those within that Discourse to speak, write, listen, believe, and interact like a native of that Discourse would. Ultimately, Gee makes the claim that children from different cultures do not know the Discourse of academia, which Gee argues is necessary to be successful in the school environment. Children unfamiliar with academic Discourse may not succeed in school - not because they don’t possess “impressive language abilities” but because they lack “specific verbal abilities tied to specific school-based practices and school-based genres

of oral and written language” (2001, p. 724). He argues that children need to “interact intensively with adult and more advanced peers” to be successful in school.

Not only is learning an academic Discourse important for young children, it is also an important goal for beginning college writers: to join the Discourse of their future area of expertise. In Kenneth Bruffee's (1984) seminal work, *The Conversation of Mankind*, he further elaborates on this idea of writing "within a community of knowledgeable peers" (1984, p. 642). Bruffee (1984) believes that teachers should “involve engaging students in conversation among themselves at as many points in both the writing and the reading process as possible” because ultimately “the way they talk...determines the way they will think and the way they will write” (p. 642).

Even though cognitive and sociocultural theories have traditionally been considered at odds, Purcell-Gates, Jacobsen, and Degener (2004) argue that the two theories are not "independent and incommensurable" because the cognitive happens within a sociocultural framework (p. 81). While sociocultural theorists view literacy as socially embedded whereas cognitivists view literacy as a set of skills to be learned, both components are necessary in achieving literacy success. Vygotsky's (1978) theory of internalization serves to bridge the gap between a cognitive view of literacy and a sociocultural one. He defines internalization as "the internal reconstruction of an external operation," which serves to use what a student learns from interacting with others and then adopting it as his own (p. 56). Vygotsky goes on to say that "every function in the child's development appears twice: first on the social level, and later on the individual level; first *between* people (*interpsychological*) and then *inside* the child (*intrapsychological*)" (1978, p. 57 emphasis in original). In application, a child learns

first through socialization, such as participating in a peer review group to discuss writing, and secondly by internalizing what he has learned and using that learning independently, such as composing and editing his own work. When put into practice in the writing classroom, peer review can offer students and teachers several benefits.

Benefits of Peer Review

A significant benefit of the peer review process is that it promotes collaboration among group members (Bruffee, 1984; Sullivan & Pratt, 1996; Meyer, 2003). Anne Ruggles Gere (1987) discusses peer review methods in what she calls "writing groups" as encouraging "negotiation rather than application of absolute standards" that causes writers to work together, using tentative language (p. 73). She believes that collaboration encourages students to question their own writing through the lens of their own experience, which triggers authentic response to what they read. Gere (1987) prizes writing groups for their ability to create and share knowledge by working together.

Another beneficial quality of peer review is the positive effect it can have on the reviewer's own writing. In general, peer review encourages the student offering feedback to "develop evaluation skills, increase their reflection ability, and develop awareness of their own work's quality" (Figl et al., 2006, para. 1), which accounts for its widespread use in writing classrooms. When a student is required to critically examine a peer's writing, he compares his writing to theirs and evaluates which is better and why. This process develops his reflexive ability, meaning his ability to evaluate his own writing through assessing another's. So even though the focus is on the benefits the author receives, it is also beneficial for the peer reviewer himself.

Peer review is a type of formative assessment, which by definition “aims to improve learning while it is happening in order to maximize success rather than merely determine success or failure only after the event” as opposed to summative assessment which assigns students a final letter grade with no chance for revision (Topping, 1998, p. 249). Peer review falls into the formative assessment category because it is usually considered low-risk and meant to scaffold the writer along the way to producing a final draft; the peer’s comments help to direct the writer and give feedback during multiple drafts. Topping (2009) further endorses participation of peers as "equal status learners" which can transfer skills learned in school to those needed in everyday life (p. 21).

Peer Review Implementation

Students do not intuitively know how to practice effective peer review; they must be taught how to critique each other as a learned behavior (Simmons, 2003; DiGiovanni & Nagaswami, 2001). Studies with college level students have shown that participants realize a greater benefit from peer review if they receive direct instruction on effective feedback strategies, regardless of the medium; the greater degree of responders’ experience, the more beneficial the peer review becomes (Bruffee, 1964; Hacker, 1996; Mangelsdorf, 1992). For example, in order to guide students effectively, “researchers offer models for the peer review process, and demonstrate the need for extensive training if the students are to be successful” (DiGiovanni & Nagaswami, 2001, p. 264). Bruffee (1964) offers a word of warning about implementing peer review groups, cautioning that "without guidance, peer influence can have a conservative effect potentially detrimental to intellectual growth" (p. 455). Teachers should not assume that students inherently know how to give helpful feedback.

In her book *Sharing Writing*, Karen Spear (1998) discusses both the theory behind peer groups as well as practical ways to implement them in the classroom. According to Spear, the success of peer review rests on three premises: that students' writing is an interpersonal process, that students' failure to thrive in peer review groups can be tied to their lack of interpersonal skills, and that writing and learning become conjoined when students learn how to successfully navigate peer review in conjunction with learning how to write. Spear (1998) sees the connection between students' interpersonal skills and the quality of their peer review as directly related and that these skills can be improved through group interaction by providing "novice writers [the opportunity] to behave like their more experienced counterparts"(p. 7). She advocates that teachers model for students how to conduct peer review by using constructive criticism in a group setting.

During peer review, students have a tendency to focus on sentence-level mistakes, what writing teachers commonly refer to as "local issues," which are minor and easily fixed. Sommers' (1980) study comparing student writers' revision strategies with those of experienced writers describes students' process as a mere "rewording activity" rather than major rewriting (p. 381). Instead, students must be encouraged to focus on global issues, those items such as content and organization that have a larger impact but are much more cognitively demanding to correct. Simmons (2003) argues that students need practice giving feedback and that it takes time for them to move beyond giving sentence-level advice and global praise. His study covered three years of high-school students' feedback during peer review and found that those students who were more experienced writers offered less editing pertaining to local issues and more focus on responding from a reader's point of view, which can be considered a global issue (Simmons, 2003).

F2F Peer Review

Traditionally, the practice of peer review has taken place in person with the author reading aloud his work to a peer and then discussing it. The method used to practice peer review, then, was almost exclusively oral rather than written.

Benefits of F2F Peer Review

F2F peer review offers several distinct benefits. DiGiovanni and Nagaswamni (2001) examined 32 ESL students' responses to OL versus F2F review. Those student authors who self-identified as preferring a F2F environment cited the main benefit as being able to further explain the meaning of their writing to their peer review partner, having the ability to immediately clarify any misunderstandings. In addition, their results showed that student reviewers in the F2F peer review had significant more questions for their author-partners than those in the online environment, 36 instances compared to only five online. DiGiovanni and Nagaswami (2001) attributed this higher number of negotiations in the F2F group because it allows for "a conversational mode of communication" that the online environment lacks (p. 267). Likewise, in another study by Figl et al., college student participants "found themselves constrained by the lack of discussion [online], which they experienced and valued in the face-to-face process" (2006, para. 1). Clearly, F2F offers some advantages to the peer review process that the online environment does not and cannot.

Limitations of F2F Peer Review

One significant criticism of F2F is that students often get off-topic during discussion and their attention therefore shifts from the paper they are reviewing to other topics of interest. Yang (2011) reports that "adopting peers' revision is low when peers

engage in non-revision-oriented feedback such as chatting and making complimentary remarks” (p. 688). Since instructors are unable to monitor all group conversations simultaneously, students may waste time by diverting their attention elsewhere.

Not only can students diverge from giving focused feedback, oftentimes when peer review is practiced in class the process can be hurried and students may not have adequate time to read a peer’s work and form thoughtful and/or substantive comments (Meyer, 2003).

OL Peer Review

Peer review practice traditionally took place in a face-to-face classroom environment. When teachers of writing offer this same practice, now referred to as “virtual peer review,” in an online environment, the process has fundamentally changed; what originally took place in a F2F classroom using oral language as the primary means of communication now takes place in an OL classroom using predominately written language (Breuch, 2004). Breuch (2004) stresses this significant change as OLPR relies solely on computers to create writing, exchange writing over the Internet, and leave feedback for the author entirely online. She further expands the definition of online peer review as “the activity of using computer technology to exchange and respond to one another’s writing for the purpose of improving writing” (2004, p. 10). This shift from traditional pen-and-paper to an online environment is also referred to as computer-mediated communication or CMC. Also related to an OL environment is the way students respond to each other, which can be categorized in two ways. The first method is called synchronous communication, where students are able to immediately respond to each other with no lag time, an example being Skype. This method is similar to F2F

communication because peers are able to respond orally and have a discussion in real-time. The second method is called asynchronous communication, where students post their feedback and then wait until their peer responds at a time that is convenient for them, an example being a Learning Management System (LMS) like Blackboard. Both synchronous and asynchronous communication pertain to an OL environment. What remains unclear, however, is how changing the environment from an oral one to a written one affects the way students participate in the practice of peer review and ultimately their written products.

Benefits of OL Peer Review

With the popularity of online courses, it is now estimated that roughly 7 million students enroll in a minimum of one online course over their academic careers (Allen & Seaman, 2014); therefore, it is necessary to examine how OLPR changes the peer review experience. Wang and Vasquez (2012) explain that these “Web 2.0 technologies exploit the participatory potential of the Web” by encouraging students to not only retrieve information, but share it. They are not only consumers of knowledge, but creators. Overall, OLPR offers significant advantages in the areas of time, degree of anonymity, amount of teacher control, and student engagement.

One benefit of OLPR pertains to time, specifically that moving discussions online frees up classroom time while simultaneously allowing students to spend more time drafting OL responses. Meyer’s (2003) study looks at the benefits moving a discussion online has in relation to time; she also includes a survey that polled students on which method they preferred, OL or F2F. Her study examines students’ threaded discussions for evidence of higher-order thinking and concludes that an online environment does

promote higher-order thinking. Her findings suggest that on the whole, students favored the OL environment over F2F because they “got more from the discussion” and could participate more fully online, which allowed students “time to reflect...who felt the constrained time of the classroom meant they could not contribute well or regularly to the discussions” (Meyer, 2003, p. 61). When moved OL, the peer review process can feel less pressured, allowing students multiple readings of a peer’s writing and the necessary time to compose a helpful response.

Another benefit of OLPR is the degree of anonymity it offers students. For example, Chong, Goff, and Dej's (2012) study compared two groups of second-year college students in the life sciences. One group participated in a F2FPR (approximately 500 students) and the other in OLPR (approximately 348 students). Based on participants' anonymous course evaluations, results showed that OL students found the peer review process more helpful in improving their essays than did the F2F students (Chong, et al., 2012). The authors attribute this preference to two reasons: 1) online students had more time to read and compose a response than did their F2F counterparts (20-40 minutes versus 5 days) and 2) online feedback was double-blind, so reviewers had no idea whose paper they were reviewing and authors had no idea who reviewed their papers, which “may permit more honest peer reviews and thus perhaps more constructive feedback” (Chong, et al., 2012, p. 71).

Similar to Chong et al.'s (2012) study using anonymous peer reviewers and identifiable peer reviewers, Lu and Bol (2007) compared both writing performance and amount of critical feedback between two groups of 92 college freshmen: anonymous and identifiable electronic peer review groups. Results showed that the anonymous group

"performed better on the writing performance task and provided more critical feedback to their peers" than the identifiable e-peer group (Lu & Bol, 2007, p. 100). While not all online forums are completely anonymous, the mere act of moving the critique online provides a degree of anonymity, even if the reviewer and reviewee are identified.

OL forums allow students who struggle with shyness or social skills to participate more than they would in a F2F classroom. Warschauer's (1996) study included 16 advanced ESL students and compared the amount of participation of each student in both F2F and OL environments. He found that more students participated OL than F2F, and hypothesized that "discomfort in speaking out" may account for a student participating less in F2FPR than OLPR (Warschauer, 1996, p. 21). In addition, his study also examined the complexity of their responses. By recording students' F2F discussions, he was able to analyze the complexity of student responses on two measures: vocabulary and sentence structure. Warshauer (1996) then compared the F2F discussions to the OL discussions and found that "electronic discussions involved significantly more complex language than the face-to-face discussions" (p. 17).

When peer review is approached online, the teacher's involvement in the virtual "classroom" can be affected. DiPardo and Freedman (1988) explore this concept of teacher-control and believe that how instructors choose to implement peer review "serves as a powerful indicator of an individual instructor's theory of what [writing process] means" (p. 127). They further suggest that teachers who follow in the footsteps of Elbow and Murray relinquish their control to their students, remaining in the background so students must devise their own methods for responding to one another's writing. Conversely, other teachers choose to maintain their involvement in the process and feel

uneasy about surrendering control to students. One way teachers attempt to maintain control is to require the use of scripted worksheets to be used during peer review. However, Dipardo and Freedman's study (1988) found that the use of worksheets hampered student collaboration and diminished the intended purpose of peer review as "truly peer-run collectives" by focusing instead on "teacher-mandated concerns" (p. 127).

In an online environment, the teacher's role can be more easily minimized since the teacher's online contribution to the discussion is one voice among many in the classroom and may appear indistinguishable from the students' contributions, supporting a learner-centered approach to learning. In Sullivan and Pratt's (1996) research comparing online and oral college-level classrooms, they found that the teacher "ultimately dominated the class discussion" during the F2F classroom in order to encourage students to participate; however, the teacher's role in the OL discussion was "drastically minimized" and she "was engaged solely as another voice in the discussion" (p. 498) as a contributor rather than the dominant voice. Overall, their study found the OL classroom to be more beneficial because students gave more suggestions for improvement and their participation was more engaged and focused (Sullivan & Pratt, 1996).

A final benefit of OLPR centers on student engagement. In a study by DiGiovanni and Nagaswami (2001), online students were found to be "on task and focused" and that OLPR has the advantage of allowing teachers to "monitor students' interaction much more closely than in face-to-face situations" (p. 268). Since online posts are easily documented, students and teachers alike can review notes and "need not depend on their

memory to revise drafts based on their peers' oral comments" (DiGiovanni & Nagaswami, 2001, p. 268).

Drawbacks of OL Peer Review

While OLPR undoubtedly has many benefits, it also has pitfalls. Students commenting asynchronously online do not experience the back-and-forth conversation as happens in a F2F setting. For example, Figl, Bauer, Mangler, and Motschnig's (2006) experimental study using F2F and OL groups in computer science courses analyzed differences between groups. Students reported that one drawback of peer reviewing online asynchronously was that they "found themselves constrained by the lack of discussion" and missed the dialogue they experienced in the F2F discussions (2006, para. 1). Students post rough drafts either via email or to an LMS, but oftentimes peers respond only once with all of their comments attached. As a result, asynchronous OLPR does not have reciprocity of a real-life conversation.

OLPR's advantages seem to outweigh its disadvantages. The benefits of offering more class time and allowing the teacher to minimize her authority are boons for OLPR, along with engaging students and providing a way to anonymously leave peer feedback. On the downside, the only notable drawback is that OLPR can lack the feel of a F2F discussion, particularly in asynchronous environments. This problem could be ameliorated by conducting OLPR using LMS forums or in real time, through services like Skype or FaceTime. The next section will specifically address studies discussing peer review in both OL and F2F environments using different populations of students.

Existing Scholarship on Peer Review

Scholarship using ESL and EFL Student Populations

Several studies compare the use of OLPR to F2FPR in the English as a second language (ESL) or English as a first language (EFL) classrooms (DiGiovanni & Nagaswami, 2001; Ho & Savignon, 2007; Liu & Sadler, 2003; Sullivan & Pratt, 1996; Warschauer, 1996). These studies focused on examining the quality of comments left by peer reviewers (Liu & Sadler, 2003; Sullivan & Pratt, 1996) while others concentrated on the revisions students made based on their peer's comments (Liu & Sadler, 2003).

Liu and Sadler's study (2003) compared both mode of commenting (Microsoft Word versus pen and paper) and mode of interaction (OL versus F2F) of 48 freshmen second language (L2) composition students. Their study also looked at the impact these differences had on students' final written product. Their statistical findings showed that those students who participated in giving feedback using Microsoft Word gave more comments, resulting in greater revision than the pen-and-paper group. However, they also included that participants felt the F2F mode of interaction was more effective "because of the nonverbal communication feature that is indispensable in intercultural communication in a peer review setting" (Liu & Sadler, 2003, p. 193). Because of students' preference for a combination of technology (Microsoft Word) and F2F interaction during review, the researchers recommended a hybrid form of peer review that includes both F2F and online activities.

Moloudi's (2011) study compared 124 Malaysian undergraduate students in distinct groups of F2F and OL. Students were given both a writing attitude scale and computer attitude scale in addition to a TOEFL pre- and post-test. In addition, students

participated in using semi-structured interview guides. Results showed no significant difference between the two groups; however, both groups showed significant gains in their writing achievement. A major weakness of this study was that no data was included for each group, only that the overall average mean performance of the 124 participants improved from 3.10 to 4.54. Another weakness was that it was unclear how many students were interviewed from the two groups, F2F and OL. Moloudi (2011) concluded by recommending a “mixed peer review” that offers “distinctive features of both face-to-face and online interaction” (p. 6).

While these studies illuminate how ESL students participate in OL and F2F peer review, it should also be noted that this population of students varies from native speakers. One difference is their level of proficiency in speaking and writing English, which can limit their ability to give helpful feedback. Another concern that relates closely to their proficiency is their comfort level; if ESL students struggle to communicate in English, they could be embarrassed during peer review or hindered to share their ideas. A final constraint could be the way Asian cultures distrust peer response in favor of a teacher-centered culture (Sengupta, 1998). For example, in a study by Mangelsdorf that used 40 advanced ESL learners, Asian students consistently rated peer review negatively, which can be a characteristic of “students not familiar with a collaborative, student-centered environment” (Mangelsdorf, 1992, p. 280).

Scholarship Using American Student Populations

Since this study will focus exclusively on first year native English students, it is necessary to include research that deals with this population. The following two studies

will be included as they consider peer review in both the OL and F2F environments and include American students as participants.

The first study by Mabrito (1991) focused on first-year composition students classified as high and low apprehensive writers and whether they responded differently as peer reviewers in OL and F2F environments. The OL method of communication was solely through electronic mail (e-mail), which is a somewhat outdated method to conduct peer review. Mabrito's (1991) multiple case study included only 8 participants; using a self-reporting survey, four students classified themselves as high-apprehensive writers, and four classified themselves as low-apprehensive writers. By coding transcribed interviews as well as email communications, Mabrito (1991) examined the "amount and type of discourse" between the two groups, concluding that high apprehensives participated more in the OL environment than F2F; commented less completely than their low apprehensive counterparts; and revised more heavily based on comments via e-mail than F2F (p. 512).

The second study by Stroble (1987) examined four areas related to students' OL or F2F participation in an English language arts class: student comments left in both types of environments, writers' reasons for revisions, quality of final compositions, and students' preference for using OL or F2F peer review. Strobel (1987) included twenty upper-level college students and divided them into four groups. Each group participated in both types of peer review (OL and F2F) and were therefore able to serve as its own comparison group. Similar to Mabrito's (1991) study, students communicated via electronic communication (e-mail) as their online method of peer review. Stroble (1987) reported that there was almost no difference between the amount of comments, writers'

reasons for revisions, and the quality of students' final compositions. She did include that students reported more comments initially using e-mail, but that the F2F group ended up commenting almost equally over the duration of their peer review session. Participants preferred F2FPR, citing the slowness of the computer and the amount involved to check if their partner had left feedback. Her conclusion was that OLPR "may function best as a complement rather than as a substitute for face-to-face sessions" (p. iv).

While the literature reveals that several studies investigate student attitudes regarding F2FPR and OLPR only one measures the effect on student writing achievement. The following section will discuss the literature regarding writing assessment and then discuss the use of automated essay scoring to evaluate student writing.

Writing Assessment

In order to determine the best way to assess student writing, educational researchers French, Carlton, and Diederich, sponsored by the Educational Testing Service, set out to determine what qualities were considered most important when assessing student writing; their 1961 study's findings serve as the bedrock for foundational writing skills to this day (Diederich, 1974). To figure out what qualities in writing were most essential, they assembled 53 judges, spanning six career fields, to assess three hundred student papers (Diederich, 1974). The researchers based their findings by tabulating the 11,018 comments these judges left on the student papers. The end result of this groundbreaking study "legitimized direct assessment of writing" because it evaluated "actual *writing* as the object of judgment" rather than previous methods of using multiple choice tests on grammar (Broad, 2003, p. 8, emphasis his). By

coding these judges' comments, five factors most common in outstanding student writing emerged.

The first factor was ideas, defined as "richness, soundness, clarity, development, and relevance to the topic and writer's purpose" (Diederich, 1961, p. 6). The second factor was mechanics, which pertained to usage, punctuation, spelling, and sentence structure. The third factor was organization and analysis, characterized by writing that demonstrates "an underlying plan" (Diederich, 1961, p. 56). Wording, the fourth factor, is "the choice and arrangement of words, including the deletion of unnecessary words" (Diederich, 1961, p. 8). The fifth and final factor cited style, which encompassed "the personal qualities" of the writer pertaining to the "individuality, originality, interest, and sincerity" of the writer (Diederich, 1961, p. 8). These five factors resulted in a rubric that "became the standard, traditional, five-point rubric...which nearly every large-scale assessment of writing since 1961 has been strictly guided" (Broad, 2003, p. 6). These key indicators of quality writing have allowed writing teachers to "reach impressively high correlations among scorers...with remarkable speed and relative ease" (Broad, 2003, p. 7).

Now more than 50 years later, these benchmarks of quality writing endure. For example, the popular traits method of today ties back to these same markers. One prime example is Vicki Spandel's 6-Trait Writing curriculum, which has its roots in Diederich's research and closely parallels the same key features necessary for students to master. Spandel (1984) furthered Diederich's research by expanding it to include 17 teachers and 5,000 students from Beaverton, Oregon. The quality markers identified by both Diederich and Spandel are the same markers used by educators today. The Common Core State

Standards Initiative for English Language Arts and Literacy, for example, listed the need for writing students to produce "clear and coherent writing" in the areas of "development, organization, and style," language that closely resembles those qualities identified in 1961 (2010, p. 46). In order to effectively measure quality indicators in a piece of writing, automated essay scoring (AES) was developed as an economical, accurate means of quickly scoring student writing.

Automated Essay Scoring

The field of AES, sometimes referred to as automated essay evaluation (AEE) or automated essay grading (AEG), developed as a practical method to score large amounts of student essays like the Scholastic Aptitude Test (SAT) or the Graduate Management Admission Test (GMAT) by using a computer rather than human raters. Rather than assess students using multiple choice or true and false questions, educators believe the better way to assess student learning is through the use of essays, which allow students "to demonstrate their range of skills and knowledge, including higher-order thinking skills, such as synthesis and analysis" (Zupanc & Bosnic, 2015, p. 383). AES allows teachers to assign frequent writing assignments, which otherwise might be impossible to grade without a team of expertly trained human graders.

AES draws on research from the fields of "computer science, cognitive psychology, educational measurement, linguistics, and writing research" in order to accurately score a piece of writing (Zupanc & Bosnic, 2015, p. 383). Originally, AES was developed for educational institutions that employed large scale assessment; however, with the advancements in technology, software based on AES is now being included in conjunction with course textbooks for use in traditional classrooms. AES

employs Latent Semantic Analysis (LAS) as a technique in natural language processing whose goal is to discover the underlying "latent" meaning of a text (Islam & Hoque, 2012). LSA works by "form[ing] a word by document matrix and then the matrix is decomposed using Singular Value Decomposition (SVD) technique" (Islam & Hoque, 2012, p. 616). Simply speaking, LSA uses a sophisticated mathematical technique to measure the similarities of words between two pieces of text. Using an "expert" piece of writing, LSA could compare the similarity between the "expert" and the "student" essay and thus render a judgement score as to the student essay's likeness to the expert.

AES addresses three problems associated with using human scoring: cost, time, and consistency. First, hiring human graders can be an expensive proposition while AES is more cost-effective. For example, the Florida Department of Education has chosen to use AES to grade the writing portion of their state standardized test, which typically uses two human graders. By replacing the graders with AES, Florida stands to save "30.5 million over the course of the six-year, \$220 million contract" (O'Connor, 2014). Second, employing this system saves time, since teachers spend approximately 30% of their time grading essays (Valenti, Neri, & Cucchiarelli, 2003). Whereas human scorers take several weeks to grade student writing in national assessments, AES can function in a matter of seconds. Third, AES has made huge gains in its scoring accuracy over the last ten years; experimental results demonstrate that when employing "LSA, the scoring accuracy can reach 88.8%" compared to human raters (Zhang, Hao, Xu, Ke, & Peng, 2014, p. 429). In addition, AES can alleviate human error: human raters may "combine their subjective judgments...leading to an inaccuracy of grading" whereas computers remain objective

and consistent (Zhang et al., 2014, p. 429). All in all, AES software has proven itself as a reliable, cost-efficient way to quickly and accurately grade student writing.

Summary

The growing popularity and outright dependency on technology in our society is undeniable, so teachers are interested in harnessing technology to benefit students, to move traditional in-class activities to an OL environment. One such activity that has traditionally taken place in the classroom has recently been offered OL: peer review. While peer review is increasingly taking place OL, little research has investigated the effect on students, particularly regarding their writing achievement. The studies that have been reviewed in this literature review are helpful in comparing F2FPR to OLPR, but among them only one study evaluated writing characteristics using an analytic, quantitative scale. In addition, no studies implemented an AES system even though they are growing in popularity because of their accuracy, affordability, and efficiency.

Of perhaps greatest concern, very few studies using American students compare differences in OL and F2F peer review even though both methods are regularly practiced in most composition classrooms. Educators need to look more closely at how differences in F2F and OL environments alter the practice of peer review since “virtual peer review has important, even fundamental, differences from [F2F] peer review in practice” (Breuch, 2004, p. 8). This current study is unique in three ways: it uses online forums to conduct peer review, it allows groups to participate exclusively in either F2F or OL peer review, and it implements a software program to measure student writing. By examining prior research on peer review and utilizing AES software as the instrument of measure,

this study will explore what impact either F2F or OL peer review has on students' writing achievement.

CHAPTER THREE

METHODOLOGY

Introduction

While traditional, in-class peer review has been used for many years in the writing classroom, growing popularity of computer technology has made it possible for students to experience peer review in an online setting. While both face-to-face (F2F) and online (OL) peer review have been implemented in writing classrooms across the world, the question of which type is more beneficial to improve students' writing remains relatively unanswered because few studies explore this comparison. Previous studies have concentrated on comparing the type of feedback students give (Hewett, 2000), student attitudes toward OL and F2F learning environments (Rendahl & Breuch, 2013), and the amount of peer feedback authors incorporate in their revised draft (Liu & Sadler, 2003). In addition, this study will also implement computer software to grade students' writing samples, a relatively new instrument to measure writing quality. This study will examine the type of peer review first-year American college students participate in and how that participation influences students' writing ability.

This study utilized two groups of students from two comparable universities. Participants were divided into two groups: those participating in online peer review (OLPR) and those participating in face-to-face peer review (F2FPR). For the purposes of this study, F2FPR is defined as a completely oral activity, with the peer review partner reading aloud the author's paper. OLPR is defined as a completely written activity, with feedback taking place asynchronously with identifiable peers. University A participated

in strictly OLPR, while University B participated exclusively in F2FPR. All other course content was kept similar. To further investigate the type of peer review on students' writing achievement, students' writing samples were scored using computer software and assigned an overall global score of 1-100. In addition, this global score was further evaluated in five content areas related to quality writing, with each content area being assigned an individual score of 1-6.

Research Questions

The data collected in this study will first answer a general question:

1. How is the overall writing score affected by type of peer review, OL versus F2F?

The research question can be further examined into the following five related questions:

2. How are the five key areas of writing affected by type of peer review, OL versus F2F?

2a. How are focus and meaning in writing affected by type of peer review?

2b. How are content and development in writing affected by type of peer review?

2c. How is organization in writing affected by type of peer review?

2d. How are language use, voice, and style in writing affected by type of peer review?

2e. How are mechanics and conventions in writing affected by type of peer review?

Research Design

A quasi-experimental nonequivalent control group design study (Fraenkel, Wallen, & Hyun, 2015) was conducted to assess the effectiveness of OLPR compared to

F2FPR. The participants were not randomized but a sample of convenience; as such, the results should not be generalized to all first year writing students. Since this design allows for a pre- and post-test of the participants, it was chosen to measure the treatment effect, comparing the type of peer review between groups and its overall impact on students' final writing achievement. In addition, this study further broke down the five areas of writing achievement to explore which of the five categories were influenced most according to peer review group.

In order to answer research questions 1 and 2, a two-way repeated measures analysis of variance (ANOVA) was used. Two main effects were tested. The first main effect was to determine if writing instruction led to higher scores from pre- to post-test (1-100). This within subjects effect will measure the change over time. This question will answer, "Is there a significant change from pre- to post-assessment?"

The second main effect was to determine differences between the two groups, F2F and OL (1-100). This question will answer, "Is there a difference between the F2F group versus the OL group?"

The interaction effect is the between-subjects effect, which is whether the change over time was different for the two groups, F2F or OL. This question will answer, "Is there a significant difference in the change from pre- to post-assessment for the two groups?"

To further explore the effect of peer review group on writing achievement, multiple repeated measures ANOVAs were run in each of the five content areas: focus and meaning, content and development; organization; language use, voice and style; mechanics and conventions. These five content areas are sub-analyses, secondary to the

overall analysis. Therefore, each one will be treated as a separate analysis, following the same structure and approach as the overall analysis.

This two-way repeated measures ANOVA design was chosen to measure both within and between group comparisons. See Table 3.1

After obtaining IRB approval for the study (see Appendix A), data were collected Winter semester of 2017, which ran from the beginning of January through mid-April. The treatment schedule is included in Table 3.1. In Week 2, students were administered the pre-test writing prompt; students were allowed approximately 45 minutes in class to complete a 500-word response and upload their essay to Write Experience, an automated essay scoring system used to grade student writing. This score was used as a baseline writing ability score before students participated in any treatment group. In weeks 3, 6, and 10, students participated in either OL or F2F peer review of three writing assignments. Again in Week 13, students were given a post-test, similar to the first, and were allocated approximately 45 minutes in class to complete and submit to Write Experience. While the scores that students received from Write Experience (WE) did not

Table 3.1

Diagram Illustrating the Two-way Repeated Measures Analysis of Variance (ANOVA)

	Pre-test	Post-test
F2F Group	1-100	1-100
OL Group	1-100	1-100

count towards their final grade in the course, students were given points for completing each writing prompt.

Both the pre-test and the post-test were graded by WE software in the overall writing score as well as the five subcategories: focus and meaning; content and development; organization; language use, voice, and style; mechanics and conventions. These WE subtest scores serve as the dependent variables in answering question 2a, b, c, d, e.

Setting and Participants

Sixty-six first year writing students from two public four-year Midwestern universities, referred to as University A and University B, contributed data in this study; University A comprised the OL group, while University B the F2F group. Table 3.2 illustrates the treatment schedule for the OL and F2F groups.

Table 3.2

Treatment Schedule for OL and F2F Groups

Treatment Schedule	
Pre-test	Week 2
Essay 1 Peer Review	Week 3
Essay 2 Peer Review	Week 6
Essay 3 Peer Review	Week 10
Post-test	Week 13

Originally, the OL group consisted of forty students, and the F2F group contained thirty-five students. However, over the course of the study, seven students dropped the class from the OL group, and two dropped from the F2F group, resulting in the final participant number of sixty-six.

The first year writing programs at both universities subscribe to a similar curriculum. Both programs emphasize that writing is a process and encourage student interaction and discussion as an integral component of producing a well-written essay. Both mission statements specifically mention creating environments that allow students to work collaboratively on writing projects and encourage revision as a necessary stage in this iterative process. Finally, both writing programs highlight that work produced in their writing classes should be structured, substantiated, and use language correctly. Structured writing is organized and easy-to-follow; substantiated writing includes support for claims; language use pertains to correct use of writing styles and the mechanics of writing.

The writing course used for this study is the second half of the first year writing program, in which occurs in the Winter semester and focuses mainly on argumentative and persuasive writing. At both universities, the class catalog and course description have nearly the same objectives and course outcomes, with academic writing and information literacy serving as fundamentals to the class. This class is also a required general education requirement for all incoming freshmen at both universities, and students must successfully pass the first semester as a prerequisite to enroll in this second semester class.

Both groups met with the same instructor twice per week, for a total class time of approximately three hours per week. The content and requirements for each group was similar, with the only difference being type of peer review group students participated in, F2F or OL. This study used convenience sampling as the four courses, two from University A and two from University B, were all taught by the researcher.

Instrumentation

To assess students' writing gains over the course of the semester, the researcher selected software created by Write Experience, which originated from the IntelliMetric (IM) Essay Scoring System and is Vantage Learning's automated essay scoring engine. IM software has been used to evaluate the Analytic Writing Assessment portion of the Graduate Management Admission Test (GMAT) and has been adopted by both universities that are used in this study. It is important to note that this software was commercially released in January 1998, and as such has proprietary trade secrets that prevent the full disclosure as to the particulars of how the software operates. Broadly speaking, IM employs artificial intelligence-based tools to offer students both feedback for improvement (subscores) as well as a final grade (overall score) for their written essays. IM software must be trained in order to operate effectively and a minimum of 300 graded essays must be entered into the software to develop a scoring model for each prompt. The system uses the scores from expert judges in order to effectively train the software, which is a "blend of artificial intelligence (AI), natural language processing and statistical technologies" (Elliot, 2003, p. 71). IM works by analyzing "more than 400 semantic-, syntactic-, and discourse-level features to form a composite sense of meaning" (Vantage Learning, 2016, para. 2).

The researcher chose the IM software for its consistency in scoring data (reliability) and its ability to accurately measure the quality indicators in a piece of writing (validity). IM relies on "multiple mathematical models, including linear analysis, Bayesian, and Latent Semantic Analysis (LSA) to predict the final score" which then combines input from these three models to assign a final score (Zupanc & Bosnic, 2015, p. 388). In order to assure the validity of IM, over 140 studies have been conducted since 1997 to test the software (Elliot, 2003) although most of these studies were funded by Vantage Learning. Based on six years of research, the findings suggest that the software is compatible with expert scorers, often exceeding their performance, that it is accurate across multiple grade levels and subjects, and that it shows reliable findings when grading multiple writing samples (Elliot, 2003).

Write Experience (WE) evaluates students' essays by assigning an overall score of 1-100 and further assigns a subscore in five key areas: focus and meaning; content and development; organization; language use, voice, and style; mechanics and conventions. These five subscores are rated using a score of one to six, with a score of one identifying the paper as "inadequate" and a score of six as "very effective." As shown in Table 3.3, the terminology used by the WE software closely resembles those characteristics of quality writing identified in French, Carlton, and Diederich's 1961 study; this seminal study was the first to identify the most important indicators in student writing and serves as the basis for modern-day rubrics used in assessment, such as Vicki Spandel's 6-Trait Writing model. (See Table 3.2)

The definitions for the five categories are taken from the WE Writer's Guide available at <https://www.myaccess.com/myaccess> and represent the characteristics most

important in quality writing as shown in Table 3.3. The first category is focus and meaning, which is defined as demonstrating a “clear understanding of the purpose of the task for the audience” and establishing “a clear controlling or central idea” that remains focused on the main subject of the essay. The second category pertains to content and development; content is the relevant detail a writer chooses to include in his essay, while development expands “ideas fully and artfully, using extensive specific, accurate, relevant details. If there is a text or texts, there is a wide variety of details from the text(s) to support ideas.” The third area is organization, described as having “a structure or pattern” as well as being unified and including transitions to connect ideas. The next area relates to language use, voice, and style in writing. Language use relates to fluency or “artful word choice and precise language.” Voice is defined as the writer’s “personality that comes through” the author’s writing. Style is determined by the decisions writers make, including word choice and sentence structure. Style is a combination of tone and voice, with tone being defined as the writer’s “manner of expression.” The fifth construct deals with mechanics and conventions, defined as the correct use of paragraphing, grammar, usage, punctuation, and spelling.

WE software offers a choice of 13 essay topics with prompts that instructors may choose to use. In order to create a new essay prompt, WE must input at least 300 expertly graded essays called *training sets* into the system that serve as training essays for the software; consequently, the software offers a limited number of essay prompts. The researcher chose two of these prompts (Appendix A & B) that were similar in nature for both the pre- and post- tests; students were assigned the same 500-word prompt and submitted their essays to WE, which assigned a percentage grade out of 100 points, with

Table 3.3

Comparison of Definitions Used to Identify Five Key Areas of Writing

Definitions from the Literature Review	Definitions from Write Experience
Ideas	Focus & Meaning
Mechanics	Mechanics
Organization	Organization
Wording	Content & Development
Style	Style

a breakdown score of one to six for each of the five categories named above, with a score of 1 being the lowest and 6 being the highest.

In a study published by the *Journal of Technology, Learning, and Assessment*, the IM software was evaluated using a two-pronged measure (Rudner, Garcia, & Welch, 2006). The first measure compared IM scores and two human raters' scores to a baseline score given by a group of expert raters to 750 responses to six prompts. Three of the prompts used analysis of an argument; the remaining three used analysis of an issue. Of the 750 responses, approximately 270 were identified as training essays and approximately 500 were labeled as validation sets. 13 essays were then added that were bogus, either being completely off topic or a repetition of the prompt. Results showed a perfect + adjacent agreement between the score of human raters to the original baseline data from .94 to .98, "resulting in slightly more than 4% of the scores needing adjudication" (Rudner, Garcia, & Welsh, 2006, p. 8). When comparing IM to the original

baseline data, results showed perfect + adjacent agreement from .96 to .98 over the six prompts, with only "3% of the scores needing adjudication" (Rudner, Garcia, & Welch, 2006, p. 9), a modest improvement over the two human raters. Of the bogus essays, IM was able to correctly identify them with 100% accuracy but did not fare as well with the essays identified as off-topic (Rudner, Garcia, & Welch, 2006).

In the second study, IM was evaluated against human raters using 500 responses to each of 101 prompts. Comparison of means using paired t-tests between average human scorer and IM score showed no significant differences at $\alpha = .05$ (Rudner, Garcia, & Welch, 2006). This study confirmed the findings from the first study, that the IM system "consistently calculates scores, closely matching those provided by human raters" (Rudner, Garcia, & Welch, 2006, p. 18).

A study using 500 international secondary students in the UK answered a single, narrative style question asking students to produce a narrative essay, which IM graded on a scale from 1-7. Scores between expert human scores and IM scores correlated by .90, with IM within only one point of human graders 100% of the time and exactly the same as human graders 62% of the time. These scores compared to the "expert-to-expert correlation of .89, adjacent agreement of 99%, and exact agreement of 64%" (Vantage Learning, 2001).

Overall, automated essay scoring software like WE appears to provide a reliable and valid measure to evaluate student writing. It has shown to outperform human raters, which is the current and traditional standard in writing assessment.

Intervention

The intervention developed for this quasi-experimental nonequivalent control group design study relied on two groups who participated in different types of peer review: F2F and OL. The OLPR group participated in three online peer review sessions while the F2F group participated in three F2F sessions. These sessions took place over the course of a 15-week semester. Both groups received instruction on how to give appropriate and helpful feedback to their peers before participating in peer review, with the five writing traits highlighted as the bedrock for quality writing. The researcher allotted one class period to explain the peer review process and how to effectively comment on a peer's paper. The F2F group received in-class modeling by the instructor, using a previous student essay as an example. The OL group also received in-class modeling; in addition, the researcher posted a student example essay and student feedback for that essay to the MLS for students to refer to as they gave their typed feedback. Both groups were provided sample papers that included examples of high quality student feedback. Global types of revision were encouraged, especially those dealing with content and organization, while local revisions like spelling and grammar were de-emphasized. Also modeled for both groups were types of questions reviewers might ask student authors. For example, a question related to content could be, "What important terms or definitions should be included in your essay?" A handout detailing possible questions related to each of the five writing traits was given to students (See Appendix C).

Peer review has been traditionally practiced in classrooms by having the peer review partner read aloud the student author's draft. As an alternate method, the student-

author reads his own piece aloud to his peer review partner. Using either method, the piece is read out loud during peer review. This practice has been referred to as a "read-aloud" strategy to improve students' writing. The intended benefit is "to let student writers experience readers other than themselves trying to make sense out of their drafts" (Rosenbaum, 1999, p. 105). According to Rosenbaum (1999), this method reinforces writing as a process by revealing "difficulties that the reader has with the draft, such as gaps in logic or inadequately explained analogy, [which] are likely to be precisely those problems that student writers are unable to see because they are too close to the draft" (p. 106). By hearing their pieces read orally, students are able to more easily identify their mistakes.

This study adopted the read-aloud strategy for F2F students, which required the peer reviewer to read aloud the author's paper. This method was used so that the student author could hear his mistakes and listen to see where the reviewer was having difficulty reading a passage, perhaps because of missing punctuation or words. For OL students, the peer review process required student authors to post their drafts to the OL forum, and then allowed the peer reviewer two days to leave written feedback. OL reviewers used the track changes feature in Microsoft Word to respond with their comments.

Independent Variables: Online Peer Review or Face-to-Face Peer Review

Thirty-three students from University A were assigned to participate in the OLPR group. OL students participated in three peer revision sessions, using the university's Learning Management System (LMS) and were randomly assigned two of their peers' papers to respond to by leaving detailed feedback. These partners were identifiable and remained the same throughout the semester. Each peer review group consisted of three

students, with each member of the group reviewing two group members' essays (author, reviewer 1, reviewer 2). In the case of an uneven number of participants, one or two groups were expanded to include four students rather than three. Since these were second semester Freshmen, students were familiar with the LMS and were well-versed in uploading files. The researcher demonstrated in class how to upload a file and comment in a forum.

Thirty-three students from University B were assigned to participate strictly in F2FPR. Similar to the online group, F2F students participated in three peer review sessions, with three students randomly assigned to each group, resulting in each student giving feedback to two of his classmates' papers. Again, students were randomly assigned partners which they maintained for each of the three peer review sessions.

Dependent Variable: Students' Writing Sample Scores

To assess the treatment effect, each cohort's post-test scores were examined to determine whether writing gains had been achieved. Next, scores were compared between groups to see if one group made greater writing gains. Gain scores based on sub-test scores in the five content areas described earlier were also examined.

Data Analysis

To compare pre- and post-test scores on writing achievement, a two-way repeated measures ANOVA was run using Statistical Package for the Social Sciences (SPSS) software (IBM, v24). To further explore the effect of peer review group on writing achievement, multiple repeated measures ANOVA were run in each of the five content areas: focus and meaning, content and development; organization; language use, voice and style; mechanics and conventions. The researcher chose this two-way repeated

measures ANOVA design to measure both within person comparison over time (pre- and post-test) and between person comparison between groups (OL and F2F).

In the overall analysis comparing pre- to post-tests, the alpha level of .05 will be used because it is the central analysis. However, the five sub-analyses will require alpha levels to be adjusted using the Bonferroni method. Therefore, alpha levels will be adjusted to .0083 for the five sub-analyses.

Summary

While peer review in the writing classroom is a ubiquitous practice, little research exists on how moving this event online affects first-year American students' writing ability. The differences between the two modes is substantial; F2FPR relies on orality to exchange feedback, but OLPR relies almost solely on writing to communicate. By fundamentally changing the way students participate in peer review, teachers could also influence how their participation affects their decisions as writers and consequently their final written products.

Using a quasi-experimental nonequivalent control group design, this quantitative study attempts to explore what impact each type of peer review, F2F or OL, has on first-year American students' writing achievement. This study implemented WE software to determine students' overall scores from pre- to post-test (1-100) as well as analyze sub-scores (1-6) related to the five writing characteristics. The two-way repeated measures ANOVA was chosen to assess the two main effects, which is the change from pre- to post-test as well as explore the differences between groups. The interaction effect will explore whether the change over time was different for the two groups, OL or F2F. The

sub-analyses will be run independently and further determine how each group scored in relation to the five content areas.

CHAPTER FOUR

RESULTS

Peer review is the process of students giving each other feedback on their written essays with the ultimate goal of improving student drafts while providing them with a genuine sense of audience (Lockhart & Ng, 1995). This activity has been a bedrock of writing classrooms and is used by practitioners as a technique to encourage social interaction while simultaneously encouraging students to offer helpful feedback to one another. Traditionally an entirely oral activity, peer review can now take place online, via computer, with comments being typed rather than spoken. Although peer review is "well established as an important theoretical component of the writing process" (Brammer & Rees, 2007, p. 71), less is known about the efficacy of moving this activity to an online environment.

Chapter Four presents the results of this quantitative study, which was to determine what effect, if any, type of peer review group students participated in, face-to-face (F2F) or online (OL), had on students' writing achievement. Specifically, this research looked at an overall writing achievement score of 1-100 and sub-scores aligning with the five key areas of writing (focus and meaning; content and development; organization; language use, voice, and style; mechanics), with scores of 1-6.

This study comprised a total of sixty-six participants. Thirty-five students began the study in the F2F group, but two students did not complete the semester and were dropped from the study, resulting in a total of thirty-three. Likewise, forty students began in the OL group, but seven students did not complete the semester, resulting in a total of

thirty-three for the OL group. Thus, each group comprised thirty-three students, with a final total of sixty-six participants.

Results for each research question are presented in order. The following two research questions are addressed in this chapter:

RQ1. How is the overall writing score affected by type of peer review, OL versus F2F?

RQ2. How are the five key areas of writing affected by type of peer review, OL versus F2F?

2a. Focus & Meaning

2b. Content & Development

2c. Organization

2d. Language Use, Voice, & Style

2e. Mechanics & Conventions

For RQ1, alpha levels are set at .05 for the main analysis. For RQ2, alpha levels are adjusted to .01 because five tests were run across the subcategories.

RQ1

Research question 1 asks, "Is there a mean difference in first-year American students' overall writing scores based on type of peer review (F2F or OL), controlling for pre-test scores?" Overall scores are based on a scale of 1-100. To answer this question, the researcher conducted a 2X2 repeated measures ANOVA to test the main effect for the format (F2F versus OL peer review), the main effect for time (writing achievement pretest scores to posttest scores), and the interaction effect, or in other words, the

combination of factors (F2F/OL and pretest writing achievement/posttest writing achievement) that yields the highest writing achievement score.

Descriptive Statistics

Results indicate the mean overall writing achievement raw score for the pre-test is slightly higher for the OL group ($M=68.00$, $SD=6.95$) than the mean overall writing achievement raw score associated with the F2F group ($M= 67.79$, $SD= 9.50$). Pre-test scores reveal that both groups began with similar writing ability. Post-test results show that the OL group performed slightly better ($M=87.17$, $SD= 11.62$) than the F2F group ($M= 82.06$, $SD= 7.09$). See Figure 4.1

Statistical Analysis

The main effect for the format of peer review group (F2F vs. OL) students participated in yielded an F ratio of $F(1,61) = 3.08$, $p>.05$, with a small effect size and weak power (partial $\eta^2_{\text{GROUP}} = .048$, observed power = .408), indicating no significant difference based on format. The other main effect based on time (pre- and post- scores) showed little difference in mean scores between groups: $M_{\text{F2F}} = 82.06$; $M_{\text{OL}} = 87.17$. However, results show that post-test writing achievement scores increased significantly from pre-test scores for both groups ($F= 143.78$; $df= 1,61$; $p< .001$). Finally, the interaction effect was not significant, ($F= 3.08$; $df= 1,61$; $p= .084$), indicating that none of the combinations of factors, format of peer review group (F2F or OL) or time (pre-post test) yielded a significantly higher writing achievement score. (See Table 4.1.)

RQ2a-e

Research question 2a-e addresses differences in students' sub-scores in the five key areas of quality writing. These areas are focus and meaning; content and development; organization; language use, voice and style; mechanics and conventions. Sub-scores were based on a scale of 1-6. Each research question 2a-e asks, "How is each key area in writing affected by type of peer review, OL versus F2F?" To answer this question, the researcher conducted a 2X2 repeated measures ANOVA to test the differences in change over time between the two groups regarding each of the five subcategories.

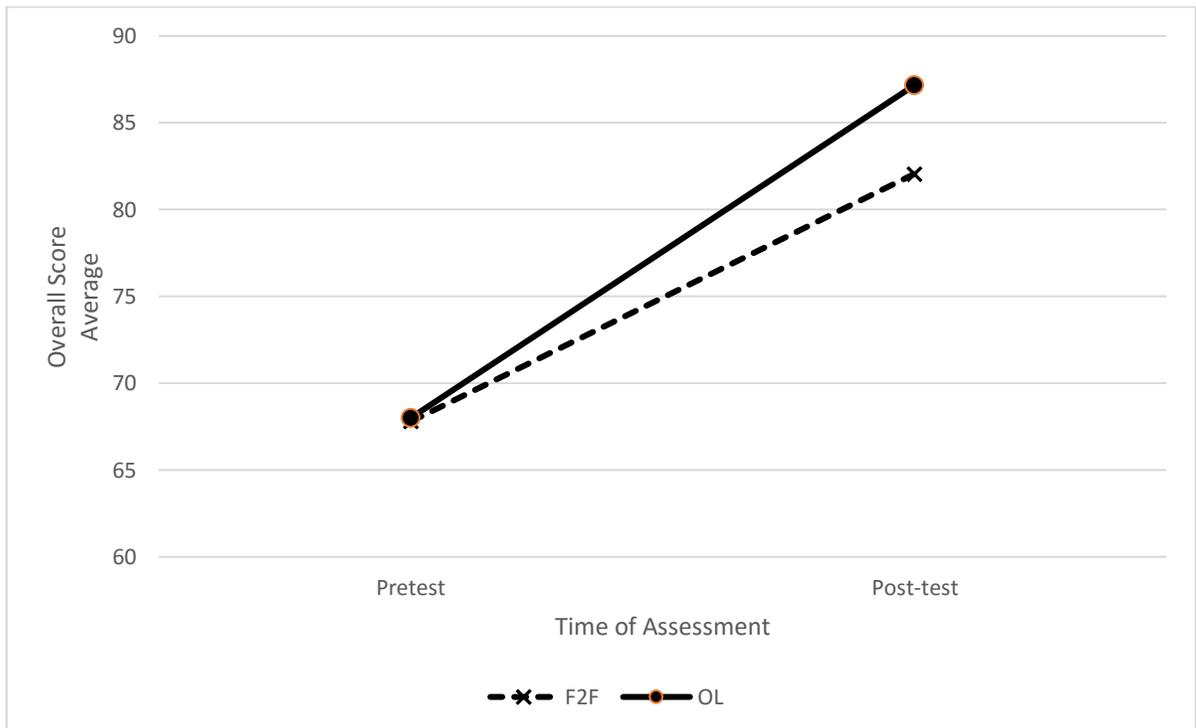


Figure 4.1. Estimated Marginal Means

Table 4.1

Results of a 2X2 Repeated Measures ANOVA Testing Differences in Change Over Time Between OL and F2F Students for the Total Score

Within Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	8785.803	1	8785.803	143.784	.000
Time * Format	188.184	1	188.184	3.080	.084
Error(time)	3727.356	61	61.104		

Between Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	730983.335	1	730983.335	7311.986	.000
Format	222.224	1	222.224	2.223	.141
Error	6098.205	61	99.971		

Descriptive Statistics

The researcher used 2X2 repeated measures ANOVA to explore differences in pre-post test scores in each of the five subcategories. The alpha level was adjusted from .05 to .01, because five tests were run across the subcategories. See Table 4.2.

Statistical Analysis

Results of the 2X2 repeated measures ANOVA regarding focus and meaning in writing suggests that there was not a statistically significant difference between groups ($F= .609$; $df= 1,61$; $p= .438$). The post scores in focus and meaning show little difference in mean scores between groups: $M_{F2F}= 5.15$; $M_{OL}= 5.47$. However, results show that post-test scores in focus and meaning increased significantly from pre-test scores for both

groups ($F= 109.641$; $df= 1,61$; $p< .001$). For content and development, results suggest that there was not a statistically significant effect between groups ($F= .539$; $df=1,61$; $p= .466$). The post scores in content and development show little difference in mean scores between groups: $M_{F2F}= 4.73$; $M_{OL}= 4.97$. However, results show that post-test scores in content and development increased significantly from pre-test scores for both groups ($F= 150.056$; $df= 1,61$; $p< .001$). Results of the 2X2 repeated measures ANOVA regarding organization in writing suggests that there was not a statistically significant difference between groups ($F=4.223$; $df=1,60$; $p=.044$). The post scores in organization show a difference in mean scores between groups: $M_{F2F}=4.50$; $M_{OL}=4.87$. In addition, results show that post-test scores in organization increased significantly from pre-test scores for both groups ($F=152.023$; $df=1,60$; $p< .001$).

For language use, voice, and style in writing, results of the 2X2 repeated measures ANOVA suggests that there was not a statistically significant difference between groups ($F= 1.162$; $df= 1,61$; $p= .285$). The post scores in language use, voice, and style show a difference in mean scores between groups: $M_{F2F}= 5.09$; $M_{OL}= 5.40$. In addition, results show that post-test scores in language use, voice, and style increased significantly from pre-test scores for both groups ($F= 103.098$; $df= 1,61$; $p< .001$).

Results of the 2X2 repeated measures ANOVA regarding mechanics and conventions in writing suggests that there was not a statistically significant difference between groups ($F= 4.450$; $df= 1,61$; $p= .039$). The post scores in mechanics and conventions show a difference in mean scores between groups: $M_{F2F}= 4.76$; $M_{OL}= 5.13$. In addition, results show that post-test scores in mechanics and conventions increased significantly from pre-test scores for both groups ($F= 81.939$; $df= 1,61$; $p< .001$).

Table 4.2

Pre- and Posttest Mean Scores for Five Key Areas of Writing in F2F and OL Groups

	Pretest		Posttest	
	M	SD	M	SD
Focus F2F group	4.06	.747	5.15	.566
Focus OL group	4.20	.551	5.47	.819
Content F2F group	3.52	.834	4.73	.517
Content OL group	3.60	.675	4.97	.809
Organization F2F group	3.50	.803	4.50	.508
Organization OL group	3.47	.571	4.87	.819
Language Use F2F group	4.12	.696	5.09	.522
Language Use OL group	4.20	.551	5.40	.814
Mechanics F2F group	3.97	.684	4.76	.614
Mechanics OL group	3.87	.571	5.13	.776

Table 4.3

Results of a 2X2 Repeated Measures ANOVA Testing Differences in 5 Areas of Writing

Research Question 2a: Focus and Meaning						
Within Subjects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Time	43.671	1	43.671	109.641	.000	
Time * Format	.243	1	.243	.609	.438	
Error(time)	24.297	61	.398			
Between Subjects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Intercept	2800.354	1	2800.354	5323.052	.000	
Format	1.623	1	1.623	3.086	.084	
Error	32.091	61	.526			
Research Question 2b: Content and Development						
Within Subjects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Time	52.251	1	52.251	150.056	.000	
Time * Format	.188	1	.188	.539	.466	
Error(time)	21.241	61	.348			
Between Subjects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Intercept	2220.001	1	2220.001	3246.421	.000	
Format	.826	1	.826	1.208	.276	
Error	41.714	61	.684			
Research Question 2c: Organization						
Within Subjects						
Source	Type III Sum of Squares	df	Mean Square	F	Sig.	
Time	44.594	1	44.594	152.023	.000	
Time * Format	1.239	1	1.239	4.223	.044	
Error(time)	17.600	60	.293			

Table 4.3 Continued

Between Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2065.376	1	2065.376	3150.574	.000
Format	.860	1	.860	1.312	.257
Error	39.333	60	.656		

Research Question 2d: Language Use, Voice, and Style

Within Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	36.988	1	36.988	103.098	.000
Time * Format	.417	1	.417	1.162	.285
Error(time)	21.885	61	.359		

Between Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2780.611	1	2780.611	5587.312	.000
Format	1.182	1	1.182	2.375	.128
Error	30.358	61	.498		

Research Question 2e: Mechanics and Conventions

Within Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Time	33.166	1	33.166	81.939	.000
Time * Format	1.801	1	.1.801	4.450	.039
Error(time)	24.691	61	.405		

Between Subjects					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	2469.156	1	2469.156	5145.353	.000
Format	.584	1	.584	1.218	.274
Error	29.273	61	.480		

Note. Significance computed using alpha = .01

Summary

The two-way repeated measures ANOVA reveal that there was not a significant effect based on the format of the peer review (F2F vs. OL). In addition, there was not a significant interaction effect (Time*Format), and therefore, not one combination of conditions had a greater impact on the students' writing achievement scores. However, there was a statistically significant growth in overall writing achievement scores over time (pretest to posttest writing achievement scores), indicating that regardless of peer review format, first-year writing students' writing achievement can improve over the course of one semester.

Results of the five additional two-way repeated measures ANOVA tests indicates that the format of the peer review (F2F vs. OL) does not significantly impact growth in the writing achievement subcategories (focus and meaning; content and development; organization; language use, voice and style; mechanics and conventions). However, similar to the findings for the overall writing achievement scores, there was statistically significant growth in each of the subcategory scores over time, indicating that all five areas of writing (focus and meaning; content and development; organization; language use, voice and style; mechanics and conventions) can improve over the course of the semester for first-year writing students. Since five content areas are sub-analyses within the larger analysis, alpha was adjusted from .05 to .01. While the data support the claim that writing scores increase as a result of participation in the writing course, the type of peer review, OL or F2F, did not show a difference in students' writing achievement scores. Chapter Five will further discuss the implications of these findings as they relate to practical writing instruction in the classroom.

CHAPTER FIVE

DISCUSSION AND CONCLUSIONS

This study centers on peer review, a common practice in writing classrooms everywhere; in fact, it is "a teaching strategy [that] is so entrenched in composition that it might almost serve as an emblematic practice of composition" (Warnock, 2009, p. 108). Peer review has been used to improve students' writing through collaboration and exchanging ideas by commenting on each other's written work. Although ubiquitous, peer review is experiencing a transition that few researchers have directly addressed. What once was a face-to-face (F2F) experience in the classroom is now being carried out in an online (OL) environment. This change dramatically affects the experience because it "reverses the primacy of oral over written communication so that written communication is king" (Breuch, 2004, p. 2). Thanks to technology, educators now have a host of digital tools that allow students to collaborate online.

Technology is influencing what is happening in today's classrooms, and composition classes are not exempt from being affected by this new form of literacy. Traditionally, peer review was practiced in the classroom, where a student could communicate directly and immediately with his or her peer review partner, within a conversational dialogue. However, this conventional practice has begun to shift, with peer review often taking place online rather than in a classroom environment. While students now have the ability to communicate synchronously using oral-electronic response such as Skype, this study uses strictly written feedback in online forums, with participants typing out their peer review comments rather than speaking to their partners

orally. This "e-feedback" transforms the spoken word to digital form, moving "oral response into the electronic arena" (Tuzi, 2004, p. 217). But what happens to the peer review conversation when it is moved online? Does a virtual environment do more to encourage a novice writer's development than a F2F experience?

As noted in Chapter Two, research suggests that online peer review offers students distinct benefits. In addition, it reinforces the social cognitive process of writing by allowing students to read and comment on classmates' work. The question of "'virtual peer review' has received very little explicit attention in writing studies" and, despite the modest results of this study, deserves a closer examination (Breuch, 2004, p. 2).

This research delves into the question of which type of peer review is more effective at improving students' writing by using a two-way repeated measures ANOVA with two groups of students participating in two types of peer review: F2F and OL. This study is somewhat unique in the way it measured students' writing by incorporating Write Experience, a digital software program, as the instrument of measure. Participants were given a pre- and post-test writing prompt that used computer intelligence to measure the quality of their writing by assigning an overall score of 1-100 as well as to evaluate their writing in five key areas: focus and meaning; content and development; organization; language use, voice, and style; and mechanics and conventions. The software graded these sub-scores on a scale of 1-6.

The results reported in Chapter Four show no statistical difference between groups, neither in overall writing achievement nor in the five key sub-scores based on the format of the peer review. In other words, the type of peer review (F2F or OL) did not significantly impact the overall writing achievement scores or the sub-scores.

Interestingly, two of the sub-scores did appear to be more sensitive to the format of peer review (OL or F2F), grammar ($p = .039$) and organization ($p = .044$). Both groups' writing scores, including overall and subcategories, improved significantly over the course of one 15-week semester. Although these findings would seem to indicate that the type of peer review does not directly influence writing achievement, investigating why grammar and organization seemed to show a greater improvement in the OL group seems helpful for future practice and research.

Improvement in Grammar

Although still considered a key area of writing, directly teaching grammar in a writing classroom is now considered ineffective. Hillocks' (1986) research suggests that "the study of traditional school grammar ...has no effect on raising the quality of student writing" and goes on to say that the traditional approach of writing teachers who place "a heavy emphasis on mechanics and usage (e.g., marking every error) results in significant losses in overall [writing] quality" (p. 74). Best practice suggests that students learn grammar through writing rather than completing grammar worksheets or skill and drill exercises. With this concept in mind, peer review sessions ideally are a conversation about the writer's ideas rather than surface-level problems. However, despite the researcher's explicit directions to avoid commenting on mechanics (grammar and spelling) in each other's online peer reviews, most student feedback did include references to mistakes in mechanics that the student-author made. For example, one student reviewer commented that his peer made "several mistakes with commas" and another suggested her peer "review the rules on punctuation and capitalization" to fix the errors in the peer's draft. Research supports the tendency for student writers to

concentrate on basic-level feedback when they are novice writers (Simmons, 2003; Straub, 1997). In other words, commenting on grammar mistakes is easier than commenting on content, so students often focus on lower-order issues, particularly for those students who are not native speakers (Sengupta, 1998; Sullivan & Pratt, 1996). Given the nature of F2F peer review, where a student reads his or her paper aloud and a peer comments on it, such attention to detail is not possible.

Perhaps practitioners can use OL forums to improve grammar, since the findings from this study show an overall increase in mean scores from the pretest to posttest as being slightly higher in the OL group than the F2F group. This suggests that OL peer review may indirectly improve students' grammar. Other research supports these findings. An online study by Villamil and de Guerrero (1998) found that "grammar was the most revised aspect" of student revisions.

Improvement in Organization

This subcategory of writing falls under what many experts consider to be one of the most important elements of a student essay. How an essay is organized has a strong impact on how readers understand the writer's message. In this study, while not statistically significant, the category of organization appears to be more sensitive to the format of peer review (OL or F2F), ($p = .044$), suggesting OL students may have made more improvement in their organization than did the F2F students. This finding is consistent with previous research that asserts the "medium shapes the talk" (Hewett, 2000, p. 265). In other words, Hewett (2000) found that the method students used to give feedback influenced the type of feedback they left. Her study reported that students who used e-feedback tended to focus on concrete revisions relating to organization and

content than did oral feedback, which centered on developing ideas. Perhaps the form students use to give feedback impacts the type of comments they choose to highlight in a peer's draft.

This question of how learning online may differ from a classroom environment is important because universities are consistently increasing the number of online classes they offer to accommodate students' busy lives. However, concern has been expressed by educators that the transition to teaching online is not a seamless one. Online teaching can be "very complex" for professors who are converting "a highly social process, that of educating students in a traditional classroom setting, to an online computerized setting with limited social interaction" (Dykman & Davis, 2008). Not only does the social dynamic change, but also the way the material is presented to the learner. Learning how to navigate this shift is necessary because OL courses are growing in popularity.

According to a national report investigating growth rates of online classes in higher education, 33.5% of all college students are enrolled in at least one online course (Allen & Seaman, 2014). The appeal of taking a course online is undeniable because these virtual classrooms allow for flexibility that today's non-traditional students expect. While online classes continue to grow, educators may want to consider how OL instruction differs from F2F and how their pedagogical practices may be affected by the medium.

Online Environment

The OL environment has important differences from its F2F counterpart (Dykman & Davis, 2008). Students can communicate in a variety of ways online, from posting in online forums to creating their own web pages or blog posts. One oft-cited benefit of the

OL world is that it is participatory in nature, meaning the medium encourages students to get involved. The following areas will highlight how the OL environment is distinct from the F2F, namely that it requires students to write, costs less to operate, and allows more time for reflection.

Requires Students to Write

Students who enroll in an OL writing course are required, by the medium, to communicate their ideas in writing. Unlike F2F students, OL students are "faced with a unique educational challenge/pressure to write to communicate almost everything in the course" (Warnock, 2009, p. xi). In addition, they are forced to read others' ideas rather than to listen to an in-class discussion. Due to the "sheer amount of writing exchanged among students and the teacher in an O[nline] W[riting] course," students are expected to use their reading and writing skills to participate in the class (Warnock, 2009, p. xi). This forced writing approach "privileges writing in ways that traditional classes cannot" (Harrington, Rickly & Day, 2000, p. 8). The 2012 Conference on College Composition and Communication (CCCC) reported findings about the "literacy load" (i.e., how much reading and writing were required) in two first-year writing courses, one conducted OL and the other F2F. The report showed that OL students were required to do 2.5 times more reading and writing than those in the identical F2F class. By its very nature, OL classes demand that students read and write to a greater degree than they would in an onsite class.

Lowers Cost

Online classes may be of interest to colleges and universities because of the challenges many institutions face with lowering costs while accommodating burgeoning

numbers of students. The appeal of online classes allows schools "to manage increasing enrollments without creating additional infrastructure" (Journell, 2012, p. 46). Today's higher education administrators allocate considerable budget dollars to building classrooms, dormitories, and general upkeep of campus buildings. Online classes may be a practical solution to these problems by offering an economical way to address these obstacles by simultaneously lowering costs for both administrators and students.

Another aspect related to cost that online instruction would alleviate is instructional costs. Francis (2012) explains, "The ability to reach students cheaply and efficiently is the great promise of online education" (p. 1b). Online classes require fewer instructors to manage than actual grounded classrooms. One example of the ability for one instructor to manage a large number of students OL can be exemplified by MOOCs (Massive Open Online Course). These courses are open to all students who are interested in learning about the course subject matter and can range in size from several hundred students to thousands. Currently, composition classes are offered in MOOC format, but these courses draw heavy criticism because of the way grades are evaluated.

For example, in Spring 2013, 67,530 students were enrolled in a Composition I class offered as a MOOC through Duke University. Rather than the instructor grading assignments, students were required to grade one another's work in a peer-assessment arrangement. While the instructor often provides students with rubrics and sample models of graded writing, he or she does not comment individually on each student's writing. In a MOOC, peers are responsible for assigning grades, with no instructor feedback. For a first-year writing course, a MOOC would probably be a less-than-ideal format for novice

writers because they omit the role of instructor as expert from the assessment process entirely.

Encourages Reflective Writing

Reflection can be traced back to John Dewey (1933) and his book *How We Think*. Broadly, reflection can be defined as making meaning and connections by looking back on one's own experience. Others define reflection as "purposeful thinking toward a goal" (Yang, 2010, p. 1202). One often-cited benefit of giving feedback OL is that students can reflect on both their peer's paper as well as their own response. Ko and Rossen (2008) term this experience as seeing an assignment "with fresh eyes" (p. 122). Reflection could be related to a benefit discussed in Chapter Two of this study, which is the advantage of time. Posting OL allows students more time: time to read a peer's draft several times, time to think about what they read, and time to compose a thoughtful response. This idea of time can be linked to the concept of reflection in writing (Breuch, 2004; Chong, Goff & Dej, 2012; Meyer, 2003). In Walther's (1992) study, Computer Mediated Communication (CMC) participants were "better able to plan" their online responses and used the "opportunity for objective self-awareness [and] reflection" (p. 229). Juan Flores (2006) echoes this statement regarding OL feedback by saying students "engage in reflective thinking over the writing of others" (p. 430). OL environments allow students time to engage in reflection, which is an element that can be lacking in a F2F scenario.

Limitations

Write Experience

This study used Write Experience (WE) software as the instrument of measure. WE grades students' writing, assigning both an overall score of 1-100 and five additional

scores of 1-6 in the five key areas of writing. Students also receive feedback from WE to improve their essays. The software uses artificial intelligence to mimic the way human scorers evaluate essay responses, by learning the characteristics of a known set of expertly graded essays and applying that rubric to ungraded essays. However, this software is not without controversy. Touted as being a much cheaper and faster alternative to human scorers, Automated Essay Scoring (AES) draws many who are critical of the software and decry that AES cannot "evaluate writing on the semantic level" as human graders can (Condon, 2013, p. 102). Although Condon (2013) is not a proponent of using the software, even he concedes that it does a "decent... job of mimicking the scores of human raters on short written samples" (p. 102).

This study employed the WE software to evaluate students' pre- and post-tests. WE assigns scores from 1-100, with each ten points on the scale representing one letter grade (Cengage Learning, 2014). Data for this study revealed many scores improved by 4 letter grades from pre- to post-test score. For example, one student's pre-test score as graded by WE was 50, but the post-test score jumped to 92, which is the equivalent of a 4-letter grade improvement. Another student's scores went from 59 to 90, a 31-point improvement. Many of the post-test scores were 20 points higher than pre-test scores and the overall average was 16 points higher on the post-test. Such unusual improvement leads the researcher to question the accuracy of the students' assigned scores, as many of them jumped as much as 40 points or more from pre- to post-test, which seems a questionable level of improvement in a brief 13-week period. One explanation of the score difference may be in the prompts themselves.

Essay Prompts

The researcher was limited in essay prompts available in the WE software, as only thirteen topics were available. However, the prompts chosen required students to take a position on a topic and argue their case. Although the researcher initially believed the prompts to be similar in nature, further investigation of the prompts may reveal a difference. For example, the pre-test prompt asked students to argue whether they felt they had a right to privacy or if they should be willing to give up a portion of their privacy in the name of security. The post-test prompt asked students whether they preferred traditional classrooms or online learning. While both prompts are argumentative in nature, they could differ in their appeal from a student's perspective or personal experience with the topic. Specifically, the post-test prompt could be more immediately relevant to students' experience taking online classes than would the pre-test prompt dealing with privacy issues. In this study, the mean post-test scores for both groups (84.49) were significantly higher than the mean pre-test scores (67.89). One way to test the hypothesis of whether the prompts themselves influenced the writing achievement scores would be to rearrange the order of prompts by reversing the order of pre- and post-test prompts.

Sample Size

Since the primary researcher was limited to the use of four of her own classes, the participants were not random which led to a quasi-experimental study. With only sixty-six participants, the data showed low observed power (.408) suggesting that the effect size is relatively low, which could explain why no statistically significant difference was

observed between the two groups, OL and F2F. A larger sample size would be needed to see a significant difference between groups.

Implications for Teaching

Although the overall improvement between the F2F and OL groups was not significant, both groups did make noteworthy gains in their writing ability over the course of the semester. Research supports the following ideas as related to classroom teaching: students need to be directed in how to give feedback, novice writers can serve as helpful peer reviewers, and teaching OL differs from conventional classroom teaching in important ways.

Students Must Be Instructed on Giving Feedback

Participants in this study were given instruction on how to give appropriate feedback to their partners during peer review. Both groups, F2F and OL, were shown transcripts of dialogue between an author and peer reviewer, a document with an example of peer feedback attached, and notes from the primary researcher encouraging global comments dealing with higher-order issues like content and organization. Peer reviewers were discouraged from focusing on local comments like grammar and mechanics during their peer review sessions. Research supports that peer reviewers need direction from the instructor before beginning peer review. In his article, "Responders are taught, not born," Jay Simmons (2003) highlights the importance of students being guided in how to give helpful feedback in his study using high school seniors. He found that students initially tended to leave comments that evaluated a peer's writing rather than responded to it as a reader. However, the more students practiced giving feedback, the more adept they

became and the more their responses "offered insight about what readers are thinking," concluding that "instruction changes students' response habits" (Simmons, 2003, p. 692).

Breuch (2004) recommends an exercise she employs in her own writing classroom, that of reading aloud mock transcripts of two example student peer review sessions. A pair of students reads aloud in class these sessions, pretending they are the actual peer reviewers as they read aloud from the transcripts. One example models a poor peer review session while the other demonstrates a fruitful session. Afterwards, the class discusses why one peer review session was more beneficial than the other. By modeling both sessions and discussing the strengths and weaknesses of each, Breuch (2004) believes students have a better idea of how to respond to each other's drafts in their own peer review groups.

Novice Writers Can Serve as Helpful Peer Reviewers

This study included first-year college writers who were enrolled in the second semester of a one-year program. As such, they could be considered as beginning writers who may be inexperienced writing in an academic genre as well as giving feedback to their peers. While giving students instruction on how to give feedback is one method to encourage profitable peer review, students do not need to be expert writers to serve as adequate reviewers. Writers need readers who serve as an interested audience, and peer review creates a genuine sense of audience by making "audience needs concrete and help[ing] writers who otherwise might not focus on those needs to do so" (DiPardo & Freedman, 1988, p. 124). In a study by Lundstrom and Baker (2009), the researchers found that those students who gave feedback without receiving feedback improved their writing ability more than those who received feedback without giving it. Interestingly,

those students who were classified as "low proficiency made more gains than those at higher proficiency levels" (Lundstrom & Baker, 2009, p. 30). This finding reinforces the idea that even those students identified as lower-level writers can improve their writing ability by giving peer feedback. Despite being first-year writers, participants in this study were able to give helpful feedback to their peers in both OL and F2F environments.

Online Teaching Differs from Conventional Classroom Teaching

This study employed participants using peer review in two types of environments, OL and F2F. The primary researcher realized differences when converting the peer review process traditionally held in class, to an OL environment. For instance, peer review sessions held in class produced no tangible results. While the primary researcher did listen in to various groups during peer review, no written work was collected that could be analyzed. However, peer reviews done OL did leave tangible evidence behind. By looking at student comments OL, the primary researcher could clearly see what feedback students gave one another. In addition, OL peer review required the instructor to be more organized by assigning peer review groups. In class, peer review partners are paired based on which students are present in class on the day of peer review. In an OL setting, partners must be paired prior to the peer review because rough drafts are posted and then reviewed asynchronously. These are significant differences to consider when moving online and are addressed in the research.

In 2013, the Conference for Composition and Communication (CCCC) published guiding principles of teaching online. Principle 3 says, "Appropriate composition teaching/learning strategies should be developed for the unique features of the online instructional environment." While the point of this research is not to argue one form of

teaching (F2F or OL) as superior to the other, the literature does caution practitioners against assuming they are identical. For the uninitiated instructor, focusing on the appeal of OL instruction rather than acknowledging the difficulties can be tempting. Dykman and Davis (2008) warn that teaching online is "foreign, uncertain, and much more difficult than teaching in the familiar conventional classroom" and that university faculty are often not supported properly in converting a F2F class to an OL class (p. 162). They discuss the need for OL teachers to be clear, at the outset, exactly what they expect of students and to follow the course syllabus throughout the semester closely since making sudden changes is much more difficult OL than F2F.

In his book *Teaching Writing Online: How and Why*, Scott Warnock (2009) recommends teachers who are converting from a F2F class "focus on what you do well in the classroom" so the "move to online teaching [will be] less difficult--and more enjoyable" (p. xiv). His focus is on taking the skills teachers have honed during their years of teaching in an onsite classroom and migrating them to the online environment. This migration requires educators to examine and reflect on their onsite pedagogical practices and how they might be translated to an online class (Warnock, 2009).

Recommendations for Future Research

Since this study used AES software as the instrument of measure, this study could be replicated. Results from this research suggested a sensitivity in two key areas of writing: organization and mechanics, with the OL group slightly outperforming the F2F group. To further explore this finding, future researchers could conduct another study using a larger sample size. By increasing the number of participants, findings may show a significance in these two areas. Another suggestion would be to follow-up this

quantitative study with a portion that is qualitative. For example, asking students to participate in a short interview or complete a brief survey would add further detail to the quantitative data and help to explain why students' performance in these two areas may have improved more so in the OL than in the F2F group.

Although this study was limited to pre- and post-test writing samples that focused solely on writing achievement, a future study could include only writing samples that have been peer-reviewed. This would allow researchers to track what type of revisions students made to their final drafts and how their participation in a specific type of peer review, OL or F2F, affect their revisions. For this study, the pre- and post- writing prompts were chosen from WE software and were not created by the instructor. As such, the prompts were not used as a graded writing assignment in the course and were therefore not peer reviewed as part of regular classroom practice. Because the prompts were used strictly to measure gains in students' writing ability, they were not peer-reviewed pieces. In other words, the essay responses submitted to WE were not peer reviewed by either group, OL or F2F. It is impossible, then, to tell if the improvement in students' writing was a result of what type of peer review they participated in. A future study could use peer-reviewed writing samples to track what type of revisions students made to their final drafts based on peer review type.

One final recommendation would be to replicate this study using the AES software but include both student writing samples that have been peer reviewed as well as those that have not. By doing so, researchers could compare the influence of peer review to students' baseline writing score that has not been peer reviewed. Including both writing that has been peer reviewed against writing that has not would highlight the influence of

peer review on a writer's performance in the key areas of writing as evaluated by the AES software.

Summary

Writing is a critical skill that allows students to successfully navigate school and the workforce. Educators have recognized its importance and it is the impetus for movements such as Writing Across the Curriculum (WAC) whose mission is to incorporate some aspect of writing into every discipline, from math to social studies. It is also listed in the Common Core State Standards which advocates that responsibility for "instruction in reading, writing, speaking, listening, and language be a shared responsibility within the school" (p. 4). Writing is no longer strictly the purview of English teachers, and peer review is not restricted to writing classrooms. One takeaway from this study is that the practice of peer review can be implemented to improve student writing in any subject, using a wide range of age groups. For example, peer review has been used successfully in children as young as third grade (Wang, et. al, 2014) all the way up to graduate students (Figl et. al., 2006). It is the current method of publishing research in academia, where scholars critique each other's work before it can be accepted and published in peer-reviewed journals. Educators value peer review because it "fosters the kind of conversation college teachers value most...in a community that approximates the one most students must eventually write for in everyday life" (Bruffee, 1984, p. 642).

Ultimately, data from this study showed little difference in outcomes, depending on the type of peer review practice students were assigned, OL or F2F. Studies have shown each type of learning environment has strengths and weaknesses. Additionally, the type of feedback students provide one another varies based on the medium. One

noteworthy difference is that OL feedback less resembles a conversation, with less back and forth dialogue between partners. However, the comments tend to be more focused and direct OL than when carried out F2F. While neither method can be declared superior, educators should be informed about the pedagogies they use in each environment.

Students themselves seem to have mixed feelings about which type of peer review they prefer. Some advocate for a hybrid of the two, sometimes referred to as a "blended course." Bauer et al. (2009) reported that students ultimately preferred a hybrid version: a combination of written peer reviews with the option to ask follow-up questions of their peer reviewer orally. Moloudi (2011) also recommended a "combined format" he dubbed "Mixed Peer Review" which allows students to "benefit from distinctive features of both face-to-face and online interaction" (p. 6). Going back three decades ago, Stroble (1987) concluded that OL peer review may function best as an addendum to F2F peer review. Undoubtedly, both OL and F2F offer advantages and disadvantages, so a mix of the two styles could offer students the best of both formats.

APPENDIX A

IRB APPROVAL LETTER FROM OAKLAND UNIVERSITY



Institutional Review Board for the Protection of Human Subjects

DATE: December 5, 2016

TO: Lisa Hine

FROM: Oakland University IRB

PROJECT TITLE: The impact of Face-to-face versus Online Peer Review on students' writing achievement

REFERENCE #: 951817-2

SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS DECISION DATE:

December 5, 2016

REVIEW CATEGORY: Exemption category # 1

Thank you for your submission of New Project materials for this project. The Oakland University IRB has determined this project is EXEMPT FROM IRB REVIEW according to federal regulations.

The exempt submission includes the following:

- IRB Application (IRBNet # 951817-2)
- Adult Consent Form Version 12/5/2016 **Please use the date stamped consent form in the recruitment and consent of all research participants.**
- Recruitment script (IRBNet # 951817-2)
- Peer review handout (IRBNet # 951817-1)
- Pre test Writing Prompt (IRBNet # 951817-1)
- Post test Writing Prompt (IRBNet # 951817-1)

The exemption is made with the understanding that NO CHANGES may be made in the procedures to be followed until such changes have been reviewed and approved by the IRB. Please use the "Protocol Amendment" form found in IRBNet to submit any proposed changes to the IRB. Do not collect data while the proposed changes are being reviewed. Data collected during this time cannot be used.

Please retain a copy of this correspondence for your record.

If you have any questions, please contact Kate Wydeven M.S. at (248) 370-4306 or kwydeven@oakland.edu. Please include your project title and reference number in all correspondence with this committee.

This letter has been electronically signed in accordance with all applicable regulations, and a copy is retained within Oakland University IRB's records.

APPENDIX B

IRB APPROVAL LETTER FROM UNIVERSITY OF MICHIGAN-FLINT

Subject: Notice of Exemption for [HUM00124390]

SUBMISSION INFORMATION:

Title: The impact of face-to-face versus online peer review on students' writing achievement
Full Study Title (if applicable): The impact of face-to-face versus online peer review on students' writing achievement

Study eResearch ID: [HUM00124390](#)

Date of this Notification from IRB: 12/16/2016

Date of IRB Exempt Determination: 12/16/2016

UM Federalwide Assurance: FWA00004969 (For the current FWA expiration date, please visit the [UM HRPP Webpage](#))

OHRP IRB Registration Number(s): IRB00000248

IRB EXEMPTION STATUS:

The IRB Flint has reviewed the study referenced above and determined that, as currently described, it is exempt from ongoing IRB review, per the following federal exemption category:

EXEMPTION #1 of the 45 CFR 46.101.(b):

Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.

Note that the study is considered exempt as long as any changes to the use of human subjects (including their data) remain within the scope of the exemption category above. Any proposed changes that may exceed the scope of this category, or the approval conditions of any other non-IRB reviewing committees, must be submitted as an amendment through eResearch.

Although an exemption determination eliminates the need for ongoing IRB review and approval, you still have an obligation to understand and abide by generally accepted principles of responsible and ethical conduct of research. Examples of these principles can be found in the Belmont Report as well as in guidance from professional societies and scientific organizations.

SUBMITTING AMENDMENTS VIA eRESEARCH:

You can access the online forms for amendments in the eResearch workspace for this exempt study, referenced above.

ACCESSING EXEMPT STUDIES IN eRESEARCH:

Click the "Exempt and Not Regulated" tab in your eResearch home workspace to access this exempt study.

Marianne McGrath

Chair, IRB Flint

APPENDIX C

PRE-TEST WRITING PROMPT FROM WRITE EXPERIENCE

Pre-test writing prompt used from Write Experience

Topic: Your Right to Privacy

“Notices from credit card companies, ads on television, warning on the evening news, release forms at the doctor’s office, and locker searches at school all address the issue of invasion of privacy. Some are opposed to what they believe is an increased invasion of privacy for individuals. Others believe that national and individual safety and security necessitate some invasion of privacy for individuals.”

“In a well-developed essay, articulate your position on your right to privacy. Include facts and examples to support your position.”

APPENDIX D

POST-TEST WRITING PROMPT FROM WRITE EXPERIENCE

Post-test writing prompt used from Write Experience

Topic: Internet or Traditional Classrooms

“With advancement in technology, some students have chosen to complete their schoolwork through the internet, rather than in traditional classrooms. Those who support a traditional classroom approach to learning argue that it provides more of an opportunity for students to interact with the teacher and other students. Those who feel that learning over the internet is better argue that it allows students to learn anytime and anywhere they choose.”

“Do you feel education is better provided in traditional classrooms or when offered over the internet? Write an essay to be read by a classroom teacher in which you persuade the reader that either traditional classroom education or internet-based learning is better.”

REFERENCES

- Allen, E.A., & Seaman, J. (2014). *Grade change: Tracking online education in the U.S.* Babson College: Babson Survey Research Group.
- Arum, R., & Roksa, J. (2011). *Academically adrift: Limited learning on college campuses.* Chicago: University of Chicago Press.
- Baker, E. A. (2010). *The new literacies: Multiple perspectives on research and practice.* New York: Guilford Press.
- Bauer, C., Figl, K., Derntl, M., Beran, P., & Kabicher, S. (2009). The student view on online peer reviews. Paper presented at the 26-30. doi:10.1145/1562877.1562892
- Bennett, S., & Barp, D. (2008). Peer observation - a case for doing it online. *Teaching in Higher Education, 13*(5), 559-570. doi:10.1080/13562510802334871
- Bourelle, A., Bourelle, T., Knutson, A. V., & Spong, S. (2016). Sites of multimodal literacy: Comparing student learning in online and face-to-face environments. *Computers and Composition, 39*, 55-70. doi:10.1016/j.compcom.2015.11.003
- Brammer, C., & Rees, M. (2007). Peer review from the students' perspective: Invaluable or invalid? *Composition Studies, 35*(2), 71-85.
- Breuch, L. K. (2004). *Virtual peer review: Teaching and learning about writing in online environments.* Albany: State University of New York Press.
- Broad, B. (2003). *What we really value: Beyond rubrics in teaching and assessing writing.* Logan: Utah State University Press.
- Bruffee, K. A. (1984). Collaborative learning and the "conversation of mankind". *College English, 46*(7), 635-652.
- Cengage Learning. (2014). Write Experience leads to improved writing skills. Retrieved from http://assets.cengage.com/pdf/wp_write-experience.pdf
- Chen, N., Wei, C., Wu, K., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education, 52*(2), 283-291. doi:10.1016/j.compedu.2008.08.007

- Cheng, K., Liang, J., & Tsai, C. (2015). Examining the role of feedback messages in undergraduate students' writing performance during an online peer assessment activity. *The Internet and Higher Education*, 25, 78-84.
doi:10.1016/j.iheduc.2015.02.001
- Chong, M. R., Goff, L., & Dej, K. (2012). Undergraduate essay writing: Online and face-to-face peer reviews. *Collected Essays on Learning and Teaching*, 5, 69.
- Conference Board, Corporate Voices for Working Families, the Partnership for 21st Century Skills, and the Society for Human Resource Management (2006). Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce. Retrieved from http://www.p21.org/storage/documents/FINAL_REPORT_PDF09-29-06.pdf
- College Board. (2004). Writing: A ticket to work...or a ticket out. Retrieved from http://www.collegeboard.com/prod_downloads/writingcom/writing-ticket-to-work.pdf
- Conference on College Composition and Communication Committee for Best Practices in Online Writing Instruction (2013). A position statement of principles and example effective practices for online writing instruction. Retrieved from <http://www.ncte.org/library/NCTEFiles/Groups/CCCC/OWIPinciples.pdf>
- Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing*, 18(1), 100-108.
doi:10.1016/j.asw.2012.11.001
- Cooper, M. M., & Selfe, C. L. (1990). Computer conferences and learning: Authority, resistance, and internally persuasive discourse. *College English*, 52(8), 847-869.
doi:10.2307/377388
- Davidson, K. (2010). The integration of cognitive and sociocultural theories of literacy development: Why? how? *Alberta Journal of Educational Research*, 56(3), 246-256.
- Deane, M., & Guasch, T. (2015). *Learning and teaching writing online: Strategies for success* doi:10.1163/9789004290846
- Deane, P. (2013). On the relation between automated essay scoring and modern views of the writing construct. *Assessing Writing*, 18(1), 7-24.
doi:<http://dx.doi.org.libproxy.umflint.edu/10.1016/j.asw.2012.10.002>
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. Boston: Heath.

- Diederich, P. B., & National Council of Teachers of English. (1974). *Measuring growth in English*. Urbana, Ill: National Council of Teachers of English.
- DiGiovanni, E., & Nagaswami, G. (2001). Online peer review: An alternative to face-to-face? *ELT Journal*, 55(3), 263.
- DiPardo, A., & Freedman, S. W. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research*, 58(2), 119-149. doi:10.2307/1170332
- Dykman, C. A., & Davis, C. K. (2008). Online education forum: Part two - teaching online versus teaching conventionally. *Journal of Information Systems Education*, 19(2), 157.
- Dykman, C. A., & Davis, C. K. (2008). Online education forum: Part two - teaching online versus teaching conventionally. *Journal of Information Systems Education*, 19(2), 157.
- Elbow, P. (1981). *Writing with power: Techniques for mastering the writing process*. New York: Oxford University Press.
- Elliot, S. (2013). *Handbook of automated essay evaluation*. Routledge.
- Emig, J. A., Goswami, D., & Butler, M. (1983). *The web of meaning: Essays on writing, teaching, learning, and thinking*. Montclair, N.J: Boynton/Cook Publishers.
- Emig, J. (1971). *The composing processes of twelfth graders*. The National Council of Teachers of English: Urbana.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, Mass: MIT Press.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322. doi:10.2307/1170785
- Figl, K., Bauer, C., Mangler, J., & Motschnig, R. (2006). Online versus face-to-face peer team reviews. 7-12. doi:10.1109/FIE.2006.322469
- Fitzpatrick, C. Y. (2006). *Peer review in an online technical writing course* (Ed.D.). Available from ProQuest Dissertations & Theses Global. (305335364).
- Flores, J. F. (2006). The first letter in individual: An alternative to collective online discussion. *Teaching English in the Two Year College*, 33(4), 430.

- Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365.
- Flynn, E. A. (2011). Re-viewing peer review. *Writing Instructor*.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2015). *How to design and evaluate research in education*. New York, NY: McGraw-Hill Education.
- Francis, E. (2012). Traditional colleges embrace online classrooms. *Arkansas Business*, 29(26), 1B.
- Gee, J. P. (1990). *Social linguistics and literacies: Ideology in discourses*. New York; London [England]: Falmer Press.
- Gee, J. P. (2001). Reading as situated language: A sociocognitive perspective. *Journal of Adolescent & Adult Literacy*, 44(8), 714-725. doi:10.1598/JAAL.44.8.3
- Gere, A. R. (1987). *Writing groups: History, theory, and implications*. Carbondale: Published for the Conference on College Composition and Communication [by] Southern Illinois University Press.
- Golub, J. N. (2005). *More ways to handle the paper load: On paper and online*. Urbana, Ill: National Council of Teachers of English.
- Graham, S., Gillespie, A., & McKeown, D. (2013). Writing: Importance, development, and instruction. *Reading and Writing*, 26(1), 1-15. doi:10.1007/s11145-012-9395-2
- Graham, S., & Harris, K. R. (2016). A path to better writing. *The Reading Teacher*, 69(4), 359-365. doi:10.1002/trtr.1432
- Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24(4), 443-461. doi:10.1016/j.compcom.2007.03.002
- Hacker, T. (1996). The effect of teacher conferences on peer response discourse. *Teaching English in the Two-Year College*, 23, 112.
- Hairston, M. (1982). The winds of change: Thomas kuhn and the revolution in the teaching of writing. *College Composition and Communication*, 33(1), 76-88. doi:10.2307/357846
- Hamilton, M. (2000). Expanding the new literacy studies: Using photographs to explore literacy as social practice, 16-34.

- Hancock, T. (2010). Web 2.0 opens worldwide possibilities in education. *Momentum*, 41(1), 14.
- Handsfield, L. J., Dean, T. R., & Cielocha, K. M. (2009). Becoming critical consumers and producers of text: Teaching literacy with web 1.0 and web 2.0. *The Reading Teacher*, 63(1), 40-50. doi:10.1598/RT.63.1.4
- Harris, M. (1992). Collaboration is not collaboration is not collaboration: Writing center tutorials vs. peer-response groups. *College Composition and Communication*, 43(3), 369-383. doi:10.2307/358228
- Hawkins, M. R. (2004). *Language learning and teacher education: A sociocultural approach*. Clevedon: Multilingual Matters.
- Hewett, B. L. (2000). Characteristics of interactive oral and computer-mediated peer group talk and its influence on revision. *Computers and Composition*, 17(3), 265-288. doi:10.1016/S8755-4615(00)00035-9
- Hillocks, G. (1986). *Research on written composition: New directions for teaching*. New York, N.Y.; Urbana, Ill: National Conference on Research in English.
- Ho, M., & Savignon, S. J. (2007). Face-to-face and computer-mediated peer review in EFL writing. *CALICO Journal*, 24(2), 269-290.
- Humphrey, R. C., Walton, M. D., & Davidson, A. J. (2014). "Im gonna tell you all about it": Authorial voice and conventional skills in writing assessment and educational practice. *The Journal of Educational Research*, 107(2), 111-122. doi:10.1080/00220671.2013.788990
- Indrisano, R., & Squire, J. R. (2000). *Perspectives on writing: Research, theory, and practice*. Newark, Dela: International Reading Association.
- Islam, M. M., & Hoque, A. S. M. L. (2012). Automated essay scoring using generalized latent semantic analysis. *Journal of Computers*, 7(3), 616. doi:10.4304/jcp.7.3.616-626
- Jiang, J., & Yu, Y. (2014). The effectiveness of internet-based peer feedback training on Chinese EFL college students' writing proficiency. *International Journal of Information and Communication Technology Education*, 10(3), 34-46. doi:10.4018/ijicte.2014070103
- Journell, W. (2012). Walk, don't run — to online learning. *The Phi Delta Kappan*, 93(7), 46-50. doi:10.1177/003172171209300711

- Kaufman, J. H., & Schunn, C. D. (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
- Kellogg, R. T. (1993). Observations on the psychology of thinking and writing. *Composition Studies/Freshman English News*, 21(1), 3.
- Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology*, 101(1), 136-160. doi:10.1037/a0013097
- Ko, S. S., & Rossen, S. (2010). *Teaching online: A practical guide* (3rd ed.). New York: Routledge.
- Lan, Y., Wang, J., Hsu, S., & Chan, T. (2011). Peer feedback in online writing system. (pp. 126-129). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-23456-9_23
- Li, M., & Zhu, W. (2013). Patterns of computer-mediated interaction in small writing groups using wikis. *Computer Assisted Language Learning*, 26(1), 61-82. doi:10.1080/09588221.2011.631142
- Liard-Muriente, C., & Robinson, C. (2015). The write experience in economics: A case study from Central Connecticut State University. *International Advances in Economic Research*, 21(4), 453-465. doi:10.1007/s11294-015-9543-y
- Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes*, 2(3), 193-227. doi:10.1016/S1475-1585(03)00025-0
- Liu, S., & Kunnan, A. J. (2016; 2015). Investigating the application of automated writing evaluation to Chinese undergraduate English majors: A case study of WriteToLearn. *CALICO Journal*, 33(1), 71. doi:10.1558/cj.v33i1.26380
- Lockhart, C., & Ng, P. (1995). Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45(4), 605-651. doi:10.1111/j.1467-1770.1995.tb00456.x
- Lu, J., & Zhang, Z. (2012). Understanding the effectiveness of online peer assessment: A path model. *Journal of Educational Computing Research*, 46(3), 313-333. doi:10.2190/EC.46.3.f
- Lu, R., & Bol, L. (2007). A comparison of anonymous versus identifiable e-peer review on college student writing performance and the extent of critical feedback. *Journal of Interactive Online Learning*, 6(2), 100-115.

- Lundstrom, K., & Baker, W. (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
- Mabrito, M. (1991). Electronic mail as a vehicle for peer response: Conversations of high- and low-apprehensive writers. *Written Communication*, 8(4), 509-532. doi:10.1177/0741088391008004004
- MacArthur, C. A., Graham, S., Fitzgerald, J. (2015). *Handbook of writing research* (Second ed.). New York: The Guilford Press.
- Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think? *ELT Journal*, 46(3), 274-284. doi:10.1093/elt/46.3.274
- McGarr, O., & Clifford, A. M. (2013). 'Just enough to make you take it seriously': Exploring students' attitudes towards peer assessment. *Higher Education*, 65(6), 677-693. doi:10.1007/s10734-012-9570-z
- Meyer, K. A. (2003). Face-to-face versus threaded discussions: The role of time and higher-order thinking. *Journal of Asynchronous Learning Network*, 7(3), 55-65.
- Miller, G. A. (2003). The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, 7(3), 141-144. doi:10.1016/S1364-6613(03)00029-9
- Moloudi, M. (2011). Online and face-to-face peer review: Measures of implementation in ESL writing classes. *Asian EFL Journal*, 52, 4-23.
- Murray, D. (1982). *Learning by teaching: Selected articles on writing and teaching*. Montclair, N.J: Boynton/Cook.
- National Center for Education Statistics. (2012). Writing 2011: National assessment of educational progress at grades 8 and 12. Retrieved from <https://nces.ed.gov/nationsreportcard>
- Nelson, M. M., & Schunn, C. D. (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
- Oliver, K. (2010). *Integrating web 2.0 across the curriculum*. Boston: Springer US. doi:10.1007/s11528-010-0382-7
- Purcell-Gates, V., Jacobson, E., & Degener, S. (2004). *Print literacy development: Uniting cognitive and social practice theories*. Cambridge, Mass: Harvard University Press.

- Qiu, M., & McDougall, D. (2013). Foster strengths and circumvent weaknesses: Advantages and disadvantages of online versus face-to-face subgroup discourse. *Computers and Education*, 67, 1-11. doi:10.1016/j.compedu.2013.02.005
- Rendahl, M., & Breuch, L. K. (2013). Toward a complexity of online learning: Learners in online first-year writing. *Computers and Composition*, 30(4), 297-314. doi:<http://dx.doi.org/10.1016/j.compcom.2013.10.002>
- Rollinson, P. (2005). Using peer feedback in the ESL writing class. *ELT Journal*, 59(1), 23-30. doi:10.1093/elt/cci003
- Rosenbaum, J. (1999). Using read-aloud protocols as a method of instruction. *Teaching Legal Research and Writing*, 7, 105-109.
- Rudner, L. M., Garcia, V., & Welch, C. (2006). An evaluation of the IntelliMetric[SM] essay scoring system. *Journal of Technology, Learning, and Assessment*, 4(4).
- Rumelhart, D. E., McClelland, J. L., & University of California, San Diego. PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, Mass: MIT Press.
- Russell, S. (1999). Clients who frequent Madam Barnett's emporium. *The Writing Center Journal*, 20(1), 61-72.
- Sadler, D. R. (1998). Formative assessment: Revisiting the territory. *Assessment in Education: Principles, Policy & Practice*, 5(1), 77-84. doi:10.1080/0969595980050104
- Selfe, C. (1989). Creating a computer-supported writing facility: A blueprint for action. *Computers & Composition*: Houghton.
- Sengupta, S. (1998). Peer evaluation: "I am not the teacher.". *ELT Journal*, 52(1), 19-28. doi:10.1093/elt/52.1.19
- Shen, D., Cho, M., Tsai, C., & Marra, R. (2013). Unpacking online learning experiences: Online learning self-efficacy and learning satisfaction. *The Internet and Higher Education*, 19, 10-17. doi:10.1016/j.iheduc.2013.04.001
- Shermis, M. D., & Burstein, J. (2013). *Handbook of automated essay evaluation: Current applications and new directions*. GB: Routledge Ltd - M.U.A. doi:10.4324/9780203122761
- Simmons, J. (2003). Responders are taught, not born. *Journal of Adolescent & Adult Literacy*, 46(8), 684-693.

- Simpson, G., & Clifton, J. (2016). Assessing postgraduate student perceptions and measures of learning in a peer review feedback process. *Assessment & Evaluation in Higher Education*, 41(4), 501-14. doi:10.1080/02602938.2015.1026874
- Sommers, N. (1982). Responding to student writing. *College Composition and Communication*, 33(2), 148-156.
- Spear, K. I. (1988). *Sharing writing: Peer response groups in English classes*. Portsmouth, NH: Boynton/Cook Publishers.
- Street, B. V. (1984). *Literacy in theory and practice*. New York: Cambridge University Press.
- Street, B. V. (2003). What's new in new literacy studies? *Current Issues in Comparative Education*, 5(2), 77-91.
- Stroble, E. J. (1987). *Use of electronic mail to facilitate peer group response during the writing process* (Ph.D.). Available from ProQuest Dissertations & Theses Global. (303613930).
- Sullivan, N., & Pratt, E. (1996). A comparative study of two ESL writing environments: A computer-assisted classroom and a traditional oral classroom. *System*, 24(4), 491-501. doi:10.1016/S0346-251X(96)00044-9
- Szeto, E. (2014). A comparison of Online/Face-to-face students' and instructor's experiences: Examining blended synchronous learning effects. *Procedia - Social and Behavioral Sciences*, 116, 4250-4254. doi:10.1016/j.sbspro.2014.01.926
- Tai, H., Lin, W., & Yang, S. C. (2015). Exploring the effects of peer review and teachers' corrective feedback on EFL students' online writing performance. *Journal of Educational Computing Research*, 53(2), 284-309. doi:10.1177/0735633115597490
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68(3), 249-276.
- Topping, K. J. (2009). Peer assessment. *Theory into Practice*, 48(1), 20-27. doi:10.1080/00405840802577569
- Tseng, S., & Tsai, C. (2007). On-line peer assessment and the role of the peer feedback: A study of high school computer course. *Computers & Education*, 49(4), 1161-1174. doi:10.1016/j.compedu.2006.01.007

- Tuzi, F. (2004). The impact of E-feedback on the revisions of L2 writers in an academic writing course. *Computers and Composition: An International Journal for Teachers of Writing*, 21(2), 217-235. doi:10.1016/S8755-4615(04)00008-8
- Valenti, S., Neri, F., & Cucchiarelli, A. (2003). An overview of current research on automated essay grading. *Journal of Information Technology Education : Research*, 2, 319-330.
- Vanderhoven, E., Raes, A., Schellens, T., & Montrieux, H. (2012). Face-to-face peer assessment in secondary education: Does anonymity matter? *Procedia - Social and Behavioral Sciences*, 69, 1340-1347. doi:10.1016/j.sbspro.2012.12.071
- Villamil, O. S., & Guerrero, M. C. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics*, 19(4), 491-514. doi:10.1093/applin/19.4.491
- Villanueva, V. (2003). *Cross-talk in comp theory: A reader*. National Council of Teachers of English: Urbana.
- Voogt, J., Erstad, O., Dede, C., & Mishra, P. (2013). Challenges to learning and schooling in the digital networked world of the 21st century: Learning and schooling in a digital world. *Journal of Computer Assisted Learning*, 29(5), 403-413. doi:10.1111/jcal.12029
- Vorobel, O. (2013). *A case study of peer review practices of four adolescent english language learners in face-to-face and online contexts* (Ph.D.). Available from ProQuest Dissertations & Theses Global. (1430500718).
- Vygotskiĭ, L. S., & Cole, M. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Walther, J. B. (1992). A longitudinal experiment on relational tone in computer-mediated and face to face interaction. *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, 4, 220-231. doi:10.1109/HICSS.1992.183433
- Wang, J., Hsu, S., Chen, S. Y., Ko, H., Ku, Y., & Chan, T. (2014). Effects of a mixed-mode peer response on student response behavior and writing performance. *Journal of Educational Computing Research*, 51(2), 233-256. doi:10.2190/EC.51.2.e
- Wang, S., & Vasquez, C. (2012). Web 2.0 and second language learning: What does the research tell us? *CALICO Journal*, 29(3), 412-430. doi:10.11139/cj.29.3.412-430
- Warnock, S. (2009). *Teaching writing online: How and why*. Urbana: National Council of Teachers of English.

- Warschauer, M. (1996). Comparing face-to-face and electronic discussion in the second language classroom. *CALICO Journal*, 13(2-3), 7-7.
- Wuensch, K. L., Aziz, S., Ozan, E., Kishore, M., & Tabrizi, M. H. N. (2008). Pedagogical characteristics of online and face-to-face classes. *International Journal on ELearning*, 7(3), 523-532.
- Yancey, K. B. (2004). Made not only in words: Composition in a new key. *College Composition and Communication*, 56(2), 297-328. doi:10.2307/4140651
- Yang, Y. (2010). Students' reflection on online self-correction and peer review to improve writing. *Computers & Education*, 55(3), 1202-1210. doi:10.1016/j.compedu.2010.05.017
- Yang, Y. (2011). A reciprocal peer review system to support college students' writing. *British Journal of Educational Technology*, 42(4), 687-700. doi:10.1111/j.1467-8535.2010.01059.x
- Zhang, M., Hao, S., Xu, Y., Ke, D., & Peng, H. (2014). Automated essay scoring using incremental latent semantic analysis. *Journal of Software*, 9(2), 429. doi:10.4304/jsw.9.2.429-436
- Zhu, W. (1994). *Effects of training for peer revision in college freshman composition classes* (Ph.D.). Available from ProQuest Dissertations & Theses Global. (304144301).
- Zupanc, K., & Bosnic, Z. (2015). Advances in the field of automated essay evaluation. *Informatica*, 39(4), 383-395.