5-1-2017

A CASE STUDY OF PEER ASSESSMENT IN A MOOC-BASED COMPOSITION COURSE: STUDENTS' PERCEPTIONS, PEERS' GRADING SCORES VERSUS INSTRUCTORS' GRADING SCORES, AND PEERS' COMMENTARY VERSUS INSTRUCTORS' COMMENTARY

Lan Thi Vu
Southern Illinois University Carbondale, vulanlequydon@gmail.com

A CASE STUDY OF PEER ASSESSMENT IN A MOOC-BASED COMPOSITION COURSE: STUDENTS' PERCEPTIONS, PEERS' GRADING SCORES VERSUS INSTRUCTORS' GRADING SCORES, AND PEERS' COMMENTARY VERSUS INSTRUCTORS' COMMENTARY

by

Lan Vu

B.A., Quy Nhon University, 2001
M.A., Southern Illinois University Carbondale, 2011

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy Degree

Department of English
College of Liberal Arts
Southern Illinois University Carbondale
May 2017

DISSERTATION APPROVAL

A CASE STUDY OF PEER ASSESSMENT IN A MOOC-BASED COMPOSITION COURSE: STUDENTS' PERCEPTIONS, PEERS' GRADING SCORES VERSUS INSTRUCTORS' GRADING SCORES, AND PEERS' COMMENTARY VERSUS INSTRUCTORS' COMMENTARY

By
Lan Thi Vu

A Dissertation
Submitted in Partial Fulfillment of the Requirements for the
Doctor of Philosophy Degree

Approved by:

Lisa McClure, Co-chair
Jane Cogie, Co-chair
Peter Fadde
Laura Halliday
Elizabeth Klaver

Graduate School
Southern Illinois University Carbondale
December 2, 2016

AN ABSTRACT OF THE DISSERTATION OF

Lan Vu, for the Doctor of Philosophy degree in English, presented on December 2, 2016 at Southern Illinois University Carbondale.

TITLE: A CASE STUDY OF PEER ASSESSMENT IN A MOOC-BASED COMPOSITION COURSE: STUDENTS' PERCEPTIONS, PEERS' GRADING SCORES VERSUS INSTRUCTORS' GRADING SCORES, AND PEERS' COMMENTARY VERSUS INSTRUCTORS' COMMENTARY

MAJOR PROFESSORS: Dr. Lisa McClure and Dr. Jane Cogie

Although the use of peer assessment in MOOCs is common, there has been little empirical research about peer assessment in MOOCs, especially composition MOOCs. This study aimed to address issues in peer assessment in a MOOC-based composition course, in particular student perceptions, peer-grading scores versus instructor-grading scores, and peer commentary versus instructor commentary. The findings provided evidence that peer assessment was well received by the majority of student participants from their perspective as both peer evaluators of other students' papers and as students being evaluated by their peers. However, many student participants also expressed negative feelings about certain aspects of peer assessment, for example peers' lack of qualifications, peers' negative and critical comments, and unfairness of peer grading. Statistical analysis of grades given by student peers and instructors revealed a consistency among grades given by peers but a low consistency between grades given by peers and those given by instructors, with the peer grades tending to be higher than those assigned by instructors. In addition, analysis of peer and instructor commentary revealed that peers' commentary differed from instructors' on specific categories of writing issues (idea development, organization, or sentence-level). For instance, on average peers focused a greater percentage of their comments (70%) on sentence-level issues than did instructors (64.7%),

though both groups devoted more comments to sentence-level issues than to the two other issue categories. Peers' commentary also differed from instructors' in the approaches their comments took to communicating the writing issue (through explanation, question, or correction). For example, in commenting on sentence-level errors, on average 85% of peers' comments included a correction as compared to 96% of instructors' comments including that approach. In every comment category (idea development, organization, sentence-level), peers used a lower percentage of explanation—at least 10% lower—than did instructors. Overall, findings and conclusions of the study have limitations due to (1) the small size of composition MOOC studied and small sample size of graded papers used for the analysis, (2) the lack of research and scarcity of document archives on issues the study discussed, (3) the lack of examination of factors (i.e. level of education, cultural background, and English language proficiency) that might affect student participants' perception of peer assessment, and (4) the lack of analysis of head notes, end notes, and length of comments. However, the study has made certain contributions to the existing literature, especially student perception of peer assessment in the composition MOOC in this study. Analysis of the grades given by peers and instructors in the study provides evidence-based information about whether online peer assessment should be used in MOOCs, especially composition MOOCs and what factors might affect the applicability and consistency of peer grading in MOOCs. In addition, analysis of the data provides insights into types of comments students in a composition MOOC made as compared to those instructors made. The findings of the study as a whole can inform the design of future research on peer assessment in composition MOOCs and indicate questions designers of peer assessment training and practice in such MOOCs could find helpful to consider.

Keywords: Peer Assessment, Composition MOOCs, Students' Perception, Peer-grading Scores, Instructor-grading Scores, Peers' Comments, Instructors' Comments.

ADKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee chairs – Dr. Jane Cogie and Dr. Lisa McClure for their valuable support and advice. Without their guidance and persistent help this dissertation would not have been possible. My special thanks are also for other members of the committee – Dr. Klaver, Dr. Fadde, and Dr. Halliday for their insightful feedback and encouragement.

My deepest appreciation is for my beloved family: my amazing spouse Phu, my wonderful sons Andy and Nam, my encouraging parents, my bothers, and my sister for supporting me spiritually throughout writing this dissertation.

To all of you, thank you!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

FIGURES

CHAPTER 1

INTRODUCTION

In the 1960s, Douglas Engelbart and his colleagues at Stanford Research Institute delivered a proposal for creating a method of massive worldwide sharing of information by making use of the widespread personalization of computers (MouseSite). The rationale behind this idea was to give access to the same education as American Ivy League students to learners around the world who are interested in high quality education. Decades later, this idea seemed to be applied when Sebastian Thrun and Peter Norvig from Stanford University offered a course on artificial intelligence to 160,000 learners (Vu et al., 2013). In 2012, Thrun and Norvig created one of the first platforms offering Massive Open Online Courses (MOOCs), and from this beginning, other big name universities such as Harvard, MIT, and Berkeley followed with their own MOOC offerings (Vu et al., 2013). A thorough search on the most popular online platforms offering MOOCs (i.e. Coursera, Udacity, EdX, OpenupEd.edu, Canvas) shows this new trend continues, with numerous courses now offered on science and engineering disciplines, humanities and liberal arts. There are also some courses on particular language skills like writing (*First-year Composition 2.0* offered by Georgia Institute of Technology; *Writing II: Rhetorical Composing* offered by Ohio State University; *English Composition I: Achieving Expertise* offered by Duke University; *College Writing 2x: Principles of Written Communication* offered by University of California, Berkeley). With a subject so specific and individuated as writing, educators might raise questions about (1) how workable MOOCs in composition can be particularly on matters related to pedagogy, assessment, course credit, as well as retention rate, (2) how MOOCs will be financed, and (3) how technology will be used to facilitate teaching and

learning. In a 2013 article in *Slate,* Rees (2013), a professor at Colorado State University–Pueblo, argued specifically against the pedagogical viability of MOOCs:

> How do you teach tens of thousands of people anything at once? You don't. What you can do over the Internet this way is deliver information, but that's not education. Education, as any real teacher will tell you, involves more than just transmitting facts. It means teaching students what to do with those facts, as well as the skills they need to go out and learn new information themselves.

Rees is not alone in concluding as she does here that the aim of MOOCs to educate thousands of people at once has little chance of succeeding, at least if the definition of education remains the construction of--rather than just the delivery of--knowledge. Pappano (2012) also presented major problems of MOOCs, for example problems with meaningful assessment, cheating, and learners' ill preparation. With this level of criticism and little to slow down the increase in the number of MOOCs, including MOOCs in composition offered by Coursera, research into whether or not these critiques are valid is needed, particularly since at the moment, little research exists. Thus, studies to explore aspects of MOOCs are needed, particularly, as in this case, in the challenging area of composition.

When MOOCs were introduced, I enrolled in several MOOCs including composition MOOCs as a student, and seriously participated in all of the activities in those classes. My purposes of taking MOOCs were (1) to learn, (2) to have access to the course materials to understand the course design, and (3) to see how MOOCs especially composition MOOCs work. The third purpose was the most important because I, together with other Fulbright scholars, administered an online learning program – E-Center for Professional Development (E-Center), and E-Center attempted to offer MOOCs. E-Center had its first chance to run and test the effectiveness and smoothness of running a MOOC in 2012 when we offered a course on Computer-assisted Language Learning open for unlimited enrollments. After two months of

course promotion and enrollment, the course had 512 learners from 23 countries. By today's definition, a course with 512 students cannot be considered a MOOC if compared with the huge number of students enrolled in Coursera, Canvas, or Edx, currently three of the largest MOOC providers in the world. On average, a MOOC at Coursera has approximately 43,000 students (Ferenstein, 2014). We had one instructor who was in charge of lecture delivery and two co-instructors who were in charge of monitoring grading, facilitating the course's forum discussion, and answering students' emails.

In 2013, E-Center continued to offer open online courses in ESL including Writing courses to thousands of students from many parts of the world. For open online writing courses, we realized that one of the biggest challenges was how to grade students' writing assignments. Instructors, limited in number, could not complete the grading mission. We created an online peer grading system in this course to solve the grading problem. With a standard grading rubric for each writing assignment and some instructions on peer assessment (called training on peer assessment), three students graded one assignment, using the same standard grading rubric. We set up our grading mechanism in which only students who submitted the assignment could do peer grading. From what we perceived - peer assessment seems to be one of the most controversial issues in composition MOOC pedagogy and feedback given by peers might be thought to be of poor quality, as a partial solution to this issue, we had instructors provide students with detailed rubrics and guidelines during the training on peer assessment. However, many students did not do peer grading; some students' assignments were graded by only one or two peers not three peers as expected. This practice may make the students' feedback on writing assignments unreliable and invalid. Moreover, the practice of using online peer assessment was still quite new in the online learning field, and certainly to us, so we were not really sure how

valid and reliable the grades offered by the fellow students were. These challenges of running composition MOOCs as well as my experiences as a MOOC student spurred me to investigate future composition MOOCs, especially aspects of assessment.

Statement of the Problem and Purpose of the Study

The evolution of traditional online learning or online learning 1.0 to online learning 2.0 has created both opportunities and challenges for higher education (Sloan C, 2013; Grosseck, 2009; McLoughlin, & Lee, 2007). In the traditional online learning, online courses are quite similar to traditional face-to-face courses in term of the ratio of students to instructors. However, in online learning 2.0, of which MOOCs, including MOOCs in composition, are a typical representative, an online instructor can have up to several thousand students in his or her course. Grading in such massive open online courses becomes a burden or a mission impossible for even the most dedicated professors, with an army of equally dedicated teaching assistants. Because not all assignments can be designed in auto-graded formats, and artificial intelligence grading programs are not well regarded by educators and researchers (Condon, 2013; Deane, 2013; Bridgeman, Trapani, & Yigal, 2012; Byrne, Tang, Truduc, & Tang, 2010; Chen, & Cheng, 2008; Cindy, 2007; Benett, 2006; Cheville, 2004; Chodorow, & Burstein, 2004), online peer grading is utilized, especially for composition and other courses in humanities. This online peer grading practice shifts the traditional grading authority from the instructor to the learners and poses many unanswered questions about the reliability and validity of online peer-reviewed grades in an open online learning setting. In the few studies on peer grading (i.e. Cho et. al, 2006; Sadler & Good, 2006; Bouzidi & Jaillet, 2009), findings show a high consistency among grades assigned by

peers and a high correlation between peer grading and teacher grading, which indicates that peer grading has been found to be a reliable and valid assessment tool. However, these findings are generally based on the context of college courses with small or moderate enrollments; in addition, none of these studies on peer grading were conducted based on the context of composition courses or courses taught by composition teachers. Until the time I conducted the present study, I have uncovered only one empirical study on peer grading in MOOC context. Lou, Robinson and Park (2014) examined peer grading assignments from a Coursera MOOC called *Maps and the Geospatial Revolution,* not a composition MOOC, and found that grading scores given by peer students were fairly consistent and highly similar to the scores given by instructors. Nevertheless, Lou, Robinson and Park' study (2014) on peer grading referred to a Coursera MOOC named Maps and the Geospatial Revolution, not a composition MOOC.

It is evident that the practice of peer assessment (peer review and peer grading) is being carried out in the context of teaching and learning writing in a massive open online setting – MOOCs in composition. Although there is a lack of evidence for the effectiveness of peer assessment in MOOCs in composition, that peer assessment has been used or suggested in other settings (i.e. face-to-face or traditional online courses) likely accounts for the decision to apply it in the MOOC setting in addition to the problem teachers would have grading the students in a MOOC solo. However, there has been no study conducted on the issue of peer assessment specifically in a MOOC in composition (at least to my knowledge). Plus, there has been no study about peer grading in composition MOOCs, looking at its viability as a learning assessment tool, and other issues like students' perception. To address this research need, I collected surveys, conducted interviews, and accumulated statistical data on students' and instructors' grades and comments from a seven–week MOOC-based composition course - ESL/EFL Writing for

Academic and Professional Development offered by E-Center for Professional Development. This present study was intended to unpack issues in peer assessment; specifically, the study sought to answer the following questions that speak to the viability of the use of peer assessment in this context:

1. What perceptions do the students in a MOOC-based composition course have toward peer assessment of their own and their peers' drafts?
2. Is there a significant difference between grades given by peers in response to student drafts and those given by the instructors?
3. To what degree and in what way do comments/feedback given by peers in response to student drafts differ from those given by the instructors?

CHAPTER 2

LITERATURE REVIEW

In light of the newness of MOOCs as a context for student learning, the literature review for this study will draw upon research from a variety of fields including education, literacy, composition, and studies of writing in other disciplines. The present study looks into certain issues of peer assessment in composition in a MOOC context, specifically students' perception and differences in grades and comments given by students and instructors. To that aim, the literature review first discusses the inception of MOOCs and MOOCs in composition and then the uses of peer assessment for education and writing instruction with the goal of improving the outcomes of student writing. The literature review continues with other aspects of peer assessment such as students' perceptions, the relationship between grades given by the students and the instructors, and the commentary by the students and the instructors.

1. Inception of MOOCs and MOOCs in Composition

MOOCs (Massive Open Online Courses) hit the popular press in 2012 and were considered one of the most important emerging developments in educational technology in 2013 (New Media Horizon, 2013). McAuley, Stewart, Siemens, and Cormier (2010) considered a MOOC as an online course characterized by (1) open and free registration, (2) a publicly shared curriculum, and (3) open-ended outcomes; they stated that a MOOC heavily relies on social networking, and accessible online resources and that it is often facilitated by leading practitioners in the field of study. According to Educause (2012), a MOOC is "a model for delivering learning content

online to any person who wants to take a course, with no limit on attendance." Adding to the definition of a MOOC as a natural byproduct of open teaching and learning, Masters (2011) further elaborated that a massive course should be open to thousands of learners simultaneously engaged in a single course. "In fall 2011 Ng's online course on Applied Machine Learning at Stanford University enrolled over 100,000 students — and that course was probably the unofficial birth of the MOOC, the moment that first caught everybody's attention" (Porter, 2013, p. 2).

Due to its early stage of development, there has not been a commonly accepted definition of a MOOC. However, MOOCs have been thought to share three main features. The first feature is their "openness" to all participants. MOOC participants do not need to be registered as students, and they are not required to pay a fee. This feature of MOOCs might be changed soon when for-profit MOOC providers such as Udacity and Coursera have joined the field, and then the "open access" of MOOCs might be defined again (Vu et al., 2013). Another feature is their onlinedness: digital designs in MOOCs allow a wide variety of course materials including images, videos, readings, and the like, which can be posted by instructors and students. The third main feature of MOOCs is their massiveness and scalability because MOOCs tend to admit an unlimited number of participants, which is different from face-to-face and traditional online courses (Vu et al., 2013).

Following this trend of education have been proposals for offering MOOCs in different disciplines. In 2012, Bill & Melinda Gates Foundation awarded 12 grants with a total of three million dollars to organizations and universities to develop MOOCs for a variety of courses – from developmental math to English Composition. Innovative institutions and faculty receiving Gates Foundation grants for developing MOOCs for composition include Denise K. Comer,

Director of First-Year Writing, Duke University; Rebecca Burnett, Director of Writing and Communication Program and Karen Head, Director of the Communication Center, Georgia Institute of Technology; Kay Halasek, Director of Second-Year Writing, The Ohio State University. Comer's course –English Composition I: Achieving Expertise – opened for over 8000 of enrollments in late December 2012. This course aimed to provide an introduction to and foundation for the academic reading and writing characteristic of college. For Head's MOOC in May 2013, the First-Year Composition 2.0 course was designed to help learners develop a better process and gain confidence in written, visual, and oral communication as well as create and critique college-level documents and presentations. In Head's course, learners would draft and revise the following assignments: a personal essay, an image, and an oral presentation. Halasek's Writing II: Rhetorical Composing course introduced a variety of rhetorical concepts, ideas and techniques to inform and persuade audiences, which help learners become more effective consumers and producers of written, visual, and multimodal texts. In Halasek's course, learners can exchange words, ideas, talents, and support. All of these composition MOOCs were created through the MOOC platform Coursera.

After these MOOCs in composition were completed, questions as to whether the courses were a success were raised. Posts by Karen Head (2013) on the freshman-composition MOOC that she taught with Gates Foundation funding appeared in *The Chronicle*. The stats were disappointing. Head wrote:

> If we define success by the raw numbers, then I would probably say No, the course was not a success. Of course, the data are problematic: Many people have observed that MOOCs often have terrible retention rates, but is retention an accurate measure of success? We had 21,934 students enrolled, 14,771 of whom were active in the course. Our 26 lecture videos were viewed 95,631 times. Students submitted work for evaluation 2,942 times and completed 19,571 peer assessments (the means by which their writing was evaluated). However, only 238 students received a completion certificate—meaning that they completed all assignments and received satisfactory scores.

Head (2013) and her team had some hypotheses for the reasons there were so few students completing the course. One of the hypotheses is students not completing all three major assignments could not pass the course. Plus, many students failed to complete the course because they had technical problems and cultural issues, or because they joined the course late. In spite of the terrible retention rates, Head (2013) claimed that the course was a success if the success is defined by lessons learned in the course design and presentation. She stated "…if we define success by lessons learned in designing and presenting the course, I would say Yes, it was a success. From a pedagogical perspective, nobody on our team will ever approach course design in the same way. We are especially interested in integrating new technologies into our traditional classes for a more hybrid approach." In regard to technology integration, Head suggested not rushing to teach another MOOC soon, especially when the technology is lacking for courses in subject areas like writing, which have such strong qualitative evaluation requirements. "Too often we found our pedagogical choices hindered by the course-delivery platform we were required to use, when we felt that the platform should serve the pedagogical requirements. Too many decisions about platform functionality seem to be arbitrary, or made by people who may be excellent programmers but, I suspect, have never been teachers" (Head, 2013). In general, from what Head (2013) pointed out, technical issues seem to have certain effects on the MOOCs' working, for example retention rate and course delivery.

Similar to Head's MOOC, the MOOC in composition offered by The Ohio State University was not perceived successful in terms of the low retention rate. According to the final report on the MOOC Writing II: Rhetorical Composing, posted on Educause (2013b), among 32,765 total enrolled students, there were 55% engaging in the course at least once and 444 receiving the completion statements. However, what was considered to be a success was the

inaugural experience using the WExMOOC platform for students to review peers' writing and to receive analytics, although training students to produce strong reviews was challenging. As the report stated, approximately one fourth of the course's students used the platform to do the peer review and to receive analytics that helped students compare their own skills with the skills of others (Educause, 2013b). Another area of this course considered to be successful is training students to make critical reviews of peers' work, although offering effective training seemed to be challenging (Educause, 2013b).

For the other MOOC in Composition offered by Duke University, although there are no official results or evaluations published, many learners have shared their experiences and comments on personal websites or blogs. Steve Krause (2013), a professor at Eastern Michigan University, posted on his blog under the title "The end of the Duke Composition MOOC: again, what did we learn here?": "If the point of the Duke Composition MOOC was to see if it could provide an *educational* experience that could compete with a traditional first year writing course taken in order to earn the credential of a college degree, then the answer is clearly no." Krause pointed out problems of this specific MOOC, for example the way the course was organized and the method of evaluating students' work. Krause wrote:

> I thought there was almost no connection between Comer's video lectures/teaching, the online discussion, and the assignments themselves, and every week there seemed to be something added to the Coursera web site to further confuse things……. A multiple choice test for a writing course wouldn't work, machine grading of writing (as it currently exists, at least) doesn't work, and peer evaluation as the *only* means of evaluation for a credit-bearing course doesn't work. So logistically, it seems to me that the way we teach first year writing right now is probably the most efficient/effective way to do it at large universities.

Apart from the first three MOOCs in composition mentioned above, there have been a great number of other MOOCs in composition recently offered by different universities. In general, the birth of MOOCs, including MOOCs in composition, has created a trend in teaching

and learning – a configuration for delivering learning content online to an unlimited number of learners who want to take the course. However, with the nature of MOOCs like massiveness and scalability, MOOCs presents certain challenges, one of which is assessment.

2. MOOC Assessment Types:

Handling Massive Student Enrollments while Sustaining Quality

Prior to discussions of assessment in MOOCs and composition MOOCs in particular, it is necessary to address assessment in education to help establish a context for evaluating the viability and quality of peer assessment in the composition course examined in this study. In education in general, assessment is crucial for learning and for providing a quality learning environment (Gardiner, 1994; Bransford, Brown, & Cocking, 1999; Bain, 2006). Gardiner (1994) stated that "assessment is essential not only to guide the development of individual students but also to monitor and continuously improve the quality of programs, inform prospective students and their parents, and provide evidence of accountability to those who pay our way " (p. 109). In a comprehensive review of the literature on assessment and feedback, Evans (2013) examined the nature of assessment in higher education and its importance in student learning. Assessment and feedback are seen as an essential way "to facilitate students' development as independent learners who are able to monitor, evaluate, and regulate their own learning, allowing them to feed-up and beyond graduation into professional practice (Ferguson, 2011, as cited in Evans, 2013, p. 72). Also in this review, Evans (2013) classified types of assessment feedback and discussed effective feedback practice including e-assessment feedback, self-assessment feedback and peer assessment feedback. E-assessment feedback, which can be

made synchronously, asynchronously, or at a distance, includes "formative and summative feedback delivered or conducted through information communication technology of any kind, encompassing various digital technologies including CD-ROM, television, interactive multimedia, mobile phones, and the Internet" (Gikandi, Morrow, & Davis, 2011, as cited in Evans, 2013, p. 85). As cited in Evans (2013), e-assessment can be applied to a large number of learners with the application of technologies, making it possible for anywhere and anytime (Gikandi et al., 2011; Juwah et al., 2004). Nicol (2008) provided an example of how technology can enhance this effectiveness of e-assessment in first-year education: "in one first-year psychology module at the University of Strathclyde, redesigned as part of the REAP project, a single teacher was able to organize rich, regular peer feedback dialogue for over 560 students on a series of online essay-writing tasks without a workload increase but with significant learning gains compared with previous years" (p.8). However, the impact of e-assessment on student performance was found to be varied, depending on variables like teaching method, nature of learning, individual strengths, training on peer online feedback, technology, etc. (Evans, 2013). For self-assessment feedback, an important component of self-regulation, Evans (2013) cited Archer's (2010) work that "there is no evidence for the effectiveness of self-assessment and [Archer] has recommended the need to move from individualized, internalized self-assessment to self-directed assessment utilizing and filtering external feedback with support" (p. 88). Therefore, if self-assessment feedback is to attain the necessary quality, according to Tara (2008), as cited in Evans (2013), requires carefully organized training for students and appropriate scaffolding. In regard to peer assessment, advocates argue that this approach enables learners to engage in learning and to develop self-assessment skills (Davies, 2006; Nicol & MacFarlane-Dick, 2006; Orsmond, 2006; Topping, 2010; Vickerman, 2009; Xiao & Lucking,

2008, as cited in Evans 2013). However, for the successful implementation of peer assessment, it is essential to explore variable impacts of peer feedback on student performance for example students' discomfort (Papinczak, Young, and Groves, 2007, as cited in Evans, 2013), students' maturity (Loddington et al., 2009, as cited in Evans, 2013), students' varied ability (Van Zundert et al., 2010, as cited in Evans, 2013). Since this present study looks into peer assessment, more discussions surrounding peer assessment feedback will be discussed in the next section, specifically studies on the students' perceptions of peer assessment, the grades given by the students and the instructors, and the commentary by the students and the instructors.

In the context of e-learning like MOOCs, types of assessments discussed in Evans (2013) such as e-assessment, self-assessment and peer assessment seem to be applied, but in varying degrees and forms. Given the massive number of students and thus the challenges of peer, let alone instructor-based assessment, MOOCs have made use of web-based technologies such as online communication forums, computer-based assessments and video presentations to help students have access to the course materials, open discussions, and assessments. Due to the high enrollments in MOOCs, assessing students' work becomes a burden for even the most dedicated professors, with an army of equally dedicated teaching assistants. Thus, it is not surprising that technological solution to the problem of high enrollment grading has been sought by some through the use of "computers to score and provide feedback on student activities and assessment and thus rely heavily on multiple choice questions, formulaic problems with correct answers, logical proofs, computer code, and vocabulary activities" (Balfour, 2013, p. 40). However, not all assignments can be designed in auto-graded formats, and thus finding a more viable and quality sustaining means for providing feedback and grading written assignments in MOOCs has become a major issue discussed in media.

Another technological application, AES (Automatic Essay Scoring) was also explored for its usefulness to MOOCs assessment. According to *The New York Times* website on April 5, 2013, EdX introduced the integration of AES application in its MOOCs (non-composition MOOCs) to assess students' writing (Balfour, 2013). In the context of MOOCs, MIT and Harvard's EdX announced that AES technology could be applied in online courses with over 150,000 students enrolled (Balfour, 2013). Actually, AES, a measurement technology in which computers evaluate written work, was first successfully developed in 1973 by colleagues at the University of Connecticut (Shermis et al., 2010). According to Shermis et al. (2010), this type of technology, developed in different models and built individually for individual writing assignments, has the "capacity to receive text via a web page and return feedback to both a student user and comprehensive data base that may be accessed by teachers" (p. 6), using statistical models to correlate with the ways humans rate essays. In reviews of AES applications, Shermis et al. (2010) claimed that AES can offer feedback on sentence level issues such as grammatical errors, word usage errors, vocabulary, as well as organization levels such as thesis statements, supporting ideas, conclusions and style. However, AES is not capable of assessing complex novels using metaphors, slang, and humor (Graesser & McNamera, 2012). In addition, although AES application has been proved positive in certain levels, the National Council of Teachers of English (NCTE) opposed the use of machines to assess students' writing because:

> Computers are unable to recognize or judge those elements that we most associate with good writing (logic, clarity, accuracy, ideas relevant to a specific topic, innovative style, effective appeals to audience, different forms of organization, types of persuasion, quality of evidence, humor or irony, and effective uses of repetition, to name just a few). Using computers to "read" and evaluate students' writing (1) denies students the chance to have anything but limited features recognized in their writing; and (2) compels teachers to ignore what is most important in writing instruction in order to teach what is least important.
>
> (NCTE Position Statement on Machine Scoring, 2013)

In light of the limitations of the AES application, a few studies (e.g. Fosmire, 2010; Likkel, 2012; and Balfour, 2013) suggested using structured computer-regulated peer evaluation for its greater benefits to students. According to Fosmire (2010), Calibrated Peer Review™ (CPR), a web-based application, is "a program that can significantly enhance the ability to integrate intensive information literacy exercises into large classroom settings. CPR is founded on a solid pedagogic base for learning, and it is formulated in such a way that information skills can easily be inserted" (p. 147). CPR manages the peer review process (including submitting, distributing, and compiling grades) and scores the performance of peer reviewers (Balfour, 2013). According to Balfour (2013), CPR allows big numbers of students to: (1) compose and submit their essays by a deadline to the CPR software server; (2) receive the training of giving feedback and learn critical points in essays by scoring instructor-provided essays with a multiple choice rubric; (3) grade their peers' essay using the same rubric; (4) evaluate their own essays; and (5) access to all the feedback from their peers who reviewed their work. In reviews of this online writing software, Likkel (2012) stated that the CPR is a useful tool for assigning writing assignments in large college classes – it helps students become more confident in evaluating their peers' writing as well as their own writing. As cited by Fosmire (2010), Russell concluded that "CPR empowers students to write to learn rather than learn to write…..When students write, they are required to organize their thoughts, make decisions about what is relevant, convey their thoughts, and arrive at conclusions. This means students are active in constructing their understanding of the material" (quoted in Fosmire, 2010, p. 150). However, in a MOOC environment, CPR is found to have technical problems scaling up to multiple tens of thousands of students (Balfour, 2013). Plus, because of the need to employ the rubric on seven essays (three peers' essays, three calibration essays, and one for self evaluation), students seemed to

have difficulty managing feedback on peers' essays and on their own, especially in essays containing more than 750 words, although they do receive some training during the calibration phase (Balfour, 2013). While Balfour's study focuses on peer evaluation in a structured computer-regulated context, similar to Evans (2013) in his exploration of assessment in a range of different settings and kinds of peer involvement, Balfour (2013) argued that it is crucial to have sufficient training for students and appropriate scaffolding.

3. Non-composition Peer Assessment in Non-MOOC and MOOC Contexts

Peer assessment, which has been researched over 30 years, is a common educational practice in a wide range of subject domains including writing in non – MOOC contexts, both classrooms and traditional online classes (with small or moderate enrolments). Baird & Northfield (1992) stated that peer assessment refers to "specific judgments of ratings made by pupils about their achievement, often in relation to teacher–designed categories" (12). Topping et al. (2000) defined peer assessment as "an arrangement for peers to consider the level, value, worth, quality or successfulness of the products or outcomes of learning of others of similar status" (p. 150). Topping (2009) stated that peer assessment activities can be operated in different areas or subjects and be involved in a wide variety of products for example oral presentation, test performance, portfolio, writing, and so forth. Since peer assessment involves numerous applications and a wide range of educational methods, educators and scholars have employed varied terminology to describe this practice. The frequently used terms include "peer assessment" (Falchikov & Goldfinch, 2000, Falchikov, 1995; Topping et. al, 2000; Cho, Schunn, & Wilson, 2006), "peer review" and "peer feedback" (Connors, 1997; Lu & Bol, 2007, Gielen et

al., 2010), "peer response" (DiPardo & Freedman, 1988), "peer marking" (Fry, 1990), "peer criticism" (Holt, 1992), "peer rating" (Magin, 1993), "peer grading" (Sadler, 2006). Despite varied terminology, all of these are used to denote a common practice – students are guided through teaching materials (i.e. checklist, rubric, questionnaires, etc.) to evaluate peers' work on a number of points, to provide feedback to other students on the quality of their work, and in some instances to assign a grade. In this study, the term *peer assessment* is used as an umbrella concept to capture the diversity of definitions and involves any activity in which students make only comments on peers' work, or students only assign a grade to peers' work, or students both make comments on peers' work and assign a grade. To have a better distinction of peer assessment activities, this study will use "peer review/peer response" to refer to an activity in which students make commentary on peers' work and "peer grading" to indicate an activity in which students assign a grade to peers' work. When authors I mention in this study use the term *peer assessment*, I clarify, as author definitions allow, if *peer assessment* refers to peer response or peer grading.

Practices of peer assessment especially in traditional classroom instruction in non-composition contexts have been found to be beneficial in a large body of literature. By gathering different sources through a literature review, Sadler & Good (2006) sorted the potential advantages of peer assessment (peer grading) over teacher grading into four perspectives: logistical, pedagogical, metacognitive, and affective. Logistically, it saves teacher time and results in immediate feedback (Bould, 1989, as cited in Sadler & Good, 2006); in addition, "peers can often spend more time and offer more detailed feedback than the teacher can provide" (Weaver & Cotrell, 1986, as cited in Sadler & Good, 2006, p. 2). Pedagogically, students have opportunities to understand the topic better, develop skills through reading another's answers

(Alexander, Schallert, & Hare, 1991; Black & Atkin, 1996, as cited in Sadler & Good, 2006). Metacognitively, grading helps students become more aware of their strengths, weakness, and progress and develop a capacity to use higher order thinking skills to make judgments. Finally, affective changes can produce a dynamic in classrooms, helping students work more productively and cooperatively (Baird & Northfield, 1992; McLeod, 2001; Pfeifer, 1981; Weaver & Cotrell, 1986; Zoller, Ben-Chaim, & Kamm, 1997, as cited in Sadler & Good, 2006). In regard to benefits of peer assessment, Topping (2009) emphasized gains that either assessors or assesses have from peer assessment (peer response), among which are cognitive and metacognitive gains, improvements in writing, and improvements in peer/or group work; however, Topping failed to specify the rationale and the basis on which the claims were made – i.e. peer assessment applied for what types of assignment in what subjects. For the work online, Bouzidi & Jaillet (2009) carried out an experiment of online peer assessment (peer grading) in which 242 students, enrolled in computer architecture and electrical engineering courses, participated. Bouzidi & Jailet (2009) stated that peer assessment (peer grading) "develop learning at high cognitive levels" that "involve the student in the revision, assessment, and feedback process of work online" (p. 257).

The reliability and validity of peer assessment have been supported in the context of face-to-face and online education (non-MOOC) by some researchers. From their study on student grading compared to teacher grading in four secondary school science classrooms, Sadler & Good (2006) pointed out very high correlations between grades assigned by students and those by the teacher. This study also shows that "the high levels of inter-rater agreement are possible between students and teacher when students grade their own or others' papers," given that "students should be trained to grade accurately" (Sadler & Good, 2006, p. 27). Bouzidi & Jaillet

(2009), carrying out an experiment of online peer assessment in three different courses with the total of 242 students, concluded that peer assessment (peer grading) is equivalent to the assessment by the professor in certain exams of computer science and electrical engineering. In another study on peer assessment of 708 students across 16 different courses from 4 universities, Cho et al. (2006) investigated the validity and reliability of peer-generated writing grades in a scaffolded reviewing context. These 16 courses included Psychology Research Methods, Health Psychology, Cognitive Science, Education, Rehabilitation Sciences, Leisure Studies, History, and the Honors College, in which the writing task that was assigned to students varied, depending on different disciplines. The assigned paper genres consisted of the introduction of a research paper, a proposal for a research study, a proposal for an application of a research finding, and a critique of a research paper. The results show that "the aggregate ratings of at least 4 peers on a piece of writing are both highly reliable and as valid as instructor ratings while (paradoxically) producing very low estimates of reliability and validity from the student perspective" (p. 891). The divergence of these two findings—on peer assessment grading and student perception of peer assessment—point to a contrast relevant to the current study designed as it was to collect both data on student perception of peer assessment for answering research question one and data on peer and instructor grades toward use in answering research question two on the comparability of peer grades to instructors'. Even though the results of Cho et al. (2006) were less reliable in its findings from the students' perspective, based mostly the instructor's perspective, they still maintained that "peer generated grades can be sufficiently reliable and valid to be used widely in university settings. In other words, concerns about reliability and validity (at least the instructor's perspective on them) are probably not a good

reason for preventing instructors from implementing this approach, at least when appropriate scaffolds for peer review are included" (p. 900).

In summary, findings in a number of studies in peer assessment – peer grading in non-composition MOOC contexts (e.g. Cho et al., 2006; Sadler & Good, 2006; Bouzidi & Jaillet, 2009) show a high consistency among grades assigned by peers and a high correlation between peer grading and teacher grading. These finings indicate that within certain contexts like non-composition traditional college classrooms and online courses (these are the contexts in which most of these studies were conducted), peer assessment (peer grading) has been found to be a reliable and valid assessment tool. However, these findings are generally from studies conducted in the context of small or moderate enrollments, and thus do little to help suggest how peer assessment use would function in a MOOC with its massive enrollment [i.e. Sadler & Good's (2006) study on 4 secondary school classes; Cho et al.'s (2006) study on 16 classes with 708 students].

As to the use of peer assessment in MOOCs, not many studies have been completed on their use, though, according to Educause (2013a), peer assessment has been viewed a widely applicable approach to MOOCs of different forms, contents, and products. From a pedagogical perspective, that "students in MOOCs grade each other's work according to the professor's specifications" seems to be "a promising practice" because "it can extend students' learning experience" as does their sharing knowledge through interactions (Boston & Helm, 2012; Kuh, 2012, as cited in Boston & Helm, 2012). In their study on peer grading in a MOOC, Lou, Robinson and Park (2014) examined 1,825 peer grading assignments from the final assignment – mapping assignment – of a Coursera MOOC named Maps and the Geospatial Revolution. The authors found that grading scores given by peer students were fairly consistent and highly similar

to the scores given by instructors. However, researchers have also recognized potential problems with the use of peer assessment in MOOCs. In a class with several thousands of students such as in the context of MOOCs, this practice of peer grading was found to be problematic. Watters (2012) pointed out problems associated with peer assessment grading such as the anonymity of feedback (students do not know who they are assessing and who is assessed), the lack of feedback on peer feedback (there's no way for students to give feedback on that feedback), the lack of community in the course (students do not know much about their peers), and the variability of feedback (many students are not well prepared to give solid feedback to peers and there are issues with English as a second or foreign language).

In a nutshell, the practice of peer assessment in MOOCs is still early in implementation as well as still in the process of being assessed by researchers and educators. In addition, whether or not peer assessments are reliable and valid—or even suggested through more qualitative studies to be a viable practice comparable to instructor assessment—in MOOCs in general and MOOCs in composition in particular in the configurations currently used is in need of further research. The findings on peer assessment in non-MOOCs and MOOCs in composition are the focus of the next section.

4. Peer Assessment in Non-MOOCs and MOOCs in Composition

Because many students have been found to lack writing skills or to display inadequate abilities in writing (Baker, Gersten, & Graham, 2003; Swales & Feak, 2004; Leal, 2012), researchers, educators, and teachers have explored various instructional strategies and techniques used in the practice of teaching and learning writing. Among these strategies is peer assessment,

specifically peer review/peer response, based on theories that emphasize the social nature of language, thought and learning, prime among them Vygotsky (1962, 1978) that learning is a cognitive activity that happens through social interactions and which requires sharing, exchanging, and negotiations. In the field of teaching and learning writing, peer review/peer response is well established as an important theoretical stage of the writing process (Bruffee, 1984; Elbow, 1981) and the activity of having students critique peers' writing becomes common in composition classrooms and composition textbooks [both first language (L1) and second language (L2)].

Theoretically, peer assessment (peer review/ peer response) that involves peer interactions and collaborations will help students construct knowledge and learn effectively (Liu et al., 2002; Lee & Lim, 2012). Also, research on peer assessment (peer review/peer response) supports the capacity of peers for helping each other when they work together, and shows the benefits of peer assessment (peer review/peer response) in helping students revise their papers and get a better understanding of their writing. Marchionda (2004), an English teacher at Massachusetts's Middlesex Community College, who developed a peer review strategy for her college composition class, stated that peer review stimulates a helpful and nonthreatening atmosphere in which students talk about their writing freely and openly. Peer review "frees up the student writers to experiment with new ideas, words, and patterns", and it helps students become aware of writing as a social construct and realize that their writing affects individual readers differently" (Marchionda, 2004). Keh (1990), in her study about feedback in the writing process, claimed that feedback from peers can "save teachers time on certain tasks, freeing them for more helpful instruction". Keh concluded that peers can "gain a greater sense of audience with several readers (i.e. readers other than the teacher)…and learn more about writing through

critically reading others' papers" (Keh, 1990, p. 296). Thompson (1981), an English teacher who faced an increased number of student essays to grade, presented some promising advantages of peer assessment (peer grading) especially when students graded peers' work in composition classes. Thompson stated that "trained students not only grade papers competently and reliably but also write better as a result of this training" (p. 172) and that "trained students can be trusted as graders in composition research" (p.172).

As peer review/peer response has been an important tool in teachers' strategies for teaching composition in general (students' providing feedback on peers' preliminary drafts so that the student writers can have a wider sense of audience and thus work toward improving their writing), it is not surprising that the practice of peer review/peer response is being carried out in the context of teaching and learning writing in a massive open online setting – MOOCs in composition. In addition to the application of peer review/peer response, MOOCs in composition do utilize peer grading. Although there is a lack of evidence for the effectiveness of peer assessment (peer review/peer response and peer grading) in MOOCs in composition, that peer assessment has been used or suggested in such settings as face-to-face and traditional online courses likely accounts for the decision to apply it in the MOOC setting. Also likely relevant to composition MOOCs turning to peer assessment is the fact that their large enrollments exceed the assessment capacity (i.e. evaluating and grading) of instructors; indeed, the inability for teachers to grade so many papers is likely one of the largest, if not the largest factor responsible for MOOCs turning to peer grading. Offering one of the most high profile MOOCs in composition, Coursera (2012) claims that peer assessment in MOOCs in composition, which guides students in using instructor-constructed rubrics to evaluate and provide feedback on peer work, "offers an effective learning tool by giving students the opportunity to play the role of both

"student" and "teacher." ….and allows students to "sharpen their critical thinking skills through the active comparison" of peers' answers to the instructor's rubric" (Coursera, 2012). Relevant to note here in relation to these claims is the kind of structure of peer assessment Coursera provides. As stated on Coursera's website, the process of peer assessment including several phrases is outlined below (Table 1). Upon completion of these phases, the grades are calculated by the median of all the peer grades.

Table 1

Phases of Peer Assessment in a composition MOOC

Submission phase: During this phase, the assignment is open, and you have unlimited chances to submit your assignment (without penalty), up until the deadline. In submitting the assignment, you agree to abide by the Honor Code.
Evaluation phase: This phase begins shortly after the submission deadline. It consists of several components:

- **Training** (not required for most classes): A small number of classes may require you to practice grading sample submissions of the assignment. If this page shows up for you, you will have to pass the exercise before moving on to evaluate your peers.
- **Peer evaluation**: In every peer assessment, you will be required to grade a predetermined number of your peers' submissions in accordance with the rubric. These submissions will be randomly selected from the class pool. Failure to complete the requisite number of evaluations will result in a grade penalty.
- **Self evaluation**: In some classes, you will also be required to assess your own submission against the instructor's rubric. If this page shows up for you, and you fail to complete a self-evaluation, you may incur a grade penalty.

Results phase: Shortly after the evaluation deadline, you will be able to see your final grade and grade breakdown, as determined by your peers' evaluations of your work.

(Coursera – Peer Assessment)

In its most basic form, Coursera's common rubrics particularly designed for each of the assignments are used for students to assess peers' papers. Student's submitted papers are randomly distributed to a handful of peer raters (usually from three to five). Written peer

comments are provided along with assigning a score. Finally, students receive peer comments and the median of peer ratings.

At least according to what a few researchers (Haaga, 1993; Rushton, Ramsey, & Rada, 1993; Rada et al., 1994) believed in peer assessment in traditional contexts (e.g. face-to-face classrooms, and traditional online courses), Coursera's peer assessment system seems to help (a) MOOC students learn how to write and write better and (b) instructors minimize their grading workload. However, further investigation into Coursera's use of peer assessment seems warranted, particularly since the training for peer assessors Coursera notes offering is limited and optional (see above). As Falchikow (2001, 2005) cautioned in considering peer assessment, though in a traditional setting, establishing good quality peer assessment for responding to student writing requires more work in organizing peer assessment activities and in training students how to provide constructive feedback. And, as noted previously, Watters (2012) suggests the need to take other aspects into account, such as the anonymity of feedback, the lack of feedback on feedback, and the lack of community.

Although there have been studies that show peer assessment of writing to be beneficial in some contexts of teaching and learning writing as discussed above, there has been no study conducted on the issue of peer assessment specifically in a MOOC in composition (at least to my knowledge). Also, there has been no study about peer grading in composition MOOCs, looking at reliability and/or validity, and other issues like students' perception. This present study was intended to address issues in peer assessment in a MOOC in composition, by looking into the practice of peer assessment in the context of MOOCs in composition, particularly the students' perceptions, the grades given by the students and the instructors, and the commentary by the students and the instructors.

5. Students' Perceptions of Peer Assessment

In composition classrooms and composition textbooks, the process of having students give feedback to others' papers has become common. There have been a lot of studies discussing the impact of peer assessment (mainly peer review/peer response) for both assessors and assesses. Of these quite a number have focused on students' perception of peer assessment in different contexts of learning, not surprisingly since students are key stakeholders in peer assessment. Thus their perception of this activity can help provide another angle from which to understand such elements of their participation as their perception of themselves and other peers as able to play this role and the credibility of the comments and grades given. This section, then, will discuss studies from various fields that have focused on student perceptions of peer assessment to provide further understanding of this aspect of the current study.

A few empirical studies in different disciplines have looked into what students themselves thought about peer assessment and their participation in that activity. In a study on 250 students' perceptions about (non-MOOC) online peer assessment (including peer review and peer grading) for undergraduate and graduate writing across the disciplines in ten courses, Kaufman and Schunn (2011) found that students sometimes thought peer assessment (peer grading) unfair and that students believed they were not qualified to grade peers' work. Kaufman and Schunn added that "students perceptions about the fairness of peer assessment drop significantly following students' experience in doing peer assessment" (p. 387). Different from Kaufman and Schunn (2011), Simkin and Ramarapu (1997) examined student perceptions in peer –reviewed grading in which computer science undergraduates assigned final grades to each others' term papers and found that students trusted peers and felt comfortable with the practice of

peer grading. Regarding students' perceptions in ESL composition, Nelson and Carson (1998) investigated Chinese and Spanish-speaking students' perceptions of peer responses in an ESL composition class and found that students expressed preference for peers' negative comments over positive comments because "all students agreed that negative comments (i.e., comments pointing out problems) were desirable because they contributed to the writers' making changes" (p. 122). In Mangelsdorf's (1992) study of peer review on 40 students in ESL composition classrooms, approximately half responses from the students thought that peers' comments were helpful and expressed positive views of peer review; about 30% of students' responses indicated ambivalence about the quality of peers' suggestions. Mangelsdorf's study (1992) also revealed that students had negative reactions – students did not trust in their peers' abilities to critique essays. Similar to Mangelsdorf's findings on students' negative reactions toward peer review, ESL students' in Zhang's (1995) study on affective advantage of peer feedback in the ESL writing class sometimes questioned the efficacy of peer responses, their ability to give constructive feedback, or their peers' competency to evaluate writing work.

For scholarship in English composition (L1 composition), there are not many studies that measure students' perceptions about peer assessment neither peer review nor peer grading, but such studies are more frequent in L2 composition courses and sometimes mixed groups of both English native students and ESL students (Brammer & Rees, 2007). Brammer and Rees (2007) further explained this lack of research on L2 student perception of the peer review, stating that "Perhaps because peer response is practically instinctive to those of us who teach writing, few have felt the need to study the student perspective. Instead, studies have focused on the quality of peer comments, their effect on the revision process, and the best methods for conducting peer review" (p. 274). In their study on peer review from the students' perspective, Brammer and

Rees (2007) stated that peer review was used in most first year writing classrooms, but "most students find peer review not very helpful' (p. 278) and many students did not trust their peers to respond to their papers. Most students "expressed concerns about classmates' dedication and ability to peer review" (Brammer & Rees, 2007, p. 283). However, Brammer and Rees also emphasized that students did express some positive impressions about peer review – "students who were prepared to carry out peer review through two or more teaching methods (e.g., handout, lecture, and paper demonstration) were more likely to find peer review helpful" ….and to be "more confident in their ability to peer review" (p. 280). In a study of both L1 and L2 students' perceptions and attitudes toward peer review, Murau (1993) found that both groups appreciate the value of peer review because it can help them with, in students' words "grammar" (L2 writer), "vocabulary" (L2 writer), "surface errors" (L1 writer), "overall organization" (L2 writer), "ideas" (L1 writer) (p. 74-75). While students recognize the benefits of peer review, 20% would prefer to review their papers with a teacher or tutor "because they respect [the teacher's] knowledge of a language better than with a peer (L1 writer) or because 'he can explain me in technical grounds' (L2 writer)"; 20% would not review their writing with anyone because they felt nervous and uncomfortable (Murau, 1993, p. 75- 76). In regard to the perceived effects about peer assessment (peer review/peer response and peer grading) in a Coursera MOOC, Lou et al. (2014) concluded that approximately 63% of students believed that the peer assessment activity (peer review/peer response and peer grading) was helpful in developing their thinking competences. The researchers also found that about 62% of the students thought the grades given by peers were fair and the feedback was useful.

Overall, much research in different disciplines has been done regarding students' perceptions and attitudes of peer assessment. In composition scholarship, students' perception

has been explored at the level of peer review/peer response as part of the writing process, not peer grading. In the MOOC context, the Lou et al. study (2014) was one of very few empirical studies found in recent literature to investigate students' perceptions of peer assessment (peer review/peer response and peer grading). Their results actually refer to a Coursera MOOC named Maps and the Geospatial Revolution, not a composition MOOC. Since empirical research on students' perceptions of peer assessment in composition MOOCs is very limited and in need of further research, this present study was intended to look into the practice of peer assessment in the context of MOOCs in composition, particularly the students' perceptions, in addition to how the grades and commentary given by the students and the instructors compare. For the purpose of the present study, the following section discusses another related aspect - how teachers and peers respond to student writing.

6. Peer Comments vs. Teacher Comments in Responding to Student Writing

In the teaching of writing, commenting on student writing is considered the most widely used method for responding to student writing, and this activity usually takes teachers a lot of time. Theoretically and ideally, teachers of writing want to help students with thoughtful commentary so that students can know if their writing has communicated their intended meaning and if they communicate their ideas effectively. Whether or not teachers follow through on this ideal was a concern focused on in a study of teacher comments by Connors and Lunsford (1993):

> Most composition teachers know what the field says is important-our public "tropes," so to speak. We talk and write often of purpose, of audience, of organization, of proof, of process and invention and revision and so on. These words fill our journals, our professional books, our conferences, and especially our textbooks. But do we really follow through? Do comments on papers show us

> acting on these public tropes, giving them more than lip service? Or do we have more genuine and less overt agenda.
>
> (Connors and Lunsford, 1993, p. 209)

In their analysis of teachers' comments on 3,000 students' papers, Connors and Lunsford (1993) found that 77% of the papers contained global comments. Of this 77%, 56% of the papers contained comments on the effectiveness and the lack of supporting details, evidence or examples; 28% of the papers had comments about overall organization such as issues of introductory and concluding sections; 11% of the papers had comments that could be considered to be about purpose; and 6% of papers mentioned something about audience considerations such as voice and tone. Connors and Lunsford (1993) also discovered that many teachers provided comments that deal with specific formal elements, including 33% of all papers mentioning sentence structure, 18% paragraph structure, 16% paper format, 7% documentation, 6% quotations, and 6% source materials. From their analysis of teachers' comments, Connors and Lunsford (1993) concluded that "teachers are genuinely involved in trying to help their students with rhetorical issues in their writing" and they [teachers whose work was looked at] clearly "cared about how their students were planning and ordering writing" (p. 218). Connors and Lunsford (1993) added what they called "bad news" to the so-called "Tropics of Commentary": "Many teachers seem still to be facing classroom situations, loads, and levels of training that keep them from communicating their rhetorical evaluations effectively…. The teachers whose comments we studied seem often to have been trained to judge student writing by rhetorical formulae that are almost as restricting as mechanical formulate" (p. 218). Overall, Connors and Lunsford's study provided interesting observations about teachers' comments; however, there's no information provided about who the teachers were, who the students were, what draft was analyzed, or what the assignment was. Sommers's (1982) study on styles of commenting – a

study analyzing the written feedback of L1 teachers - revealed that "teachers' comments can take students' attention away from their own purposes in writing a particular text and focus that attention on the teachers' purpose in commenting" (p. 149). In other words, teachers can appropriate student writing especially when they mark up errors in style, diction, and usage in a first draft, which make students understand that these errors need to be addressed before the meaning of the text is attended to (Sommers, 1982). Sommers (1982) suggested that teachers of writing should respond to student writing as any reader would and that comments should focus on logic, disruptions of meaning, or missing information. However, Sommers' (1982) did not explain in her article how the teacher commentary was analyzed and on what basis she approached to the conclusions. Regarding teachers' comments on student writing, Kehl (1970) urged teachers to respond "in a distinctly human voice, with sincere respect for the writer as a person and a sincere interest in his improvement as a writer (976). In a similar qualitative study to Sommers', but in L2 settings, Zamel (1985) examined 15 teachers' comments on 105 L2 student essays and found that teachers in the study treat students' texts as final products to be edited and that "the teachers overwhelmingly view themselves as language teachers rather than writing teachers; they attend primarily to surface-level features of writing and seem to read and react to a text as a series of separate sentences or even clauses, rather than as a whole unit of discourse" (p. 86). In general, these studies cumulatively accentuate the tension in the literature on teacher response to student writing between the ideal as defined by composition theory and the actual follow-through in teachers' responses. Yet whatever distance there may be between these two, there are findings, as will be clear below, that reveal interesting distinctions between teacher and student comments, depending on the student population graded.

In regard to student comments in responding to peer writing, there is a general assumption that students have a tendency to read for surface and mechanical errors, which are usually referred to as lower order concerns (LOCs), and thus tend to ignore higher order concerns (HOCs) such as audience, organization, and development of ideas (Keh, 1990). Huff & Kline (1987), through introducing the process of holistic peer grading, explained that more mature composition students were likely to have confidence to critique such concerns as development, ideas, and mechanical errors, and that advanced composition students could recognize and solve problems with style and content. However, Huff and Kline (1987) noted that peer responses can be problematic because the feedback can be "blatantly useless, uniformed, and often thoroughly unconstructive" (p. 150). In a study on comparison of teacher and student responses to written work in an ESL writing class, Caulk (1994) found that teacher comments "tended to be general and were often aimed at the whole piece of writing, rather than one part" whereas student comments "tended to be very specific and rarely contained suggestions for the whole piece of writing" (p. 184).

In conclusion, as was discussed above in the relevant sections of this literature review of peer assessment (sections 4, 5, and 6), the previous research that does exist on this topic is limited and from a variety of fields and tends to focus on a range of different perspectives on peer review relevant to this study such as students' responding to peer writing in relationship to students' perceptions of peer responses, students' roles in peer response, training students for peer responses, and strategies for successful peer response activity. For example, Carson and Nelson (1998) looked into ESL students' perceptions of effectiveness in peer response groups; Min's (2005) study was about training students to become successful peer reviewers; Stanley (1992) explored aspects of coaching student writers to be effective evaluators. However there

was not sufficient research on student perception of peer assessment in composition classes. In addition, very little research compares teacher and student comments: the one article I found that does compare teacher and student comments is Caulk's (1994) noted above, focusing on L2 student and teacher comments. Also, previous literature does not approach peer assessment from as many angles in that, unlike the current study, they do not combine a comparison of student and teacher comments in relationship to analysis of student perceptions, and teacher versus student grading. Because there has been a lack of empirical research in a MOOC composition context on student perceptions of peer assessment, on characteristics of student comments in responding to peer writing as well as on comparison of student comments versus teacher comments, the present study was intended to explore these concerns: to what degree and in what way comments/feedback given by the online peer reviewers in response to student drafts differ from those given by the instructors, in addiiton to other aspects of the students' perceptions as well as the grades given by the students and the instructors.

CHAPTER 3

RESEARCH METHODOLOGY

In this chapter, I will discuss in detail the research methodology that has been adopted in this study of the use of peer assessment in a MOOC-based Composition Course. The MOOC-based Composition Course that the present study examines, ESL/EFL Writing for Academic and Professional Success, was offered by the E-Center for Professional Development (E-Center) from December 1, 2014 to January 18, 2015. The E- Center, founded in 2011 through a grant from the U.S Embassy in Vietnam, has provided open online learning programs with a series of different courses to learners in Vietnam and in other countries, including the ESL/EFL Writing for Academic and Professional Success course. I chose to study this course because it had features essential to researching issues related to the MOOCs' practice of using peer assessment as opposed to assessment by instructors: (1) it had characteristics of a MOOC – openness, onlinedness, and massiveness, (2) it used peer assessment as a learning assessment tool typical of larger composition MOOCs offered at a range of large universities. In addition, as an administrator of the E-Center, I was more easily able to gain access to the course's data needed for the purpose of the study. For a clear presentation of the course in which the present study was conducted, in this Research Methodology chapter, I will first explain the research methods used to study this composition MOOC and then provide an overview of the context for the study, the E-Center and the MOOC-based composition course – ESL/EFL Writing for Academic and Professional Success offered in the E-Center. With this overview of the study's context in place, I will present the participants and other details specific to the composition MOOC section I studied. From there, I will describe the data collection and data analysis.

1. Overview of Research Methods

The purpose of the present study was to investigate issues of peer assessment in a composition MOOC context focusing in particular on questions concerning (1) students' perception of peer assessment, (2) differences in grades given by peers and instructors, and (3) differences in comments made by peers and instructors. To best address this combination of research questions, this study employed a mixed methods approach, combining paradigms of quantitative and qualitative research to look into connections and disconnections among the different data sets and provide the differing perspectives from which to answer the research questions. According to Creswell and Clark (2007), "Mixed methods research is a research design with philosophical assumptions as well as methods of inquiry. As a methodology, it involves philosophical assumptions that guide the direction of the collection and analysis of data and the mixture of qualitative and quantitative approaches in many phases in the research process. As a method, it focuses on collecting, analyzing, and mixing both quantitative and qualitative data in a single study or series of studies" (p.5). As noted by Johnson and Onwuegbuzie (2004), a mixed methods approach offers the researcher a better understanding of the problem than if any one dataset is used alone.

In this study, then, of a composition MOOC in the E-Center, for research question one about students' perceptions on peer assessment in a composition MOOC, the mixed methods involved data collection in an iterative process (survey to interview), which means data collected in one phase contributes to those collected in another phase (Creswell & Clark, 2007; Creswell et al., 2003). Specifically, in phase one, I collected survey data with both closed and open-ended questions, and in phase two, I collected the interview data. The collection and analysis of

qualitative data from individualized interview questions would help explain, or elaborate on, the quantitative results in phase one more in depth (Creswell, 2003). For research question two as to differences between grades given by peers and grades given by instructors, quantitative data was collected from grades assigned by both the instructors and peer students in the course. For research question three about differences in comments given by peers and instructors, I collected the qualitative data from the comments by peer graders and instructors on a selective sample of students' graded essays. These qualitative data were then coded for a comprehensive dataset through the frequency of themes (Onwuegbuzie & Teddlie, 2003).

In sum, research questions of the study were answered through analysis of peers' survey responses, interviewees' responses, and grades and comments given by student and instructor participants. I employed the mixed methods approach to look into connections and disconnections of data sets and to understand better the problems and benefits associated with peer assessment. The following section provides an overview of the context for the study, the E-Center and the MOOC-based composition course – ESL/EFL Writing for Academic and Professional Success offered in the E-Center

2. Overview of Procedures and Configuration of the E-Center's MOOC-based Composition Course Studied

In order to study this section of the E-Center's composition MOOC, I asked the executive manager of E-center for Professional Development (E-center) for approval to conduct the study on a MOOC-based composition course offered by the E-center. The approval from the E-Center's executive provided me with access to learners' and instructors' database including logs

of learning activities, discussion forums, students' submitted assignments, grades and comments provided by both peer students and instructors, demographic information, and email addresses. The following (sections 1.1 – 1.4) present an overview of ESL/EFL Writing for Academic and Professional Success MOOC-based Course.

2.1. Course Objectives and Course Work

As noted above, The MOOC-based composition course studied for this research, ESL/EFL Writing for Academic and Professional Success, ran for seven-weeks, from December 1, 2014 to January 18, 2015. Offered by the E-Center for Professional Development (E-Center), it was designed to help students develop the fluency, focus, analytical skills needed to become successful writers. In this course, students were expected to learn and practice the strategies and processes that successful writers employ as they work to accomplish specific purposes, for example comprehension, instruction, entertainment, persuasion, investigation, problem-resolution, evaluation, explanation, and refutation, all of which would help prepare learners for academic and professional communication. There were three major writing assignments that the students needed to complete (1) Application Essays, (2) Summary and Evaluation Essays, and (3) Argumentative Essays.

Students taking this course had two options. If they were interested in receiving a Statement of Accomplishment, they had to score at least 60% in the course and participate in peer assessment. If students were not trying to receive a Statement of Accomplishment, it was fine for them to audit the course and only review the materials. The course materials remained available two months after the course ended, so students could also go through the course at a later date or slower pace. During the course, there were live virtual sessions in which keynote

speakers talked about particular topics relevant to the writing assignments students were working on. These sessions were held live at a specific time, and students were provided with the link to attend these sessions.

Students' grades were based on participation including completion of quizzes, material access (15%), three major writing assignments (25% per assignment), and a multiple-choice final exam (10%). The grading scales were A (90% - 100%), B (80% - 89%), C (70% - 79%), D (60% - 69%), and F (0% - 59%). Discussion forum was available within each writing assignment, so if students had questions on the video lectures or quizzes in each writing task, they could post their questions. The following is the course syllabus (Table 2).

Table 2.

The Course Syllabus

Weeks	Topics / Activities	Assignments Due
Week 1 (Dec 1 – Dec 7, 2014)	Introduction to the course; syllabus An overview of writing and writing process: - Pre-test - Thinking about audience and purpose - Paragraph/Essay organization - Writing as a process: planning, drafting, revising	Do the pretest and read all of the materials by Dec 7, 2014
Week 2 (Dec 8 – 14, 2014)	**Writing: Application Essay** - Lectures on application essays (videos + attachments) - Writing prompt (attachment) - First draft of Application Essay - Peer assessment and self-assessment (rubric, guidelines of assessment) **Live Virtual Session:** Questions & Answers (Scheduled time: To be announced)	**First Draft of Application Essay** (Due: by midnight Wednesday, Dec 10, 2014) **Peer and self assessment of Application Essay (**Due: by Saturday, Dec 13, 2014)
Week 3 (Dec 15 - 21, 2014)	**Writing: Application Essay** (Contd.) - Checklist of Application Essay Assignment - Second draft of Application Essay (revise the first draft for the polished one) - Peer assessment of the polished draft of	**Polished Draft of Application Essay** (Due: by midnight Tuesday, Dec 16, 2014)

	Application Essay *** Live Virtual Session - Keynote Speaker:** How to be successful in writing application essays? (Scheduled time: To be announced)	**Peer Assessment of Application Essay – Polished Draft** (Due: by midnight Saturday Dec 20, 2014)
Week 4 (Dec 22- 28, 2014)	**Writing: Summary and Evaluation Essay** - Lectures on Summary/Evaluation Essay (videos + attachments) - Writing prompt (attachment) - First draft of Summary/Evaluation Essay - Peer assessment and self-assessment (rubric, guidelines of assessment) *** Live Virtual Session - Keynote Speaker:** Avoiding plagiarism (Scheduled time: To be announced)	**First Draft of Summary/ Evaluation Essay** (Due: by midnight Friday, Dec 26, 2014) **Peer and self assessment of Summary/ Evaluation Essay (**Due: by Sunday, December 28, 2014)
Week 5 (Dec 29, 2014 – Jan 4, 2015)	**Writing: Summary/Evaluation Essay** (Contd.) - Checklist of Summary/Evaluation Essay Assignment - Second draft of Summary/Evaluation Essay (revise the first draft for the polished one) - Peer assessment of the polished draft of Summary/Evaluation Essay *** Live Virtual Session - Keynote Speaker:** Evaluation Criteria (Scheduled time: To be announced)	**Polished Draft of Summary/ Evaluation Essay** (Due: by midnight Tuesday, Dec 30, 2014 **Peer Assessment of Summary/ Evaluation Essay – Polished Draft** (Due: by midnight Saturday, Jan 3, 2015)
Week 6 (Jan 5 – 11, 2015)	**Writing: Argumentative Essay** - Lectures on Argumentative Essay (videos + attachments) - Writing prompt (attachment) - First draft of Argumentative Essay - Peer assessment and self-assessment (rubric, guidelines of assessment) *** Live Virtual Session - Keynote Speaker:** Ethos Pathos Logos: Appeals for Effective Arguments (Scheduled time: To be announced)	**First Draft of Argumentative Essay** (Due: by midnight Wednesday, Jan 7, 2015) **Peer and self assessment of Argumentative Essay (**Due: by Saturday, Jan 10, 2015
Week 7 (Jan 12-18, 2015)	**Writing: Argumentative Essay** (Contd.) - Checklist of Argumentative Essay Assignment - Second draft of Argumentative Essay (revise the first draft for the polished one) - Peer assessment of the polished draft of Argumentative Essay	**Polished Draft of Argumentative Essay** (Due: by midnight Wednesday, Jan 14, 2015)

	*** Live Virtual Session - Keynote Speaker:** Making Counter Arguments (Scheduled time: To be announced)	**Peer Assessment of Argumentative Essay – Polished Draft** (Due: by midnight Saturday, Jan 17, 2015
	Final Exam	**Sunday, Jan 18, 2015**

(E-Center)

Each of the major writing assignments in this course followed the same structure including four phases in sequence (beginning at a predefined starting time and ending at a predefined deadline time). First, students prepared and submitted Draft 1 of the assignment. Second, students peer reviewed and assessed Draft 1 of three other students. Third, students submitted the revised version of the assignment (Final Draft) based on the feedback from peers. Last, students assessed peers' Final Draft. Prior to carrying out the peer assessment, students were instructed to go through a training process for understanding and using the peer assessment tool within the MOODLE platform as well as for evaluating their peers' work. Students were encouraged to self-review their work, but the self-assessment was optional. If students failed to submit the assignment, they would not be able to participate in the peer assessment. If students failed to complete the required peer assessment, they would receive no credit for the associated assignment. The following explains typical cycles of each of the major writing assignments:

1. Prepare and submit the essay (Draft 1)
2. Complete the peer assessment process (required training phase, required peer assessing, optional self-assessing)
3. Prepare and submit the essay (Revised version – Final Draft)
4. Complete the peer assessment process (No training phase, required peer assessing, optional self-assessing)

In regard to peer assessment presented in the cycles above, for Draft 1, peer reviewers were required to make comments (marginal comments are required; head and end comments are encouraged but not required) and assign a score corresponding to the criteria designed in the rubric. Based on the feedback from peers on the Draft 1, students submitted the revised version of the assignment (Final Draft) and then participated in assessing peers' final drafts. For final drafts, students were required to assign a score in accordance with the rubric. Making comments was optional at this phase.

2.2. Students and Instructors of the Course

The MOOC-based course the researcher studied– ESL/EFL Writing for Academic and Professional Development - had 4582 registered students with 4521 students actually accessing the course. As shown in Figure 1 below, the majority of students were from Vietnam (55.18%) and China (21.41%), and the others were from Taiwan (8.18%), Indonesia (5.46%), Malaysia (3.45%), Singapore (2.76%), Japan (1.88%), and South Korea (1.68%). All of the students were non-native English speakers.

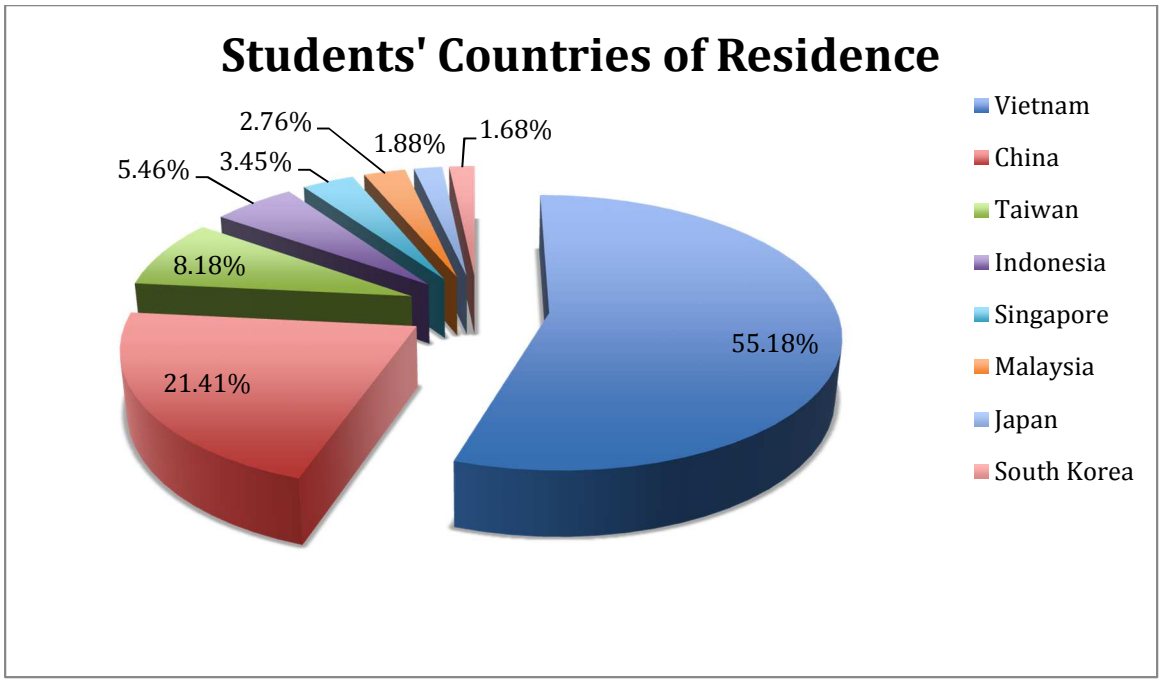

Figure 1. *Students' Countries of Residence*

As explained in Section 3 of this chapter on the participant selection process, only a portion of the total students who enrolled and accessed this course participated all three components of this study. This and all other subsections of Section 1's overview of the course is intended to provide a full context for the workings of the course from which participants were selected for the study.

In this course, there were two instructors (females) in charge of lecture delivery and three instructors (one male and two females) in charge of monitoring peer assessment, facilitating the course's forum discussion, and answering students' emails. All of the instructors were English native speakers with Master degrees in TESOL (Teaching English to speakers of other languages) and at least two years of experience in teaching ESL/EFL writing. Moreover, because the instructors had taught E-Center's writing courses before, they were familiar with the online platform, live virtual classroom (LVC), Moodle-based learning management system (LMS), etc.

used by E-Center. Due to the large enrollment far exceeding the grading capacity of instructors, instructors in this course did not provide grades and feedback to students' work. Instead, the course utilized an online peer assessment process to grade the learners' papers. At the end of the course, there were 334 students (out of the initial 4,521) receiving a completion certificate, which means that they submitted all assignments, did all assigned peer assessments, and received satisfactory scores (Cs and up). Among these 334 students, 115 students had their papers graded by one or two peers and 219 students got all of their papers graded by three peers as expected.

For the purpose of the course and program evaluation of the E-center, as in other MOOC-based course offered by the E-center, in this composition course, the administrators selected a certain number of students' papers and had them graded by the instructors, using the same rubric given to students. The E-Center had five instructors grade papers from the students who received completion certificates, more specifically, from 100 out of 219 students whose papers were graded by three peers. The 100 papers graded by both peer students and instructors were randomly selected from the pool by a computer. Similar to what students assessing their peers' papers were required to do, the instructors both made comments on and assigned a score to students' Draft 1, and assigned a score only to students' final drafts. Overall, the number of papers the instructors were required to grade was 600 (100 Draft 1 and 100 final drafts, for each of the three writing assignments). The instructors' feedback and grades assigned to student drafts were not given to the writers of these drafts but were used for course and program evaluation purposes at the E-center, in addition to the use of the instructor grades and comments built into design of this study.

2.3. Peer Assessment Used in the Course

Part of the required work in the ESL/EFL Writing for Academic and Professional course, the MOOC-based course the researcher studied, was peer assessment. Prior to the process of peer assessment, there were discussions on how to be a constructive critic of other people's work and how to respond to criticism. Students were encouraged to avoid general or vague judgments about the whole work or about its component parts – and encouraged instead to highlight specific ways to improve the work. Discussions on giving feedback were held both through LVCs and the forum. Besides, students were also provided relevant reading materials that helped them understand the qualities of constructive criticism emphasized in training. Within MOODLE platform, students were guided through a training process for understanding and using the peer assessment tool so that they would be then able to use this tool to assess the work of their peers in the course. For each of the major writing assignments, through LVCs and LMS, the instructors discussed with the students criteria defined in the rubric (see Appendix B) and then reviewed two actual student essays to help students understand the process better. Through the review of two actual student essays, students were advised to focus first on global concerns such as idea development, organization, and purpose and second on local concerns such as grammar, vocabulary, and mechanics. In addition, students were assisted to generate constructive and specific feedback. For example, in the LVC on peer assessment training for Writing Assignment 3, one student's comment on a sample essay used in training was limited to the one word *awkward*. The instructor asked students to explain what was awkward and to make that comment more specific. Then, how specific comments could be made was discussed among students, followed by the instructor's model. Students were welcome to ask questions related to what was

discussed by the instructors. During the training phase, to help apply the evaluation criteria on a particular writing assignment more fairly and critically, students were asked to evaluate two to three sample essays. Students' scores on the sample essays were compared with those given by the instructors. Students were the ones to do the comparing through being given copies of the instructors' feedback. Students were considered to pass the training phase if students' scores on each of the categories of the rubric for each major writing assignment were within one point of the instructors' scores. The passing of the training phase was self-evaluated by the students being trained. This training process, used for the course in the study, was part of the regular course design for composition MOOCs in the E-Center.

Through MOODLE platform, students submitted their work corresponding to modules designed for a particular writing assignment. The system randomly assigned these essays to be assessed by other students who had submitted their work, and there might have been a new random assignment of peer graders for each assignment. By that, students could potentially be graded by different peer assessors for each draft. Each peer assessor was supposed to grade three essays from three different students. Also peer assessors were required to assign a score based on particular categories of the rubric. Information from the assignment rubric was "translated" into easy-to-use bulletin menus. For peer assessment on draft 1 for all three writing assignments, in addition to assigning a score in accordance with the rubric, it was mandatory that students make comments on peers' writing (marginal comments were required; head and end comments were optional). Students made comments on peers' drafts using the "tracking" and "comment" commands in MS Word. However, for the peer assessment on the final Draft, it was optional for students to make comments. Depending how quickly students read, the process of peer assessment could take 20 – 30 minutes per essay. Completion of peer assessment was required,

with students not completing the peer assessment receiving 50% of the final score for that assignment.

2.4. Configuration of Students' Participation in the Course Activities

The MOOC-based course the researcher studied– ESL/EFL Writing for Academic and Professional Development –- had a discussion forum where students and instructors could discuss issues related to the course. A total of 236 threads of discussions were posted in the course's discussion forum, most of which were about the activity of peer assessment, technical issues, assignment completion, and assignment discussions.

The course had a total of 27 lecture videos and 7 recorded live virtual sessions (i.e. live sessions during which the teacher was available to discuss specific topics or questions with the students) viewed a total of 24,436 times. In regard to assignment submission, students submitted their drafts for evaluation 1756 times for Application Essay Assignment, 1050 times for Argumentative Essay Assignment, and 768 times for Summary and Evaluation Essay Assignment. Throughout the course, students completed 9065 peer assessments. Many students submitted their writing assignments but did not do peer assessment. As a result, some students' assignments were graded by only one or two peers, not three peers as the course configuration required. MOOCs have been cited in studies as having terrible retention rates (Colman 2013). The particular section of ESL/EFL Writing for Academic and Professional Development in this study is not an exception. At the end of the course, there were 334 students receiving a completion certificate, which means that they submitted all assignments, did all assigned peer assessments, and received satisfactory scores (Cs and up). Possible reasons accounting for the low retention rate would include students unable to pass the class because (1) they chose to

access the materials but not to do assignments, (2) they failed to complete all three major assignments, (3) they submitted their assignments but did not assess peers' work, (4) they struggled with technology – submitting assignments and doing peer assessment online through MOODLE platform, and (5) they had other issues related to cultural reasons, deadline missing, personal busy schedule, etc. Among 334 students receiving a completion certificate, 115 students had their papers graded by one or two peers; the remaining 219 students had all of their papers graded by three peers as expected.

3. Participants of the Study

In the study, the participants included students who enrolled in the MOOC-based composition course – ESL/EFL Writing for Academic and Professional Success - offered by E-center for Professional Development in December 2014. The number of the participants varied according to different components of the study (e.g. survey and interview).

3.1. Survey participants

The selection of participants for the study for different portions of the study was done through the following procedure. It began with identifying the total number of students who registered for the course, which was 4,582. A list of 4,582 students with general information such as nationality, email address, and name was then obtained. Then, these students were sent an email of invitation for participation in the survey portion of the study for three times. The interval of each time between the three invitations was two weeks. There were 119 incorrect

emails with failure notice responses, so the total number of students actually receiving the survey was 4,463. After two and a half months of conducting the survey, statistics on Google Drive reported 1,290 respondents submitting their answers, accounting for 28.9% of the total students the researcher of the study sent the survey. The researcher then eliminated 75 participants who did not answer all the questions in the survey completely. All of these 75 participants, who had incomplete responses, were removed from the pool of the participants. Eventually, the study used responses from 1215 participants for the survey portion.

Among 1215 participants in the survey portion, there were 573 males (47.1%) and 642 females (52.9%). In regard to education levels, 625 participants (51.4%) had high school diplomas; 240 participants (19.8%) held Bachelor's degrees; 252 participants (20.7%) received some college with no degree; and 98 participants (8.1%) had Master's or Doctoral degree. Most of the participants were Vietnamese (58.1%), and others were Chinese (22.2%), Indonesian (5.8%), Taiwanese (5.4%), Singaporean (3.9%), Japanese (1.8%), South Korean (1.5%) and Malaysian (1.3%). Table 3 and Figure 2 below summarize the demographic information about the participants.

Table 3.

Demographics of the survey participants

Categories		Number of participants	Percentages
Gender	Male	573	47.1%
	Female	642	52.9%
Education	High School Diploma	625	51.4%
	Bachelor's Degree	240	19.8%
	Some College, No Degree	252	20.7%
	Master's or Doctoral Degree	98	8.1%
Total		1,215	100%

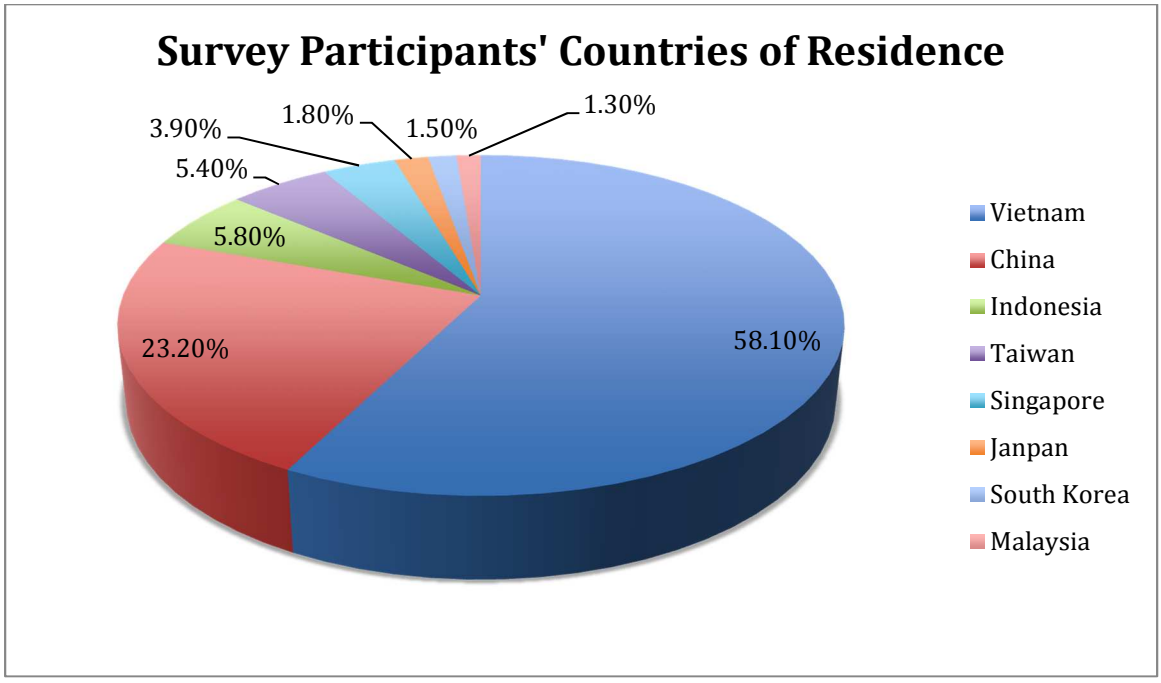

Figure 2. *Survey Participants' Countries of Residence*

3.2. Interview participants

Among 1,290 participants submitted their responses to the survey, 339 participants agreed to participate in Skype interviews. From this pool of 339 participants, the researcher selected 20 participants from a stratified random sample for the interviews —participants equally represented from different levels of grade performances in the course. Students' course grades in this pool of participants varied from A, B, C, to D. From each of the grade groups, 20 participants were randomly selected, including 5 whose final grade was D, 5 whose final grade was C, 5 whose final grade was B, and 5 whose final grade was A. Each of the interviews, which lasted approximately 25 minutes, was about participants' perception and opinions about peer assessment used in ESL/EFL Writing for Academic and Professional Success course.

4. Data Collection

As noted in section 1, given the range of research concerns, the study employed a mixed method approach combining paradigms of quantitative and qualitative because mixed methods would offer the researcher a better understanding of peer assessment in this context than if any one dataset has used alone. Each of the three research questions, then, was addressed through the collection of different types of data sets. In the subsections that follow, I will elaborate on each of these methods used to collect data and then turn to explaining the types of data analysis used.

4.1. Online Survey

The online survey included both Likert scale questions and open-ended questions to collect learners' opinions about online peer reviewed grading (see Appendix C for a copy of the online survey.) Responses from the survey were used to answer research question 1 "What perceptions do the students in a MOOC-based composition course have toward peer assessment of their own and their peers' drafts?" The survey was sent to 4,582 participants at the beginning, but the study used survey responses from 1,215 participants due to failure of email communication and incomplete responses.

The survey consisted of three parts. The first part was about demographic information such as gender, nationality, educational level, participation in previous online writing courses and peer assessment. The second part of the survey, as suggested by Dornyei (2007), were 15 questions employing a 4-point Likert scale with the following end points (1) Strongly Disagree and (4) Strongly Agree. A few of this study's Likert scale questions were drawn from Kaufman and Schunn (2010). Kaufman and Schunn (2010) investigated students' perceptions about online peer assessment for undergraduate writing across the disciplines, and they used survey items with measuring constructs like usefulness of own feedback, positive nature of own feedback, validity of own feedback, reliability of own feedback, usefulness of peers' feedback, positive nature of peers' feedback, validity of peers' feedback, reliability of peers' feedback, fairness of peers' feedback. Although Kaufman and Schunn (2010) looked into students' perceptions about online peer assessment in a non MOOC context, I found their survey items, especially questions about the nature of feedback, were relevant and could be employed for the investigation of students' perceptions about peer assessment in a composition MOOC context.

The following examples illustrate the type of questions I used from Kaufman and Schunn's survey, given their applicability to the MOOC setting and also the value of comparing the findings on perceptions of peer assessment within the present MOOC-based study and the findings within Kaufman and Schunn's non-MOOC study:

The feedback my peers gave me on my writing for this class was useful.

The feedback I gave my peers on their writing was too negative or critical.

In the present study, the Likert scale survey questions were particularly about the perceptions of (1) peer assessment training and peer assessment tool, (2) participation in peer assessment training, (3) usefulness of own and peers' feedback, (4) nature of own and peers' feedback, (5) reliability of peers' feedback, (6) fairness of peers' feedback, (7) the use of peer feedback in revision, and (8) own and peers' qualification. These aspects of peer assessment covered in the survey questions would help reveal the perceptions of peer assessment, by the participants in their roles both as students being assessed by their peers and students assessing and being prepared to assess their peers. While the participants' perceptions of peer assessment do not speak to quality of the participants' actual peer assessment, these perceptions (both the Likert questions translated into percentages and the open-ended questions coded for emergent themes) have the potential to shed light on issues related to peer assessment that might affect student readiness to benefit from it and thus potentially also its viability within the context of the particular E-Center course studied - ESL/EFL Writing for Academic and Professional. The following are two examples of the present study's survey questions employing the Likert scale:

I participated in peer assessment training and referred to the training instructions and rubric when commenting on and grading my peers' papers.

- *Strongly disagree*
- *Disagree*
- *Agree*
- *Strongly agree*

The guidelines for peer assessment were clearly outlined as to how to grade and to make comments.

- *Strongly disagree*
- *Disagree*
- *Agree*
- *Strongly agree*

The first example Likert scale question above asks the survey participants to weigh in on their participation in the peer assessment training offered for the MOOC course in this study and on the extent to which they did or did not apply the training instructions and rubric in their assessment of their peers' drafts. Likewise, the second example Likert question asks the participants to weigh in on whether or not they perceived the peer assessment guidelines as clearly presented. Each of the other Likert scale questions in the survey (see Appendix C), were similarly designed to draw from the survey participants their perception of each component of their involvement in the course as they experienced it.

The final part of the survey included four open-ended questions in order to gain more specific, more fleshed-out responses from each participant in addition to responses from closed

questions concerning the perception of peer assessment. The following is example of the open-ended question:

> *Please state in the space below your opinion regarding the following aspects of peer assessment (that is, peers providing the feedback and the grades) in the ESL/EFL Writing for Academic and Professional Success course offered by E-center.*
>
> - *Difficulties / challenges you experienced as commentator and grader for your peers*
> - *Peer assessment's usefulness in helping improve your writing performance.*

4.2. Skype Interviews

The Skype interviews were conducted with 20 participants selected from a stratified random sample - participants equally represented from different levels of grade performances in the course. Each of the interviews, which lasted approximately 25 minutes, was about participants' perception and opinions about peer assessment used in the ESL/EFL Writing for Academic and Professional Success course studied. Because students submitted their completed surveys anonymously, the survey responses submitted by individual interviewees could not be referenced and thus could not be used as a basis for the interview questions or used in combination with the interview responses to gain fuller insight into the individual interviewee's perceptions. Instead the interviewees' responses were intended to provide a sampling of how the survey participants might elaborate on their perceptions. In other words, responses from the interviews were used to provide further explanations and elaborations for the survey responses so as to flesh them out and confirm more about the students' perceptions.

4.3. Data on Peer & Instructor Grades for Submitted Assignments

The E-Center database used by E-Center to house course data included the peer and instructor grades for all the main assignments. These compiled quantitative grading data were subsequently analyzed to answer research question 2 - "Is there a significant difference between grades given by peers in response to student drafts and those given by the instructors?" Peer and instructor grades for all the graded drafts submitted provide data to add further statistical dimension to this study of the viability of peer grading in the E-Center class researched, as it compares to the grading of the E-Center instructors, and also to offer a more fine tuned analysis of how the grading for these two groups compare and contrast. In other words, the comparison of peers' grades and instructors' grades would reveal what portion of the peers' grades are similar to the instructors' grades and to what degree and what portion are not, and this would also help raise questions about issues found from other data sets (i.e. analysis of survey responses and analysis of comments). For example, a lack of difference between the grades given by individual peers and the instructors would be of particular interest should a percentage of responses by survey respondents and interviewees see grading by peers as unfair.

Based on E-center's data, the E-Center, as usual for quality check purposes as well as for the purpose of providing data for this study, had five instructors grade papers of 100 of the 219 students who both received the completion certificates and had their papers graded by three peers (as noted in section 1.4 above, a total of 334 students received completion certificates, whereas only 219 of the 334 also had three peers responding to their three papers). The number of papers the instructors were required to grade was 600 from 100 students (100 Draft 1 and 100 Final Draft, applied for three writing assignments). For research question 2, in order to have a balance

between the number of papers graded by students and the number of papers graded by the instructors, the study did not include all the grades given by peer assessors to all the essays for the course but rather limited the peer grades included in the comparison to the grades given to papers of the 100 students also graded by the instructors. Thus, the total number of graded essays by both peer assessors and instructors used to answer research question 2 would include 600 graded essays from 100 students (that is, 100 Draft 1 and 100 Final Draft, for the three writing assignments).

4.4. Peer and Instructor Comments on First Drafts for Assignment

Another portion of the database - the essay drafts with feedback (i.e. comments on essay drafts) provided by both student peers and instructors - were used to answer research question 3 "To what degree and in what way do comments/feedback given by peers in response to student drafts differ from those given by the instructors?" For research question 3, to narrow the focus of comment analysis due to time and labor constraints, the study looked into comments made by peer students and instructors from a selected sample of first drafts of the course's third writing assignment – the Argumentative Essay. The rationale for choosing the third writing assignment (versus the first or second) was that the third writing assignment might show the peer assessors at their most advanced in benefiting from their training and experience as writers and peer assessors. I limited the sample to 20 essay first drafts out of the 100 drafts total (for assignment #3) and selected 5 essays from a grouping of essays according to which of the four grades each essay was given (A, B, C, D). The 20 essay drafts included all 5 of the A papers, 5 of the 41 B papers, 5 of the 39 C papers, and 5 of the 15 D papers. To choose these five essay drafts out of

the total for each grade group, instead of random sampling, I used selective sampling, with the number of comments on each paper individually as the criteria for the selection - in other words, choosing from each grade group the five essays with the most comments. The rationale for using this approach is that it would provide me with the most comments to code and include in the results and analyses comparing the types and numbers of peer comments to the same information on the teachers' comments. In addition, selecting the essays with the most comments would give the maximum number of comments through which to assess the level of responses the peers were capable of making and provide evidence for answering the third research question.

5. Data Analysis

In this data analysis section, I will elaborate on the procedures through which the data collected were analyzed, with a separate section devoted to each analysis procedure.

5.1. Analysis of Survey Data

Survey responses by participants were automatically processed in Google Drive (a web portal where the researcher of this study created a survey and reported participants' responses). The data were exported into the Microsoft Excel program.

For survey responses of the fifteen Likert questions, the data were entered into the Statistical Package for the Social Sciences (SPSS 16.0) and analyzed through descriptive statistics in order to identify patterns of agreement and disagreement. Specifically, the percentages of responses for each of the four point scale questions were calculated. In addition,

for fifteen Likert scale questions in the survey, a Reliability Statistics Test was conducted to identify whether the items on the survey had an acceptable internal consistency. The fifteen items on the second part of the survey were tested using Cronbach's alpha (Table 4). The resulting alpha coefficient for the fifteen items was .806, suggesting that the items have relatively high internal consistency. According to George & Mallery (2009), a reliability coefficient of .70 or higher is considered "acceptable", which indicates good internal consistency reliability.

Table 4

Reliability Statistics Test

N of Items	Cronbach's Alpha	Standard Deviation	Mean
15	0.806790276	4.734402997	41.2844188

For responses of open-ended questions of the survey, I used the WordCounter in Microsoft Word to identify the highest frequency words in the responses and then manually conducted the qualitative analysis. The open-ended questions were analyzed through content analysis for common answers and categorized into themes, specifically through a coding process including open coding and selective coding. For open coding, I read through the data several times, created tentative labels based on participants' lexical uses (e.g. not feeling qualified; having limited competence/knowledge; not receiving sufficient training on peer assessment), and recorded examples of participants' words. For selective coding, I reread the data, figured out the core variable that emerges, and selectively coded any data that related to the identified core variable. A few examples of selective coding that emerged included lack of confidence and

experience, lack of qualifications, technical issues, anonymity of peer assessment, types of writing assignments, emotion, peer assessment tools, writing comments and assigning grades.

During the coding process, I was aware of the potential for overlap among categories. To deal with the overlaps, I strictly based the categories into which the data were coded on literal meanings of participants' lexical and syntactic uses. For example, when a participant wrote "*I have never graded anyone's writing before*", one might understand that to mean "*practice makes qualified*" and would consider placing this response in the category "*lack of qualification*". However, in this case, based on literal meaning of "*never graded….before*," I chose "*lack of experience*." In other cases where participants gave different ideas in one response, I separated each of the ideas and put them in relevant categories. For instance, one participant responded "*It was a very new experience to me and I felt ill-equipped to undertake the assessment*." For such response, I put the "*It was a very new experience to me"* in "lack of experience" and the "*I felt ill-equipped to undertake the assessment"* in the "*lack of qualification*." All of the coded analyses were then scored by frequencies and converted into percentages for discussion.
The following (see Table 5) is an example of open coding and selective coding for an open-ended question in the survey that asked the respondents to describe difficulties and challenges of peer assessment *"Difficulties/challenges you experienced as commentator and grader for your peers."*

Table 5

Examples of open codes and selective codes in survey's open-ended question analysis

Open codes	Examples of participants' words	Selective code
- not feeling qualified; - having limited competence/knowledge - not receiving sufficient training on peer assessment	I felt ill-equipped to undertake the assessment; I felt not qualified for the peer assessment; I am just a student, not a teacher; I don't think I have good knowledge to do the job; I should have participated in the training on peer assessment to be more qualified	Lack of qualification
- not having enough experience - new experience	I have never graded anyone's writing before; I don't do this very often, so I don't think I'm experienced enough in doing the job	Lack of experience
- Not knowing whose writing is assessed; - Not having a chance to talk about feedback. - Giving less quality feedback.	Grading was anonymous so there was no force to do the job effectively; I wish there was discussion after peer assessment; if I had known whom I gave feedback to, I would have given more constructive comments.	Anonymity of peer assessment

5.2. Analysis of Skype Interviews

The interviews were not designed to serve as triangulation of the other collected data collected from specific participants because the surveys were completed anonymously and only a small number of interviews were completed relative to the total number of survey respondents. Rather the interviews were designed to provide samples of the way in which the student participating might further define their perceptions beyond the more generally worded survey questions.

Responses from the Skype interviews were transcribed and typed in Microsoft Word. Similar to the procedure for coding the open-ended survey questions, the interview responses were also analyzed through content analysis for common answers and categorized into emergent themes using open codes and selective codes. The interview data were then reviewed in relationship to the responses for Likert scale questions as well as themes emerging from the open-ended survey responses, which altogether helped gather more in-depth insights on participants' attitudes, thoughts, and actions toward different aspects of peer assessment within the specific context of the present study.

To draw out within the interview greater information and understanding about the students' survey responses, questions that might overlap significantly with some Likert scale statements and open ended questions in the online survey, or that might be different enough in emphasis or tone to provide more insight on the implications like the following were asked: "*When you graded your peers' writing, did you strictly follow the rubric?*" (survey's Likert scale statements: *"I participated in peer assessment training and referred to the training instructions and rubric when commenting on and grading my peers' papers*"*;* "*For each of the assignments, I followed the guidelines and rubrics for commenting on peers' writing*"*)*; and "*What do you think about the comments you made on peers' writing? Were they constructive and comprehensive or negative and limited?*" (survey's Likert scale statements: "*The feedback I gave my peers on their writing for this class was useful*"; "*The feedback I gave my peers on their writing for this class was thorough and constructive."*)

Responses from the interview in review in the relationship with the survey's responses could provide a more fleshed out understanding of the student participants' perception of peer assessment and their involvement in it both as peer responders and student writers receiving

feedback on their writing. In reporting the results, more specifically, the analysis of survey data, I will also provide a sampling of the comments from the interviews in order to suggest some possible ways the significantly larger group of survey participants might further define some key terms used in the relevant survey questions.

5.3. Analysis of Peer Students' and Instructors' Grades

To answer research question two of the study - if there were any significant differences in grades given by peer students and those given by instructors and where the significant differences were, I conducted an analysis of a total of 600 essays from 100 students, including 100 Draft 1 and 100 Final Draft of the three writing assignments (these 600 essays were graded by both student peers and instructors). With the essay sample, I looked at the similarities and differences among grades given and examined whether the students tended to give the grades within a certain range that differs from the range of the grades given by the teachers. In addition, to examine statistical differences between student peers' grades and instructors,' I entered peer students' and instructors' grades into SPSS 16.0, and then carried out paired t-tests (also known as dependent t-tests) to calculate the mean of differences and to report whether this mean of differences is statically significant. More specifically, six paired t-tests were run on each of three writing assignments' grades given by the student peers and the instructors to compare the following: the mean of overall grades of the first drafts and that of the final drafts given by peers and the overall grades given by the instructors in three writing assignments (Application Essay, Summary and Evaluation Essay, and Argumentative Essay).

5.4. Analysis of Peer Students' and Instructors' Comments

To answer research question three on the differences between peer students' comments and instructors' comments, I decided to use a selected sample of 20 graded first drafts of the third writing assignment – the Argumentative Essays (comments included). I chose the third writing assignment (versus the first or second) because it might show the peer students at their most advanced in benefiting from their training and experience as writers and peer assessors.

. Using that sample, I worked to develop a coding scheme that reflects the connection between terms the peers used in their comments and key terms representing criteria the students' written products and peer comments were expected to meet as stated in the rubric and the peer training on productive grading practices. To forward the study's purpose of seeing if peer assessment is viable (as compared to instructors' assessment and as compared to the criteria set forth in the assignment rubric emphasized in the training), I, in considering how initially to approach the coding, examined the criteria and standards highlighted in the training on peer assessment and in the rubric. In the training, students were asked to focus on global issues including thesis, organization, idea development first and then local issues such as grammar, vocabulary, spelling, and punctuation. In the rubric (see Appendix B), grading criteria included focus, idea development, organization, grammar, punctuation, spelling, mechanics, and word usage.

Based on the standards presented in the peer assessment training and the rubric criteria, combined with preliminary observations of the types of comments in analyzed essays, the coding scheme broadly conceived in its initial stage took shape from these broad criteria of Idea Development, Organization, and Sentence Level, in sum called the WHAT category. Comments

that expressed concerns about thesis, focus, development of ideas and clarity of ideas were marked as "Idea Development." Comments that expressed concerns about the ordering of and relationship between parts of the essay (i.e. introduction, body, and conclusion), text structures, transitions between the ideas in writing, and logic of the ideas were marked as "Organization." Comments that expressed concerns about grammar, punctuation, spelling, mechanics, and vocabulary were marked as "Sentence Level." The three criteria of the WHAT then served as umbrella categories for sorting and finding patterns within the comments. I used these categories to further fine tune coding that could help provide results and tell ways in which and the extent to which the peer assessor students did or didn't register key factors from the rubric and their training in assessing their peers' essays.

The value of peer and instructor comments (i.e. the clarity with which they are communicated) depends not just on what aspects of the essay are commented on but also on HOW those aspects are communicated to the writer. Thus in the training, students were advised to write comments in complete, detailed sentences for clarity as well as to provide specific, comprehensive and constructive feedback on peers' writing by using different forms and functions of language such as explanations, corrections, suggestions, commands, questions, and so on. For example, in the live virtual classroom (LVC) on peer assessment training for Writing Assignment 3, one student's comment on a sample essay used in training said "awkward " to the sentence "*Friendship mark an important place in my life.*" The instructor asked students to explain what was awkward and to make that comment more specific so that the writer could understand what was really meant. Then, how specific comments could be communicated more clearly and specifically was discussed among students, followed by the instructor's model. In light of the importance of training students in how the WHAT of comments should be

communicated and the importance of this HOW to the clarity of the comments, the HOW of the comments from these 20 essays also needed coding.

To register the HOW choice for communicating the WHAT, I worked toward coding for the main HOW categories by looking for patterns in how the WHAT of the comments was communicated. Given the expectations from the training on how comments should be made, the pattern I initially arrived at for the HOW category included explanation, question, correction, suggestion, and command. Based on lexical and syntactical features, the HOW preliminarily fell into the following categories: *Explanation* indicates comments that describe or explain to students what they have written and/or that includes a "why" pointing the source of the problem. *Question* indicates comments that raise issues with the student's logic or claim or language use; *Correction* indicates comments that correct specific segments of the student's writing; *Suggestion* indicates comments that give suggestions for revising a paper; and *Command* indicates comments that instruct the student to do something.

What began to become clear, though, through this process was that this initial set of HOW categories led to coding according to the literal wording and word order used for a specific syntactic form used to express the assessor's comment rather than according to the specificity or specific tack taken in how the issue was communicated. For example, a command would involve use of the imperative or use of a modal like "*should*" or "*could*"; a suggestion would involve use of a qualifier such as "*may*" or "*might*". In addition, the literal wording or syntactic form that qualifies a comment as a command versus a suggestion could be significant for a researcher were the researcher's focus on the student writers' perceptions of comments. For instance, the peer comment wording "use a punctuation before 'and'" which is worded as a command (given the imperative verb form "use") as distinct from the wording that includes a qualifier in the peer

comment "*I would suggest you use a punctuation before 'and'*." Given that, wording comments as a suggestion might well be perceived differently by writers receiving comments than a comment worded as a command, such as in the possible greater impact of perceived commands (versus perceived suggestions) on writers' sense of their ability in revision to keep their own voice. However, this current study's focus is not on student perceptions of those comments but instead on how the peer comments compare to the instructors' in the specificity and manner in which they communicate writing issues to the writer. Looking over the sample, I found that comments can take very different form, such as one comment being a suggestion and another being a command, and still be offering the same type of feedback substantively. For instance, *"Use 'with' instead of 'to'"* (a command) and "*You might want to use a semicolon"* (a suggestion) offer a correction. In addition, I noticed that a few comments in the syntactic form of a question did not pose a question for real consideration but rather to assert a correction, as in such comments *"Should we use 'valuable' here? Is there another word for this same meaning?*"; *"Should it be a plural noun—memories?"*. These comments in question forms were all forwarding a solution – in other words – a correction. Therefore, I came back to the HOW category and adjusted the elements of the HOW, making it more relevant to the purpose of the study but still remaining connected to the main highlights from the participants' training and the rubric. The HOW I used, then, for the analysis included (1) Explanation (indicates comments that describe or explain to students an aspect of what they have written and/or that includes a "why" pointing the source of the problem); (2) Question (indicates comments that raise issues for the student to consider about the student's logic or claim or language use); and (3) Correction (indicates comments that correct specific segments of the student's writing, whether or not the correction is posed as a suggestion for revision or a command to the writer of something to

do/not to do or a question to forward a correction). Table 6 below shows some examples of the WHAT and the HOW of specific comments on student Assignment 3 essays:

Table 6.

Examples of the WHAT and the HOW

Comment	The "WHAT"	The "HOW"	Notes
"Check your dictionary for the differences between "reminds" and "reminders".	Sentence level	Correction	Comment made by the instructor.
"What is your thesis statement?"	Idea Development	Question	Comment by peer student
"The idea in this sentence seems to contradict what is said in the beginning of the paragraph."	Idea Development	Explanation	Comment by peer student
"she sat me down"	Sentence level	Correction	Comment by instructor. Student wrote "*she sit in down side*"
"You need to add a verb after 'want'. "	Sentence level	Correction	Comment by peer student

For the data analysis, each comment was given one point in accordance with the criteria in the HOW. When more than one category in the HOW might seem relevant, one point would be given to each of the corresponding criteria. For example, in the comment made by a peer student "*The flow of ideas is problematic. What position do you want to stick with?*" there was more than one category in the HOW. The first part of the comment "*The flow of ideas is problematic.*" was marked as *Explanation,* and the second part "*What position do you want to stick with?*" was considered *Question.* Although the second part of the comment ("*What position*

do you want to stick with?") implies an explanation of the problem cited in the first half of the comment (*The flow of ideas is problematic)*, the question seems the dominant coding given the lack of explicit explanation. More examples of the same manner of coding are provided in the following tables (Table 7, Table 8, and Table 9):

Table 7

Examples of the same manner of coding: WHAT (Idea Development) and HOW

Comments	WHAT	Idea Development		
	HOW	Explanation	Question	Correction
"The idea in this sentence seems to contradict what is said in the beginning of the paragraph."		1	0	0
"The flow of ideas is problematic. What position do you want to stick with?"		1	1	0

Table 8

Examples of the same manner of coding: WHAT (Organization) and HOW

Comments	WHAT	Organization		
	HOW	Explanation	Question	Correction
"What is your conclusion? You should add a concluding paragraph to wrap the ideas presented before."		1	1	1
"Does the transitional element use connect the various components of the essay and clearly express relationships between ideas? If here you use "First", the next paragraph is supposed the follow-up of this argument. However, it's not. The next should be the contrast. Suggested word: "On the one hand,""		1	1	1

Table 9

Examples of the same manner of coding: WHAT (Sentence Level) and HOW

Comments	WHAT	Sentence Level		
	HOW	Explanation	Question	Correction
"Use 'with' instead of 'to'."		0	0	1
"More than one idea here. You should revise this sentence to avoid the run-on problem."		1	0	1

Because each of student papers was assessed by three different peers, points for students' comments by each peer were counted, and were then compared with points for instructor's comments. In other words, the comments of each of the three peers responding to each paper were individually compared with the comments of the instructor who graded the same papers. This comparison approach is worthwhile because it is able to reveal variations in the WHAT and the HOW of the peers' comments as compared to the instructors'. Carrying out the comparison of peer versus instructor comments one peer at a time would help make the results be more precise, rather than comparing using the average of the comment characteristics of the three peers responding to each paper. In other words, the peer-by-peer comparison would make clear what portion of the WHAT and the HOW in each individual peer's comments are similar to or are different from the WHAT and the HOW of the instructor's comments.

Also, in coding sets of comments for patterns, I noted key words that connected with words and emphasis in the Writing Assignment rubric and looked for patterns that emerged from the rubric that peer students used. This would help bring out any patterns that revealed the peer assessors following (or not following) with reference to specifics within the rubric and instructions in the peer assessment training. For example, a comment by peer student "*What is*

your thesis statement?" had key words "*thesis statement*" precisely connected with words and emphasis in one of the criteria in the Writing Assignment # 3 rubric – the *Introduction* criterion, which stated "*Introductory section provides a strong opening, adequate context, and a clear thesis statement."* For comments that didn't have the precise words from the training or the rubric but used wording that in essence connects the comment to the rubric and training, I looked at the meaning of the content words in the comments and compared the wording with terms used in the rubric and training. For instance, the comment "*You should add a concluding paragraph to wrap the ideas presented before"* had words such as "*concluding paragraph*" and "*wrap the ideas presented before*" generally connected with but didn't directly use words "*concluding statement*" or "*summarizes the major points of the argument,*" which are stated in the *Conclusion* criterion of the rubric: "*Essay provides a concluding statement that accurately summarizes the major points of the argument and explains their significance fully*." In another example of a peer student's comment "*Are you contrasting yourself? The way you support your main points seems opposite,"* the meaning of content words "*contrasting yourself*" and "*support your main points*" indicated that the peer student addressed the issue of idea development – an emphasis in the peer assessment training and the Idea Development criterion in the rubric --"*Claims are well supported with logical reasoning and relevant evidence, demonstrating a complete understanding of the topic*."

In conclusion, this chapter discussed the methodology of the present study, including the participants, research methods, and data analysis. The number of the participants varied according to different portions of the study (1215 participants for the survey portion and 20 participants for the interview portion.) The study used a mixed methods approach combining paradigms of quantitative and qualitative research. Each of the research questions was addressed

through the collection of different types of data sets. The analysis of peer and instructor grades employed grades from 600 essays from 100 students (i.e. 100 Draft 1 and 100 Final Draft, for the three writing assignments), and the analysis of peer and instructor comments was conducted on 20 essay first drafts of the writing assignment # 3 – the Argumentative Essay. The following chapter will present the results of the data analysis in accordance with each of the research questions.

CHAPTER 4

RESULTS

As stated in Chapter 1, this study aimed to investigate aspects of peer assessment in a MOOC-based composition course, ESL/EFL Writing for Academic and Professional Success offered by E-Center, including students' perceptions of peer assessment, differences between grades given by students and those given by the instructors and differences between comments made by peer students and instructors. In chapter 3, the research methodology, I discussed the participants, the research methods, and data analysis. The study employed a mixed methods approach with the combination of quantitative and qualitative research through the collection of different data sets (i.e. online survey, Skype interviews, peer and instructor grades for submitted assignments, and peer and instructor comments on first drafts for assignment.) This chapter 4 describes the results of the data analyses, which includes three separate sections: (1) Results from the survey and interviews targeted to answer Research Question 1; 2) Results from the comparison of peer and instructor grading of student essay drafts for answering Research Question 2; and 3) Results from the coding of peer and instructor comments on the student essay drafts for the writing assignment three to answer Research Question 3.

1. Results for Research Question 1

Research Question 1. What perceptions do the students in a MOOC-based composition course have toward peer assessment of their own and their peers' drafts?

Research Question 1 was about the perceptions of students in a MOOC-based composition course toward online peer reviewed grading of their own and their peers' drafts. To

answer Research Question 1, I used an online survey with both Likert scale questions and open-ended questions to collect learners' opinions about online peer reviewed grading. The aspects of peer assessment covered in the Likert scale questions (15 questions – questions 5 through 19) were particularly about (1) peer assessment training and peer assessment tool, (2) participation in peer assessment training, (3) usefulness of own and peers' feedback, (4) nature of own and peers' feedback, (7) validity of peers' feedback, (8) fairness of peers' feedback, (9) the use of peer feedback in revision, and (10) own and peers' qualification. The open-ended questions (questions 20a through 20d) asked about their perceptions of (1) difficulties/challenges peer reviewers experienced as commentators and graders, (2) difficulties/challenges peer reviewers experienced as writer being commented on and graded, (3) peer assessment's usefulness in helping improve writing performance, and (4) similarities or differences among peers' comments and peers' grades. In reporting the analysis of survey data, I will also provide a sampling of the comments from the interviews with 20 of the survey participants in order to suggest some possible ways the significantly larger group of survey participants (1215) might further define some key terms used in the relevant Likert Scale questions. For instance, the interviewee comments on characteristics of peer feedback could serve to suggest specific terms and examples of how survey respondents might have perceived terms such as "too negative and too critical" used in Likert Scale question 13 in rating their degree of agreement or disagreement with survey respondents' rating of Likert Scale questions on the same issue. The results from the survey's open-ended questions (questions 20a – 20d), even if there is some overlap with the Likert Scale question findings, will be reported at the end of this section, given the virtue of having the results follow the order of the survey. As with the reporting of the Likert Scale question results, a sampling of comments from the 20 interviews will be included in the results for the open-ended survey questions (20a-d) I will

include a sampling of comments from the 20 interviews that help clarify or amplify or overlap with the open-ended survey question responses.

In the subsections that follow, I will present Likert scale question results and relevant sampling of interview comments, section 1.1, and results for open-ended questions and relevant sampling of interview comments, section 1.2.

1.1. Likert Scale Question Results and Relevant Sampling of Interview Comments

The bulk of this section provides the results for Likert Scale questions 5 through 19, processed and analyzed in Google Docs, are spelled out for the Likert Scale questions each within a separate table. Table 10 summarizes the participants' responses to questions on peer assessment training, peer assessment tool and instruction following. As shown in Table 10 below, 85.6% of the survey participants responded to Q5 that they took part in peer assessment training and referred to the training instructions and rubric when commenting on and grading their peers' papers. For responses to Q6, more participants agreed (64.5%) or strongly agreed (31.7%) that the guidelines for peer assessment were clearly outlined as to how to grade and to make comments, whereas a total of 3.8% of the participants showed their disagreement and strong disagreement. In terms of following the guidelines and rubrics during peer assessment (as referred to Q9 and Q10), more than 88% of the participants agreed or strongly agreed that for each of the assignments they followed the guidelines provided during training for commenting on and grading peers' writing and followed the rubrics for grading peers' writing, in comparison with approximately 11% of participants who either disagreed or strongly disagreed that they had followed the guidelines. In the interviews, a question that asked the participants about similar

aspects raised in the Likert scale questions was "*When you graded your peers' writing, did you strictly follow the rubric*?" Responses from the interviews showed that the majority of interviewees expressed that they paid attention to instructions, and followed the guidelines and the rubrics when assessing peers' work. For example, one of the interviewees responded to that question: *"I always followed the instructions and read the rubrics carefully because I think that the rubric is the guide for the assessment*" (I – 2). And another responded *"I followed the rubrics and used some terms from the rubrics when I wrote the comments. I tried to make my comments clear to my peers"* (I-13). The peers interviewed add to the Likert scale questions (e.g. Q10) details on what the interviewees might mean by "following the rubric" and why they follow. In the first quoted comment (I-2), the peer interviewed indicated the reason he followed the rubric – "*the rubric is the guide for the assessment*". In the second quoted comment (I-13), the interviewee not only explained the reason to follow the rubric –"*to make my comments clear to my peers*", but also described how to follow the rubric – that is, using terms in the rubrics when commenting.

Table 10

Survey participants' opinions on peer assessment training, peer assessment tool and instruction following

	Strongly Disagree	Disagree	Agree	Strongly Agree
I participated in peer assessment training and referred to the training instructions and rubric when commenting on and grading my peers' papers (Q5)	5.1%	9.3%	31.4%	54.2%
The guidelines for peer assessment were clearly outlined as to how to grade and to make comments (Q6)	1.2%	2.6%	64.5%	31.7%
For each of the assignments, I followed the guidelines provided during training for commenting on and grading peers' writing (Q9)	2.3%	8.1%	58.1%	31.4%
For each of the assignments, I followed the rubric for grading peers' writing (Q10).	2.2%	8.3%	33.7%	55.8%

In regard to survey participants' opinions on the feedback they gave to peers' writing (survey questions 7, 8, and 11), 62.8% of the participants believed the feedback they gave their peers on peers' writing was useful whereas 35.9 % of the participants disagreed and 1.3% strongly disagreed (Q7). 86% of the participants showed their disagreement and strong disagreement when asked if the feedback they gave their peers on peers' writing was too negative or critical (Q8). In response to another question on the quality of the feedback the respondents gave to their peers (Q11), the majority of the participants indicated their belief that

their feedback on peers' writing was thorough and constructive while 23.3 % of the participants had the opposite opinions. Table 11 below reveals the participants' opinions on own feedback given to peers' writing.

Table 11

Survey participants' opinions on own feedback given to peers' writing

	Strongly disagree	Disagree	Agree	Strongly Agree
The feedback I gave my peers on their writing for this class was useful (Q7).	1.3%	35.9%	59.3%%	3.5%
The feedback I gave my peers on their writing for this class was too negative or critical (Q8)	45.3%%	40.7%	8.1%	5.9%
The feedback I gave my peers on their writing for this class was thorough and constructive (Q 11)	4.7%	18.6%	70.9%	5.8%

Survey responses as to the feedback the particpants gave to peers' writing were echoed in the interviews. When asked in the interviews "*What do you think about the comments you made on peers' writing? Were they constructive and comprehensive or negative and limited?*" two of 20 respondents believed that their comments were constructive and comprehensive. One explained,

> "*I always avoided making too general and vague comments. My comments had explanations and suggestions for improvement. I tried to show my peers "here's what's wrong and here's how to fix it*." (I-2)

The other responded:

> "*I don't expect to receive vague and limited feedback from my peers, so I tried my best to make my comments clear and useful so that peers can find my comments useful for their revision. In all of my comments, I pointed what is wrong, what does not make sense. If I knew how to correct the errors, I would suggest the correction."* *(I – 15)*

In addition, the majority of interview respondents replied that they tried to give constructive comments to peers and make their comments comprehenisve. They also noted that some of their comments were constructive and comprehensive and some were general and negative, depending on the errors. One of the respondents explained how she commented on peers' papers:

> *"When I noticed errors like grammar and punctuations, I often suggest the corrections. For points that I think there might be problematic but I don't know how to correct them, I just highlighted the points and puts questions there. For question marks, my peers might think my comments were limited and not clear, but I think that can make my peers think about the points I highlighted."* (I – 1)

Through the interview comments quoted above, the peers interviewed adds to the Likert Scale questions (Q7, 8, and 11) details on what the interviewees might mean by the comments being "*useful*" or "*thorough*" and "*constructive*." The interviewees (I-2 and I-15) provided a definittion of what they might have interpreted to be "useful," "constructive," and "comprehensive": that is, making comments clear by marking the errors with possible explanations and making suggestions for corrections. Similar to I-2 and I-15, the interviewee I-1 also commented on usefulness as including her as assessor indicating corrections, though, unlike the previous interviewees (I-2 and I-15), this interviewee limited the corrections to a specific type of error, grammar and punctuation: *"When I noticed errors like grammar and punctuations, I often suggest the corrections…"* and went on to explain the rationale for commenting in cases in which she believed there's a problem but isn't sure how to explain it: *"…For question marks, my*

peers might think my comments were limited and not clear, but I think that can make my peers think about the points I highlighted."

For the survey participants' opinions on peers' feedback given to own writing, as shown in Table 12, only 12.7% of the participants perceived that the feedback they got from peers was negative and critical while the rest of 87.3% had the opposite opinion - i.e. the feedback from peers were not crtical or negative (survey question # 13). Table 12 also shows that the majority of the participants (80%) believed that the feedback they got from peers connected very clearly to the standards set forth in the assignment rubrics for the course assignment (survey question #12).

Table 12

Survey participants's opinions on peers' feedback given to own writing

	Strongly disagree	Disagree	Agree	Strongly Agree
The feedback my peers gave on my writing was negative and critical (Q13)	45.4%	41.9%	9.3%	3.4%
The feedback provided to me by my peers connected very clearly to the standards set forth in the assignment rubrics for the course assignment (Q12)	1.4%	18.6%	77.9%	2.1%

Responses from interviews regarding what they thought about comments made by peers added some specific components to what was found in the survey responses. When asked in the interviews "*What do you think about the comments that your peers made on your writing? Were they constructive and comprehensive or negative and limited?"* (this interview question is relevant to survey question #13), all of the interviewees responded that comments that peers made on their writing were varied – they received both comprehensive, constructive comments and negative, general ones. The following quotes are some examples that illustrate what the interviewees indicate:

> "*If you ask me whether peers gave me constructive, comprehensive, negative, or limited, I would say yes to all of these. By this I mean some comments were very comprehensive and constructive, that is explaining the problems in details and suggesting corrections. However, some other comments were so limited and negative. By limited I mean too short comments like "wrong grammar" or "revise the thesis". Such comments can not tell C. And by negative I mean direct criticisms that pointing at me not my paper for instance I got a comment that said "Your grammar is problematic." I would be okay if my peer said "The paper is grammatically problematic."*
>
> (I-18)

> *"Well…some comments were very detailed and I like it when my peers provided suggestions and corrections because these would surely help me. Some other comments were very general – simply their personal opinions like 'I like it' or 'well said'."*
>
> (I – 10)

> *"Most of the comments I got from peers were comprehensible. I'm glad that my peers specified the problems and make suggestions for correction. However, some comments were really general, and I don't know what my peers really meant."* (I – 7)

From the comments quoted above, the interviewees provided more details to Likert scale question (i.e. Q13) on what they thought about peers' comments. For I-18, "comprehensive" and "constructive" means giving specific explanations and suggesting corrections while "limited" means "short" (not telling "*what I should do in revision*") and "negative" means "direct

criticism." Similar to I-18, I-10 and I-7 indicated "*comprehensive"* and "*constructive*" comments as "*being detailed*", "*specifying the problems*", and "*making suggestions and corrections.*" In addition, both I-7 and I-10 explained what they thought of "general" comments: that is, "*personal opinions like 'I like it' or 'well said'"* (I-10) and not knowing "*what my peers really meant"* (I-7). When further asked to give an example of peers' comment, I – 7 said of the peer's comment, *"…. You can look at my paper – the assignment two – the summary evaluation essay…..The comment is "thesis" with three question marks."* The following excerpt is the follow-up questions raised for I-7:

Interviewer: *How do you understand that comment?*
Interviewee: *I don't know. I have the thesis in the introduction. Maybe the thesis is not clear enough.*
Interviewer: *What do you think about that comment?*
Interviewee: *I think the comment should be more specific.*

These follow-up answers by I-7 show that the interviewee was by implication defining an example of an unclear comment and suggesting too (in "I have the thesis in the introduction") the basis for not being able to follow what the peer comment might mean and making I-7 wonder what else the commenter might be pointing to with the comment: *Thesis???*.

In regard to the usefulness of peer's feedback (see Table 13 below), the participants expressed positive attitudes. Almost 96% of the survey participants felt that the feedback from peers helped improve their writing (survey question # 16) and approximately 89% found the feedback from peers helpful and used peers' feedback when they revised their writing (survey question #17).

Table 13

Survey participants' opinions on usefulness of peers' feedback

	Strongly disagree	Disagree	Agree	Strongly Agree
Feedback from peers helped me improve my writing (Q16).	0%	3.4%	34.1%%	62.5%
I found feedback from my peers helpful, and I used their feedback when revising my writing for this class (Q17).	1.6%	9.3%	51.2%	37.9%

In terms of survey participants' perception of the validity and fairness of peers' feedback, Table 14 reveals that the survey participants expressed the highest disagreement (65.8%) and highest agreement (28.4%) with the statement, "The feedback I got from one peer was similar to the feedback I got from other peers on the same paper" (survey question #14). However, most of the participants agreed (54.4%) and strongly agreed (7.15%) with the statement, "Peers gave me fair grades on my writing" whereas 38.5% expressed their disagreement (survey question 15).

Table 14

Survey participants' opinions on the validity and fairness of peers' feedback

	Strongly disagree	Disagree	Agree	Strongly Agree
The feedback I got from one peer was similar to the feedback I got from other peers on the same paper (Q14).	3.5%	65.8%	28.4%	2.3%
Peers gave me fair grades on my writing (Q15).	6.8%	31.7%	54.4%	7.1%

In terms of fairness of peers' grades, interview participants expressed different points of view, most of the interview respondents believed that peers gave fair grades while others stated that peers' grades were too higher or too lower than expected. When asked in the interview "Do you think your peers gave fair grades?", a respondent said, *"I do think so. I did evaluate my writing on my own, and I got the same grade to what peers gave me."* However, another replied to this same interview question, *"I remember I got a C for my first assignment. I think I should have an A for that paper."* The interview comments add details to what the rating of the related Likert scale question in the way the student participants might define what fair grades are - if the peer assessor gave the same grade as the interviewee gave to the interviewee's writing (which is a different measure than measuring the fairness in relation to other possible factors, such as what grades the interviewee writer might have heard other peer writers might have received).

For survey participants' opinions on their own and peers' qualification to give feedback and grades, only 39.5% of the participants felt that they were qualified to give feedback and grades on peers' writing while 61.5% expressed the opposite opinions (survey question # 18).

Despite expressing little expectation in their own qualification, 66.3% of the participants expressed that peers were qualified in giving feedback and grades (survey question # 19). Figure 3 below shows survey participants' opinions on own and peers' qualification to give feedback and grades (survey question #18 and 19).

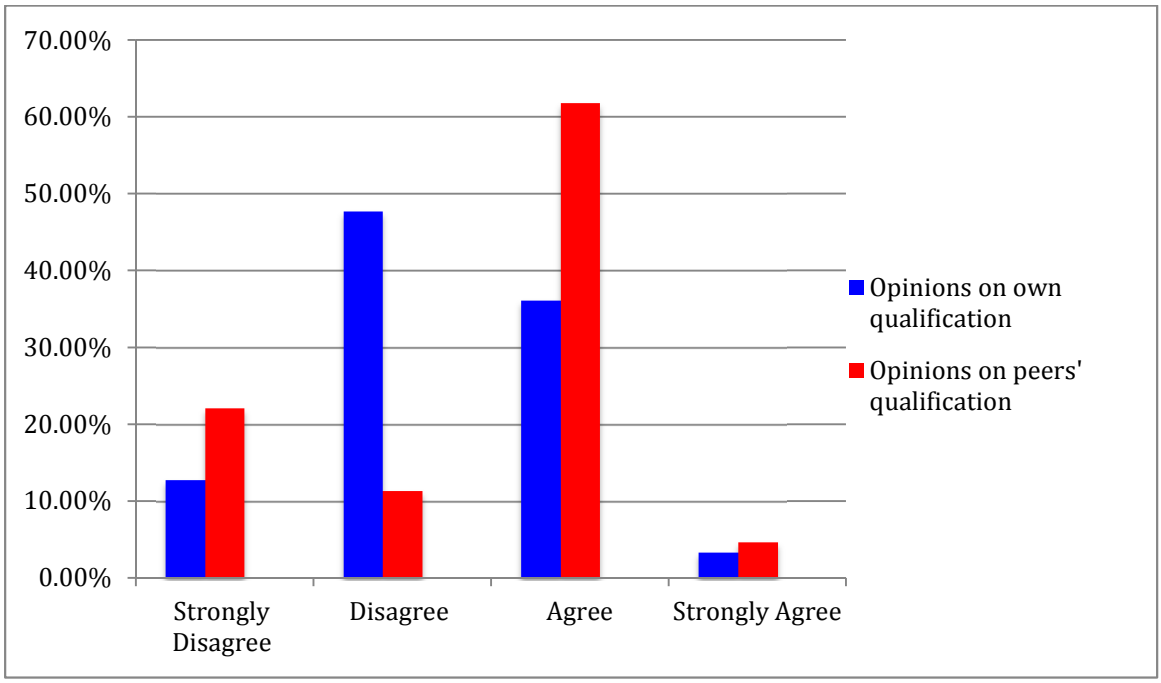

Figure 3. *Survey participants' opinions on own and peers' qualification to give feedback and grades.*

Regarding this same aspect mentioned in the survey, interviews included questions as to interviewees' own qualifications of giving feedback on peer writing. When asked in the interviews if they were well-prepared for the peer assessment, interviewees provided varied answers: some interviewees responded that they were well-prepared while others responded they were not. It was also revealed in the interviews that the feelings of some interview participants' about qualifications for peer assessment might have been changed or influenced by the training and instructions on peer assessment, whereas changes and influence could not be observed from

the survey responses (survey question #18). The following interview responses are examples of the changes in participants' feeling during the course of the class:

> *"To tell the truth, I have never given feedback to peers before, neither comments nor grades, and I did not really feel confident in the peer assessment activity at all. However, after I watched posted videos on how to do the peer assessment, followed instructions on peer assessment and tried grading and commenting on two writing samples, I felt just a little bit more confident in doing the job, but I don't think I was well prepared."* (I – 4)

> *"I think I was well-prepared after I followed the instructions about doing peer assessment and participated in the training session. Although I did not assess other people's writing before, I think I was able to do the job as long as I followed the rubric and the instructions."* (I – 2)

1.2. Results for Open-Ended Questions 20a-d and Relevant Sampling of Interview Comments

As will become clear in the results from the open-ended survey questions 20a-d presented below, survey participants indicated a variety of views on the post-training confidence and ability of peers commenting on and grading other student participants. The responses to the open-ended survey questions by the 1215 survey participants provided a great deal of data on a number of peer review issues. The open-ended questions asked about (a) difficulties/challenges peer reviewers experienced as commentators and graders, (b) difficulties/challenges peer reviewers experienced as a writer being commented on and graded, (c) peer assessment's usefulness in helping improve writing performance, and (d) similarities or differences among peers' comments and peers' grades. Major patterns and trends of responses to the open-ended questions were identified, labeled, and then categorized into themes through a coding process including open coding and selective coding (see more in Chapter 3, section 4).

Regarding responses to the survey's open-ended questions 20a-b about the participants' perception of difficulties/challenges peer reviewers experienced as commentators and graders, Figure 4 below shows prevalent themes.

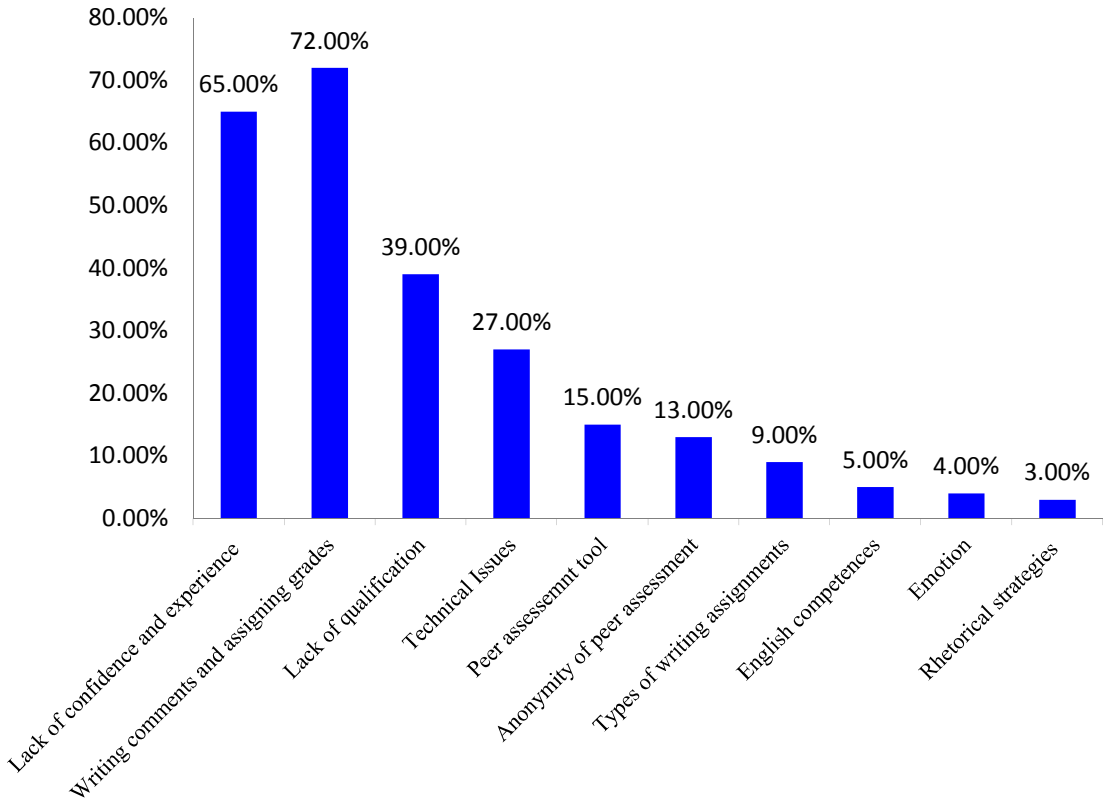

Figure 4. *Participants' perception of difficulties/challenges peer reviewers experienced as commentators and graders*

As can be seen in Figure 4 above, about 65% of the respondents spoke about their lack of confidence and lack of experience. More than 72% of the respondents expressed that they had difficulty writing comments and assigning grades, and approximately 39% of the respondents felt that they were not qualified enough to do the job. In addition, prevalent themes raised included technical issues e.g. assess to the papers, computer skills (27%), peer assessment tools i.e. the rubrics (15%), the anonymity of peer assessment (13%), the types of writing assignments

(9%), and other difficulties and challenges like one's English competences (5%), emotional factors (4%), different rhetorical strategies (3%) and so forth. The following comments from respondents are some examples that illustrate these categories:

Lack of Confidence and Lack of Experience:

> *Although there was some training on peer assessment, I never felt confident when making comments on other people's writing. This was my first time ever to do this so I was not experienced at all.*
>
> (Female #45, China)

> *Grading peers' writing was a very new experience to me. At the beginning I did not know what to do. Later I was better aware of the activity but I did not feel confident in what I did. If I had some experience in grading, I would feel more confident.*
>
> (Male #86, Vietnam)

Writing Comments and Assigning Grades

> *Writing comments was the most challenging to me. I took a lot of time to practice from the training how to make comments. Many times I got stuck at not being able to think of what comments should be made. I found it very hard to write relevant and constructive comments. Plus, translating the comments into a grade was super hard, even when I followed the rubric.*
>
> (Female #69, Japan)

Lack of Qualifications

> *I just felt I was not qualified to make comments and to give grades to my peers' papers. I was afraid my comments were not comprehensive, constructive, and helpful, and the grades would not reflect the true quality of peers' papers. I was not professionally trained to assess others' writing. I received the training from this course, but I think I need more than that to be more qualified.*
>
> (Female #220, China)

Technical Issues

> *I was challenged by assessing the papers online. I'm not good at doing things with computer and the internet, so it really took me a lot of time and effort to follow guided steps in order to be able to access the assessment page. Moreover, I almost surrendered when I had to insert my comments on margin in Microsoft Word and upload the graded papers. I appreciate the detailed instructions of doing these stuff, but these really challenged me.*
>
> (Male #472, South Korea)

Peer Assessment Tools

> *It is difficult for me to understand the rubric thoroughly. I read the rubric, and also watched the videos explaining the rubric and important points I should pay attention when doing the peer assessment. However, I did not understand some words and phrases in the rubric, maybe because my English is not good enough. Therefore, it's difficult for me to match the graded papers with the rubric criteria.*
>
> (Female #325, Vietnam)

Anonymity of Peer Assessment

> *I did not know the authors whose papers I graded and vice versa; therefore, it would be difficult to gauge how the authors would react when they got the comments and grades from me. If I knew the authors, it would be better for me to tell the authors about their strengths and weaknesses.*
>
> (Female #16, Singapore)

Types of Writing Assignments

> *I find it hard to grade a writing assignment when I am not familiar with that type of writing. In this course, the summary and evaluation assignment was something I never knew of before. I graded such papers just because I was required to do so. I don't think I did a good job when assessing something I didn't know of.*
>
> (Female #198, Taiwan)

English Competences

> *One of the challenges was that my English is not good enough. It's my low proficiency in English that limited my ability in evaluating others' work. I hardly knew if peers had good word choices or used correct grammar, hence I could neither point out peers' errors nor made suggestions.*
>
> (Male #315, Indonesia)

Emotion

> *Personal emotion affected my grading. I don't want to hurt my peers' feeling and fail them, so I tended to make positive comments and praises. I acted generous when I gave higher grades the papers actually deserved.*
>
> (Male #117, Singapore)

Different Rhetorical Strategies

> *Sometimes I had difficulty following the ideas presented in peers' papers. The way they organized evidences and provided necessary information seemed different from the way I did, so it was hard to grade such papers.*
>
> (Female #613, South Korea)

In regard to the perception of difficulties and challenges participants experienced as writers being commented on and graded (see Figure 5 below), among those who were surveyed, approximately 68% said that they had difficulty interpreting and understanding the comments given by peers. Some other most prevalent difficulties/challenges raised in the survey responses included perceived comment and grade contradiction (34%), fairness (23%), peers' qualification (19%), the anonymity of peer assessment (15%) and others (10%) like trust, conflicting comments, cultures.

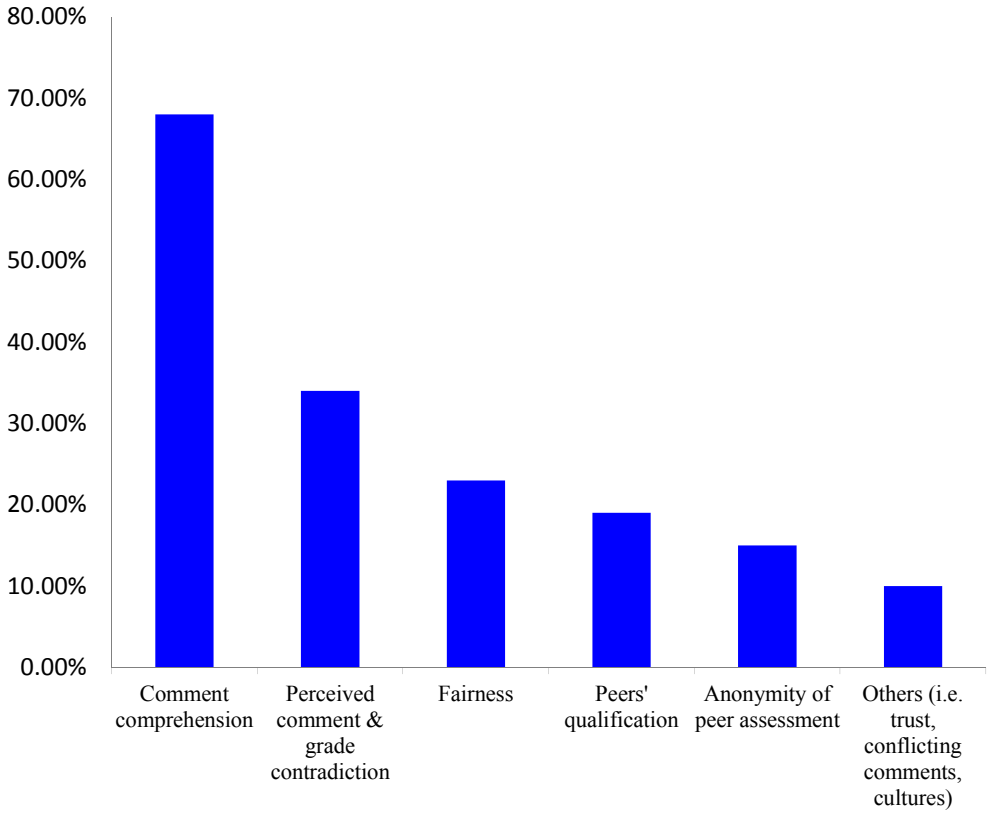

Figure 5. *Participants' perception of difficulties/challenges peer reviewers experienced as a writer being commented on and graded*

Comment Comprehension: Out of 68% of survey respondents spoke in their responses to open-ended question 20b about their difficulties in understanding comments given by peers, more than 80% of the respondents claimed that sometimes peers' comments were too vague and

general, which made it hard to comprehend what peer reviewers really meant; around 20% found peers' comments were not relevant to the writing assignments' requirements. The following are some examples from the respondents.

> *"For some peers, they gave me very short comments like 'thesis', 'grammar', 'vocabulary'. I completely understand the meaning of such comments but these were too general. I did not know what about 'thesis', 'grammar', or 'vocabulary' my peers really meant."*
>
> *(Female #75, Vietnam)*

> *"I got a few comments which were very vague and not specific enough. I also got some comments, which were not relevant to the assignment at all. To me, these kinds of comments were hard to understand and to follow."*
>
> (Female #115, Indonesia)

Perceived Comment & Grade Contradiction: Around 34% of the respondents said that the comments and grades given by peers did not reinforce each other and they were even opposite to each other (i.e. comments does not reflect the grades). The respondents explained that comments were sometimes positive but the grades were significantly low or vice versa. A participant stated:

> *"In one of my papers, I got most of positive comments about organization, idea development, vocabulary, etc. Nevertheless, my final grade was very low. There should have been a match between comments and grades."*
>
> (Male 725, Singapore)

Fairness: Approximately 23% of the respondents felt that the comments especially the grades given by peers were not fair enough. Among these respondents, many explained that peers tended to over-score the writing and that the grades they got from peers were too lower or higher than the grades that they themselves thought they deserved. Below is one of the participants' statements:

> *"Of course, I like it when I get high grades for all of my papers. I always graded my own papers using the given rubric, and I noticed that more than one of my peers scored much higher than what I though I would get. With self-assessment, I expected a C, but I got an A then. Well… that peer cheered me up, but to be honest I think that was not fair enough. Plus, the high grade might make me not motivated enough to revise my writing."*
>
> (Female #537, Japan)

Peers' qualification: Almost 19% of the respondents surveyed registered in their answers that peers were not qualified enough to make comments on and assign grades to their papers. They explained that peers' comments were sometimes vague, general, negative, and even rude, and that many comments were not comprehensive, constructive, and relevant to the writing assignments. In addition, the majority of the respondents (among the 19%) claimed that the grades given by peers did not reflect the actual quality of the papers. One of the respondents explained his difficulties /challenges that he had due to peers' qualification as follows:

> *"I had doubts about some peers' ability to do the assessment. They seemed not experienced in assessing others' writing. Their comments were not clear and constructive enough. I had difficulty interpreting their comments and I did not find those comments useful for my revision."*
>
> (Male #821, Malaysia)

Anonymity of peer assessment: Approximately 15% of the respondents spoke about the anonymity of peer assessment, explaining that they did not have chances to discuss feedback with peers or to ask peers for clarifications because they did not know who assessed their papers. A respondent stated:

> *"I think the challenge was the blind peer assessment when I didn't know who graded my papers and the graders didn't know whose papers they were grading. Not all of the feedback I got from peers were clear to me, so I would love to discuss with the graders about their feedback to clarify certain points. However, I didn't have any chance."*

(Male #67, South Korea)

Other difficulties/challenges: In addition to difficulties and challenges stated above, the participants (10%) also talked about others like conflicting comments trust in peers, different cultures, etc. Some participants expressed their doubts about the accuracy of peers' comments. Some had difficulties deciding between conflicting comments from different peers. Others talked about differences in feedback because of peers coming from different cultures. Examples illustrating these points are below:

> *"When reading marginal comments that peers wrote on my papers, I was not so sure if my peers were right and if I should do what my peers suggested for correction. That's my challenge."*
>
> (Female #2, Vietnam)
>
> *"I got mixed comments on the same point. It was really hard for me to choose which advice to follow."*
>
> (Female #105, Singapore)
>
> *"It is confusing and challenging when I got two different feedbacks on the same paper. One peer gave only positive feedback and scored high, and the other peer had more critical feedback and scored lower. I know that students in this class were from many different countries, so in this case I guess cultures made such differences."*
>
> (Female #943, Japan)

Regarding this same issue, when asked in the interviews *"What difficulties/challenges did you encounter when you assessed your peers' writing – both grading and commenting on peers' papers?"* the interviewees gave answers that were also reflected in the survey responses. Among these challenges/difficulties were lack of experience and confidence, writing comments, and understanding comments. One of the interviewees replied, "*I would say the most difficult part is how to write comments I mean how to make my comments understandable and helpful for my peers*" (I-9). Another interviewee responded: "*I sometimes found it hard to understand comments*

peers gave to me. I wish those comments would have been more detailed so that I could know what peers meant to say" (I-14).

For responses to the other open-ended questions in the survey about peer assessment's usefulness in helping improve writing performance, the first half of the responses were tallied into three categories "Yes", "No", and "Somewhat"; the second half of the responses - the reasons and explanations that the participants provided were categorized into themes through a coding process including open coding and selective coding. The following chart presents the percentages of the first half of the responses.

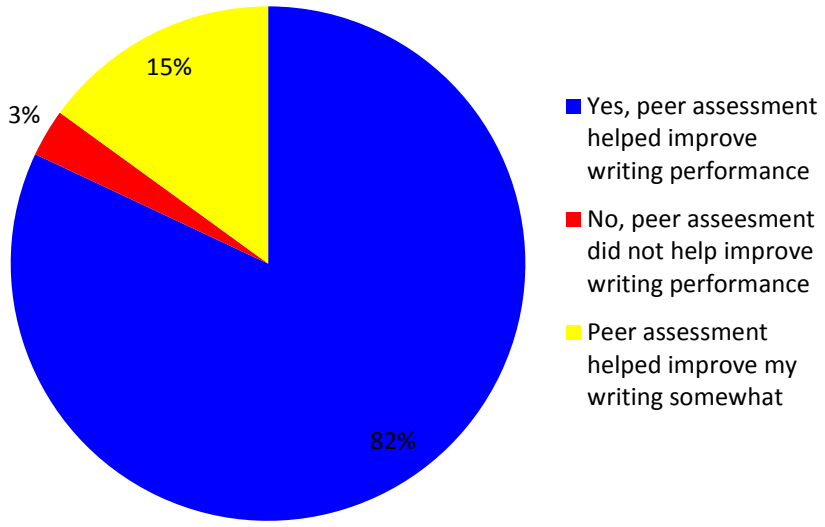

Figure 6. *Participants' opinions on whether peer assessment helps improve writing performance*

As shown in Figure 6 above, the majority of the participants (82%) responded that peer assessment was useful in helping improve their writing performance whereas 3% of the participants opposed the usefulness of peer assessment. Fifteen percent of the participants

revealed both positive and negative opinions about the usefulness of peer assessment in helping improve writing performance. The answer to the first half of this open-ended question is quite consistent with the responses of the participants to the Likert scale question. Specifically, in the Likert scale responses, almost 96% of the participants felt that the feedback from peers helped improve their writing while around 4% of the partiticpant felt the opposite, and approximately 89% found the feedback from peers helpful and used peers' feedback when they revised their writing. Overall, the results show high agreement on the usefullness of peer assessment in both Likert scale responses and open-ended responses.

Regarding the second half of the responses to the open-ended question about the usefulness of peer assessment in helping improve writing performance, approximately 45% of the participants only responded whether or not peer assessment was useful and did not give further explanations (called group A). In this group A, around 80% of the participants stated that peer assessment helped improve their writing; about 5% opposed the usefulness of peer assessment; 14% expressed both agreement and disagreement. The following are examples of the participants' responses:

> *"Peer assessment is definitely helpful."*
>
> (Female #14, Indonesia)

> *"I think peer assessment does not contribute to the improvement of my writing performance."*
>
> (Male #162, China)

> "*Not always helpful. Sometimes peer assessment helps, but sometimes it does not*."
>
> (Male #217, China)

There were about 55% of the participants who both responded whether or not peer assessment was useful and gave further explanations (called group B). Similar to group A, most of the

participants in group B expressed the agreement on the usefulness of peer assessment, accounting for approximately 78%; 6% expressed the disagreement; and 16% expressed both agreement and disagreement.

As far as the reasons and explanations for the "Agreement" are concerned, the majority of the survey participants in group B (approximately 65%) explained that peer assessment helped them revise drafts better and score higher on revised drafts because they received advice from peers. Many participants (about 40%) stated that peer assessment helped them understand what constitutes 'good work" through their seeing different approaches other peers took in responding to a writing assignment as well as knowing the assessment criteria. Another common reason that the participants (about 25%) provided was peer assessment helped promote a deep approach of learning because they had chances for better understanding of the writing assignments and the assessment criteria. Below are several participants' statements that express how peer assessment helps improve their writing:

> *"My peers pointed out my errors and also gave suggestions on how to correct the errors. With their feedback, I got higher grades for my final drafts. My grammar looked better and my essay organization looked better. I found peer assessment very useful"*
>
> (Male #67, Indonesia)

> *"Peer assessment helps a lot. I read more carefully with attention to the details of my writing, and by taking into account the peers' responses, I could strengthen my writing."*
>
> (Male #420, China)

> *"Peer assessment is very helpful. When I went through the rubric criteria for peer assessment, I was more aware of what I should do with my papers to have good writing products. Especially, after reading my peers' papers, I had better understanding of what is good and what needs improving. I knew what I would do for my revised papers."*
>
> (Male #912, South Korea)

Regarding the disagreement in responses to the open-ended question 20c on the usefulness of peer assessment in improving writing performance, the majority of participants shared the one same reason – poor quality of the comments. They explained that feedback from peers was not clear and constructive enough so they did not use peer feedback in revision. Some others argued that they did not learn anything from other peers when they assessed peers' papers. One participant said:

> *"I found peer assessment not helpful. I did not learn anything from my peers because their papers were badly written. Besides, I could not use any of the comments peers gave me for the revision because their comments did not make sense to me"*
>
> (Male #811, China)

Overall, in the open-ended question section (20c), survey's participants expressed more agreement than disagreement that peer assessment was useful in improving writing performance. This more dominant positive response to question 20c was echoed in the interviews when the interviewees were asked a similar question "*Do you think your writing improved over the period of the ESL/EFL Writing for Academic and Professional Success course? If so, do you see peer assessment—your peers' comments and grading—as having contributed to that improvement? And if so, in what ways did the peers' comments and grades help improve your writing?"* According to interview responses, all of the interviewees answered that peer assessment helped improve their writing in different ways – improvement thanks to peers' comment, improvement thanks to learning from peers as a grader, and improvement thanks to learning from the assignment guidelines and the rubrics. Many of the interviewees expressed that after peer assessment, their second drafts were generally better than the first drafts in terms of organization, idea development, and grammar. In addition, some of the interviewees believed that comments and grades given by peers made them read and write more carefully. Below are two examples of

interviewees' responses about the usefulness of peer assessment in improving writing performance.

> *"So far yes [peer assessment helped improved my writing]. I got specific comments from my peers, and also corrections and suggestions. These helped a lot when I revised my papers, especially grammatical stuff."* (I – 8)

> *"I think my writing improved. I'm more aware of issues like thesis, organization, idea development. When I graded others' writing, I had a chance to read the rubric criteria. Then I checked those in my own writing."* (I -20)

Among those responding that peer assessment was helpful, many interviewees added that in some instances peer assessment did not help improve their writing, explaining that general, vague, or even brutal comments given by peers made them more confused about what was going on with their writing. The following quotation helps illustrates this point:

> *"I think peer assessment really helped. I revised my papers according to what my peers suggested if I think peers are right. My second drafts were better than my first drafts. However, some comments especially vague ones did not help at all" (I – 3)*

For responses to the open-ended questions in the survey about participants' perception of similarities or differences among peers' comments and peers' grades, the responses were measured into three categories "Yes", "No", and "Somewhat". According to Figure 7 below, approximately 19% of the survey participants responded that there were similarities among peers' comments and peers' grades whereas around 73% of the participants stated the differences. Eight percent of the participants expressed that peers' comments and peers' grades were both similar and different.

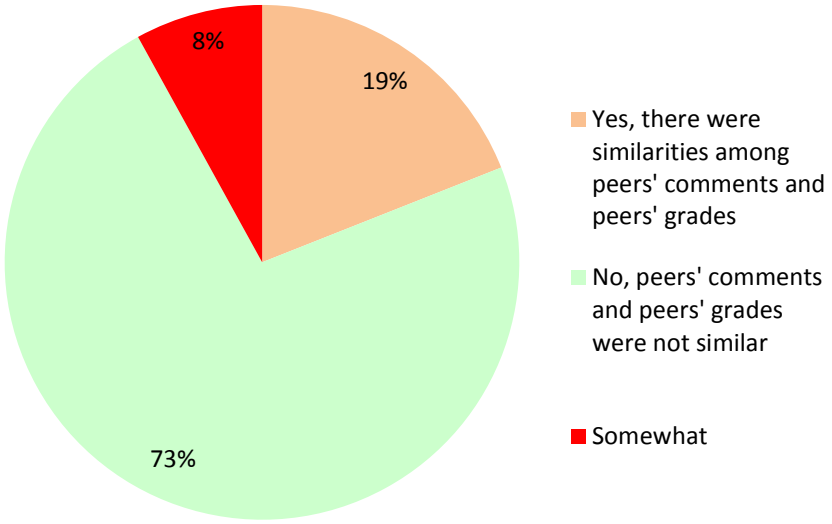

Figure 7. *Survey participants' opinions on similarities and differences in peers' comments and peers' grades.*

Below are examples of the participants' statements:

> *"Very different. I got all kinds of different comments from peers – general, specific, positive, negative. The grades were different, too. One peer have me low grade, and another gave me higher grade."*
>
> (Female 1007, Malaysia)

> *"The grades I got from peers were quite close, so I think peers' grades were similar. I also got similar comments from peers – I mean peers wrote in different words but they pointed out the same errors and had similar tone."*
>
> (Female 714, China)

> *"Sometime I received similar comments and close grades, and sometimes peer feedback was extremely different."*
>
> (Male 351, Japan)

2. Results for Research Question 2

Research Question 2. Is there a significant difference between grades given by peers in response to student drafts and those given by the instructors?

Question two examined whether there was a significant difference between grades given by student peers in response to student drafts and those given by the instructors. To answer this question, I conducted an analysis of a total of 600 student essays, including 100 Draft 1 and 100 Final Draft of the three writing assignments (these 600 essays were graded by both student peers and instructors). According to the course policy, each student's draft was graded by three different peers, and the final grade that the student received was the average of grades by each of the three student peers. The grades the students received were in the number forms (from 0 to 10), which were then translated in letter grades as follows: Grade A (9.0 – 10.0), Grade B (8.0-8.9), Grade C (7.0-7.9), Grade D (6.0-6.9) and Grade F (0.0 – 5.9). To report the results of the analysis here, I look at the similarities and differences among grades given by students according to draft 1 and final draft of three writing assignments to find out if there was any consistency among the grades given by the student peers and whether the students tended to give the grades within a certain range that differs from the range of the grades given by the teachers. I also looked at the average grades of every draft of each of the writing assignments given by student peers, and then compared with grades by instructors.

Table 15 below shows Writing Assignment 1's grades given by student peers and instructors. As can be seen in Table 15, in 24 students' Draft 1 (24%), all of the three peers gave the same grades within a certain range of letter grade. For example, SP5, SP8, SP13, SP17, SP62, and SP71 had all three student peers giving grade A (9.0 – 10.0) to the same paper they

assessed; SP14, SP15, SP75, SP85, and SP96 had all three peers giving grade B (8.0-8.9) to the same paper they assessed. Also seen in Table 15, 61 students' Draft 1 (61%) had two student peers gave the same grade range while the other gave different grade range (either lower or higher). For example, in SP1 peer 1 and peer 2 gave grade C (7.0 – 7.9) whereas peer 3 gave grade B (8.0 – 8.9); in SP38, peer 1 gave grade C (7.0-7.9) whereas peer 2 and peer 3 gave grade A (9.0-10.0). On the other hand, fifteen students' Draft 1 (15%) had peers give all different grade ranges. For example, in SP30, peer 1 gave grade B (8.0 – 8.9), peer 2 gave grade D (6.0-6.9) and peer 3 gave grade C (7.0-7.9); in SP79, peer 1 gave grade C (7.0-7.9), peer 2 gave grade A (9.0-10.0), and peer 3 gave grade D (6.0-6.9).

As to the comparison between grades given by student peers and grades given by instructors on the same papers, in 11 students' Draft 1 (11%), all three peers' grades were in the same range with the instructors' grades; for example, peers and instructors gave grade A to SP13; grade B to SP14, SP15, SP42, SP75, and SP85; grade C to SP6, SP18, SP34, SP51, SP93. In 21 students' Draft 1 (21%), two peers gave the same grade range as the instructors' while the other peer gave grades in the range that differed from the instructors' (either lower or higher). For example, in SP64, grades given by peer 1 and peer 2 were in the same range with the instructor's grade - grade C (7.0-7.9) whereas peer 3 gave a higher grade – grade B (8.0-8.9). In addition, there were 53 students' Draft 1 (53%) with one peer whose grade range given was the same to the instructor's and two peers whose grade range given were either higher or lower than the instructors' grade range. For example, in SP12, peer 3 gave grade B (8.0-8.9) that was in the same range to the instructor while peer 1 and peer 2 gave grade A (9.0-10.0) – higher than the instructor's grade range. Also, there were 15 students' Draft 1 (15%) in which none of the student peers gave the same grade range with instructors. For instance, in SP23, peer 1 and peer

2 gave grade A (9.0-10.0), peer 2 gave grade C (7.0-7.9) and instructor gave grade B (8.0-8.9); in SP59, peer 1 gave grade A (9.0-10.0), peer 2 and peer 3 gave grade B (8.0-8.9), and instructor gave grade C (7.0-7.9).

For the grades of Final Draft of Writing Assignment 1, Table 15 below shows that in 21 students' Final Draft (21%), all of the three peers gave the same grades within a certain range of grade (i.e. grade A, B, C, or D). For example, SP9, SP23, SP63 had all three student peers giving grade A (9.0 – 10.0) to the same paper they assessed; SP51, SP60, SP68 had all three peers giving grade B (8.0-8.9) to the same paper they assessed. Seventy-three students' Final Draft (73%) had two student peers who gave the same grade range while the other gave a grade in a different grade range (either lower or higher). For example, in SP38 peer 1 and peer 3 gave grade B (8.0 – 8.9) whereas peer 3 gave grade A (9.0-10.0); in SP91, peer 1 and peer 2 gave grade C (7.0-7.9) whereas peer 3 gave grade D (6.0-6.9). Six students' Final Draft (6%) had peers gave all different grade ranges. For example, in SP11, peer 1 gave grade A (9.0-10.0), peer 2 gave grade B (8.0-8.9), and peer 3 gave grade C (7.0-7.9); in SP86, peer 1 gave grade C (7.0-7.9), peer 2 gave grade D (6.0-6.9), peer 3 gave grade B (8.0 – 8.9).

As to the comparison between grades given by student peers and grades given by instructors on the same papers (see Table 15 below), in 9 students' Final Draft (9%), all three peers' grades were in the same range with the instructors' grades. For example, peers and instructors gave grade A to SP23; grade B to SP51; grade C to SP4. In 31 students' Draft 1 (31%), two peers gave the same grade range with the instructors' while the other peer gave grades in the range that differed from the instructors' (either lower or higher). For example, in SP40, grades given by peer 1 and peer 2 were in the same range with the instructor's grade - grade C (7.0-7.9) whereas peer 3 gave a lower grade – grade D (6.0-6.9). In addition, there were

46 students' Final Draft (46%) with one peer whose grade range given was the same as the instructor's and two peers whose grade range given were either higher or lower the instructors' grade range. For example, in SP59, peer 2 gave grade B (8.0-8.9), which was in the same range as the instructor's while peer 1 and peer 3 gave grade A (9.0-10.0) – higher than the instructor's grade range. Also, there were 14 students' Final Draft (14%) in which none of the student peers gave the same grade range with instructors. For instance, in SP2, peer 1 and peer 2 gave grade D (6.0-6.9), peer 2 gave grade B (8.0-8.9) and instructor gave grade C (7.0-7.9); in SP54, peer 1 gave grade D (6.0-6.9), peer 2 and peer 3 gave grade B (8.0-8.9), and instructor gave grade C (7.0-7.9).

Table 15

Writing Assignment 1's grades given by peers and instructors

Draft 1's Grades										Final Draft's Grades									
SP	**P1**	**P2**	**P3**	**Inst.**		**P1**	**P2**	**P3**	**Inst.**	**SP**	**P1**	**P2**	**P3**	**Inst.**		**P1**	**P2**	**P3**	**Inst.**
SP1	7.5	7.7	8.5	8	**SP51**	7.8	7.8	7.8	7.7	**SP1**	7.8	8	8	8	**SP51**	8.3	8.6	8.1	7.7
SP2	6	6	8	7.1	**SP52**	9	9	7	7.8	**SP2**	6.4	6.5	8.1	7.3	**SP52**	8	9	8	8.1
SP3	7.5	8.4	8.5	8	**SP53**	7.2	6	7.8	6.5	**SP3**	8.5	9	9	8	**SP53**	8	7	7	7
SP4	6.3	7.3	9	6.9	**SP54**	8.2	9.4	6.6	6.4	**SP4**	7	7.8	7	7.2	**SP54**	6.4	8.6	8	7
SP5	9.2	9.5	9.5	8.5	**SP55**	7.4	7.8	8.5	7.8	**SP5**	8.2	8.7	8.4	8.7	**SP55**	7.9	7.8	8.1	8
SP6	7.8	7.6	7.8	7.5	**SP56**	9	8	7	8	**SP6**	9	8	8	7.7	**SP56**	7.5	6.5	8.2	8
SP7	6	6.3	8	6.5	**SP57**	8	7.4	8.5	8	**SP7**	7	6.5	6.5	7	**SP57**	8.5	9.5	8.5	8
SP8	9	9.5	9.5	8.2	**SP58**	6.6	7	9	6.8	**SP8**	8	9	9	8.5	**SP58**	6.9	6.1	7	7.2
SP9	9	9	7.9	9	**SP59**	9.1	8	8.4	7.4	**SP9**	9	10	9	9.2	**SP59**	9	8	9.7	8.1
SP10	7.8	7.5	8	7.5	**SP60**	7.8	7.7	7.9	7.5	**SP10**	8	9	8	7.7	**SP60**	8.4	8.6	8	7.7
SP11	7	6.3	7	6.8	**SP61**	7.2	6.2	6.1	6.5	**SP11**	9	8	7	7	**SP61**	7	8	6	6.9
SP12	9	9	8.7	8.2	**SP62**	9	9.8	9.2	7.3	**SP12**	7.5	9.5	9	8	**SP62**	8.7	8.3	9	7.9
SP13	9	9	9	9	**SP63**	9	9.3	7.5	9	**SP13**	7.9	9.5	9.6	9.2	**SP63**	9.6	9.5	9	9.2
SP14	8	8.5	8.5	8	**SP64**	7.8	7.6	8	7.5	**SP14**	8.5	9	8.7	8.2	**SP64**	9	8	8	7.7
SP15	8.5	8.4	8.5	8.3	**SP65**	7.9	6.3	8.3	6.8	**SP15**	8.5	8.4	9	8	**SP65**	8.7	6.3	7	7.1
SP16	6.4	7.2	9	7	**SP66**	8.9	9.4	9.7	8.2	**SP16**	7.3	7.3	8	7.2	**SP66**	8.2	9	8.8	8.5
SP17	9.1	9.4	9.5	8	**SP67**	8	9	9.4	7.9	**SP17**	9.4	9.1	9.5	9	**SP67**	8.4	9.5	9	8.8
SP18	7.7	7.7	7.2	7.5	**SP68**	8.1	8.4	8.5	7.6	**SP18**	6.8	8.6	8	7.5	**SP68**	8.7	8.8	8.7	8.2
SP19	7.5	7.7	8.5	8	**SP69**	8.7	8.2	8.5	8	**SP19**	7	8.9	8.1	8	**SP69**	8.5	8.6	9.3	8.2
SP20	6	6	7.5	7.1	**SP70**	7.1	7.1	9	6.9	**SP20**	6.5	6.5	7.5	7	**SP70**	6.6	7.4	8.9	7
SP21	8.6	8.4	8.4	7.2	**SP71**	9.5	9.4	9	8.1	**SP21**	9.2	9.3	9	8	**SP71**	9	9	8.5	8.5
SP22	6.3	8.3	8	7.9	**SP72**	8	7.6	7.8	7.4	**SP22**	9	7	7	8	**SP72**	7.9	7.5	8	7.5
SP23	9	7	9.6	8.5	**SP73**	7.4	7.8	8.5	7.9	**SP23**	9.6	9.9	9	9	**SP73**	7.8	7.9	8.1	8
SP24	7.5	7.9	8	7.2	**SP74**	7.5	7.6	9	8	**SP24**	8.3	8.7	8	7.7	**SP74**	6.5	7	7	6.6
SP25	8	8.5	7	7.5	**SP75**	8.5	8.2	8.7	8	**SP25**	7.5	7.6	7.5	8	**SP75**	9	9	9	8
SP26	6	6.5	7.8	7	**SP76**	9	7.1	6.5	7.5	**SP26**	6.7	6.6	7.5	7.2	**SP76**	7.5	8	7.5	7.8
SP27	9	7.5	9.5	8.6	**SP77**	9	8.8	8.5	8.5	**SP27**	9.5	8.9	9.4	8.5	**SP77**	8	10	8.5	8.4

SP28	7.5	7.6	8	7.4	**SP78**	7.8	8.6	7	7.1	**SP28**	7.8	8.6	7	7.6	**SP78**	9	7	9	7.7
SP29	7.5	7.7	8.5	8	**SP79**	7.5	9	6	7.5	**SP29**	7.8	7.9	8.1	8	**SP79**	7.2	7.2	8.2	8
SP30	8	6.7	7.2	6.9	**SP80**	7	6.5	7	7	**SP30**	7	6.5	7	7	**SP80**	6.9	6.5	7.4	7.1
SP31	8.5	8.4	8.5	7.1	**SP81**	9.6	8.9	9.2	7.8	**SP31**	8.5	8.7	9.3	8	**SP81**	8	9.2	9.5	8.2
SP32	6.4	7.4	8.8	7	**SP82**	8	6.6	8	7.3	**SP32**	7.8	7.2	8	7.4	**SP82**	8	7.6	7.8	7.8
SP33	9	8	9.1	9	**SP83**	8	7.2	8.5	7.8	**SP33**	9.4	9	9.1	8.8	**SP83**	7.8	8	8	8
SP34	7.7	7.6	7.9	7.8	**SP84**	8	7	6.5	7	**SP34**	9	8	8	8	**SP84**	7.2	6.7	7	7.5
SP35	6.8	7.2	7.5	7.8	**SP85**	8	8.9	8.5	8	**SP35**	7.9	7.2	7.5	7.9	**SP85**	8.5	8.5	9	8
SP36	7	9	8	7	**SP86**	6.9	7.7	8	7	**SP36**	7.2	6.5	7.1	7.2	**SP86**	7.7	6	8	7
SP37	8.3	7.3	7	7	**SP87**	8.7	9	9.5	8.5	**SP37**	6.2	7.8	6	7	**SP87**	7.5	9	9.2	8.6
SP38	7.9	9	9.5	8.2	**SP88**	6.8	8.6	8	7	**SP38**	8	10	8	8.7	**SP88**	8.4	8.6	8	8
SP39	7.8	7.4	8	7.5	**SP89**	7.5	8	7.5	7.8	**SP39**	8.5	8.5	8	7.9	**SP89**	7.5	7.5	7.6	8
SP40	8	6.2	8.1	7	**SP90**	8	8	7	7	**SP40**	7	7	6	7	**SP90**	7	7	6.8	7.2
SP41	9.1	9.4	9.5	8.2	**SP91**	8.3	7.2	8.1	7.3	**SP41**	9	8	9	8.3	**SP91**	7	7	6	7.4
SP42	8.5	8.5	8.3	8	**SP92**	9.4	9.6	9	8.4	**SP42**	8	8.4	7.5	8	**SP92**	8.4	9.3	8.5	8.4
SP43	6	7.6	9	7	**SP93**	7.8	7.9	7.7	7	**SP43**	8.3	7.3	7	7	**SP93**	9.1	9	8	7.2
SP44	9.5	8.4	9.1	8	**SP94**	9.2	9.3	8.5	7.5	**SP44**	9.6	9.8	9.8	8.5	**SP94**	6.8	7	6.8	8.1
SP45	7.7	8.1	7.6	7.5	**SP95**	9.6	8.9	9.5	8.2	**SP45**	7.8	8.1	7.5	7.4	**SP95**	8.2	8.8	9	8.5
SP46	7.5	7.9	8.3	8	**SP96**	8.5	8.8	8.1	7.9	**SP46**	8.8	7.9	7.1	8	**SP96**	7.1	8.4	8.4	8
SP47	6	8	7.5	7.4	**SP97**	7.9	7.3	9.7	7.1	**SP47**	7	6.9	7.3	8	**SP97**	8.3	7.3	7	8
SP48	8.5	8.7	8.2	7.1	**SP98**	9.5	8.7	9.3	8	**SP48**	9.8	9.2	9	8	**SP98**	9	7.5	9.5	8
SP49	6.3	8.3	8	7	**SP99**	7.8	7	8.6	7.6	**SP49**	8	8	7	7	**SP99**	8	7.5	7.9	7.5
SP50	9	8	9.5	8.7	**SP100**	7.6	8	8.1	7.9	**SP50**	9.4	9.9	8	8.5	**SP100**	8.2	7.9	7.7	8.5

Notes:

SP: Student paper
P1: Peer 1
Inst.: Instructor
P2: Peer 2
P3: Peer 3

Table 16 below shows Writing Assignment 2's grades given by student peers and instructors. According to Table 11, in 16 students' Draft 1 (16%), all of the three peers gave the same grades within a certain range of grade (i.e. grade A, B, C, or D). For example, SP3, SP33, SP34 had all three peers that gave grade A (9.0 – 10.0) to the same paper; SP71 had all three peers giving grade C (7.0-7.9) to the same paper they assessed. Besides, 66 students' Draft 1 (66%) had two student peers give the same grade range while the other gave a different grade range (either lower or higher). For example, in SP20 peer 1 and peer 2 gave grade B (8.0 – 8.9) whereas peer 3 gave grade C (7.0-7.9); in SP77, peer 1 gave grade C (7.0-7.9) whereas peer 2 and peer 3 gave grade A (9.0-10.0). Eighteen students' Draft 1(18%) had peers give all different grade ranges. For example, in SP72, peer 1 gave grade A (9.0-10.0), peer 2 gave grade B (8.0-8.9), and peer 3 gave grade C (7.0-7.9); in SP92, peer 1 gave grade D (6.0-6.9), peer 2 gave grade C (7.0-7.9) and peer 3 gave grade B (8.0-8.9).

As to comparison between grades given by student peers and grades given by instructors on the same papers, in 10 students' Draft 1 (10%), all three peers' grades were in the same range with the instructors' grades; for example, peers and instructors gave grade B to SP5, SP6, SP14, SP35, SP44, SP95, SP96; grade C to SP9, SP48, SP71. In 32 students' Draft 1 (32%), two peers gave the same grade range with the instructors' while the other peer gave grades in the range that differed from the instructors' (either lower or higher). For example, in SP97, grades given by peer 1 and peer 2 were in the same range with the instructor's grade - grade C (7.0-7.9) whereas peer 3 gave a higher grade – grade A (9.0-10.0). In addition, there were 49 students' Draft 1 (49%) with one peer whose grade range given was the same to the instructor's and two peers whose grade range given were either higher or lower the instructors' grade range. For example, in SP69, peer 2 gave grade C (7.0-7.9) that was in the same range to the instructor while peer 1

and peer 3 gave grade B (8.0-8.9) – higher than the instructor's grade range. Also, there were 9 students' Draft 1(9%) in which none of the student peers gave the same grade range as instructors. For instance, in SP57, peer 1 gave grade C (7.0-7.9), peer 2 and peer 3 gave grade A (9.0-10.0) and the instructor gave grade B (8.0-8.9).

For the grades of Final Draft of Writing Assignment 2, Table 16 below shows that in 13 students' Final Draft (13%), all of the three peers gave the same grades within a certain range of grade (i.e. grade A, B, C, or D). For example, SP50 had SP3 had all three student peers giving grade C (7.0 – 7.9) to the same paper they all graded; SP91 had all three peers giving grade B (8.0-8.9) to the same paper they all assessed. Besides, 58 students' Final Draft (58%) had two student peers give the same grade range while the other gave a grade from a different grade range (either lower or higher). For example, in SP22 peer 2 and peer 3 gave grade C (7.0 – 7.9) whereas peer 1 gave grade A (9.0-10.0); in SP72, peer 1 gave grade B (8.0-8.9) whereas peer 2 and peer 3 gave grade A (9.0-10.0). Twenty-nine students' Final Draft (29%) had peers give all different grade ranges. For example, in SP49, peer 1 gave grade A (9.0-10.0), peer 2 gave grade C (7.0-7.9), and peer 3 gave grade B (8.0-8.9); in SP88, peer 1 gave grade C (7.0-7.9), peer 2 gave grade A(9.0-10.0) and peer 3 gave grade B (8.0-8.9).

As to comparison between grades given by student peers and grades given by instructors on the same papers (again see Table 16 below), in 4 students' Final Draft (4%), all three peers' grades were in the same range with the instructors' grades; for example, peers and instructors gave grade C to SP2, SP67; grade B to SP75, SP83. In 23 students' Final Draft (23%), two peers gave the same grade range with the instructors' while the other peer gave grades in the range that differed from the instructors' (either lower or higher). For example, in SP100, grades given by peer 1 and peer 3 were in the same range with the instructor's grade - grade B (8.0-8.9) whereas

peer 2 gave a lower grade – grade C (7.0-7.9). There were 55 students' Final Draft (55%) with one peer whose grade range given was the same as the instructor's and two peers whose grade range given were either higher or lower the instructors' grade range. For example, in SP4, peer 2 gave grade B (8.0-8.9) that was in the same range to the instructor while peer 1 and peer 3 gave grade A(9.0-10.0) – higher than the instructor's grade range. Also, there were 18 students' Final Draft (18%) in which none of the student peers gave the same grade range with instructors. For instance, in SP52, peer 1 gave grade A (9.0-10.0), peer 2 gave grade C (7.0-7.9), peer 3 gave grade B (8.0-8.9) instructor gave grade D (6.0-6.9).

Table 16

Writing Assignment 2's grades given by peers and instructors.

Draft 1's Grades										Final Draft's Grades									
SP	**P1**	**P2**	**P3**	**Inst.**		**P1**	**P2**	**P3**	**Inst.**	**SP**	**P1**	**P2**	**P3**	**Instr.**		**P1**	**P2**	**P3**	**Instr.**
SP1	8	7.6	8	7.5	**SP51**	7	8	8.5	7.7	**SP1**	8	7.8	9	8	**SP51**	8.1	9	9	8.1
SP2	7	6.3	7	6.5	**SP52**	8.1	7.3	9	7	**SP2**	7	7.2	7	7	**SP52**	9	7.2	8.7	6.9
SP3	9	9.5	9.5	8.2	**SP53**	8	8	9.5	8	**SP3**	9	9	9	8.5	**SP53**	9.1	9	9	8.7
SP4	8.8	9	9	8.5	**SP54**	8	8	8	7.5	**SP4**	9	8	9	8.2	**SP54**	9	7.6	8	7.7
SP5	8	8.5	8.5	8	**SP55**	8	7	7	7.5	**SP5**	8.3	9	8.8	8.2	**SP55**	8.5	7.2	9	7.5
SP6	8.5	8.4	8.5	8	**SP56**	9	6.5	8	7.5	**SP6**	9	8.4	8.6	8	**SP56**	9	6.8	7.5	7.9
SP7	7	7.3	9	7	**SP57**	7	9.5	9.5	8.5	**SP7**	7.4	7.3	8	7	**SP57**	9	7	8.6	8.6
SP8	9.1	9.5	9	8.5	**SP58**	9	7.6	8	7.5	**SP8**	9	7.7	9	8.5	**SP58**	7.6	9.4	8	7.5
SP9	7.8	7.5	7	7.5	**SP59**	8.5	6.5	8.5	8	**SP9**	7.9	7.6	8.1	7.5	**SP59**	9	8	8.1	8
SP10	7.5	7.7	8.5	8	**SP60**	7	6	7	6	**SP10**	8	7.9	9	8	**SP60**	8	7.9	7.5	6.5
SP11	7	7.5	8	7	**SP61**	7.7	7.6	8	7.5	**SP11**	8	9	7.4	7.2	**SP61**	8.2	9	8	7.6
SP12	9.3	8	7	8.5	**SP62**	8.5	6.3	6.8	6.8	**SP12**	9	9.1	9	8.5	**SP62**	8	6.6	7	7
SP13	7	7.8	9	7.5	**SP63**	8.9	9.5	9.5	8.2	**SP13**	9	7	9	7.7	**SP63**	7.6	9	9	8.4
SP14	8.2	8	8.5	8	**SP64**	7.9	9	9	8.5	**SP14**	7.5	9	8.7	8	**SP64**	9	8.4	9	8.7
SP15	7.1	8.4	8.5	8	**SP65**	7.5	8.5	8.5	8	**SP15**	8	8	9	8	**SP65**	7.8	9.4	8.7	8.2
SP16	7	8	9	7	**SP66**	9.8	8.4	8.5	8	**SP16**	6.7	9	9	7.4	**SP66**	8.5	8.5	7.2	8
SP17	7.8	9.5	9.5	8.5	**SP67**	6.2	7.3	9	7	**SP17**	9	9	8.5	8.6	**SP67**	7	7.3	7.9	7.2
SP18	7.8	7.6	6.5	7.5	**SP68**	8.7	9.5	9.5	8.5	**SP18**	7.8	8.3	8.1	7.5	**SP68**	8.8	9.2	9.5	8.5
SP19	9.2	7.7	8.5	8	**SP69**	8	7.6	8	7.5	**SP19**	8.6	7.9	8.1	8	**SP69**	8	8	8	7.6
SP20	8	8	7.8	7	**SP70**	8.1	7.7	8.5	8	**SP20**	6.9	7.5	7.5	7	**SP70**	7.4	8.3	8.1	8
SP21	7.6	8	8.5	7.7	**SP71**	7.2	7.5	7.2	7	**SP21**	7.2	9	8.8	8	**SP71**	8	9	7.6	7.2
SP22	8.4	7.3	9	7	**SP72**	9	8	7	8	**SP22**	9	7.6	7	7	**SP72**	8	9	9	8.3
SP23	8	8	9.5	8.5	**SP73**	8	7.8	9	7.5	**SP23**	8	9.4	9	8.7	**SP73**	9	9	8	7.7
SP24	8	8	8	7.5	**SP74**	7	8	8.5	7.5	**SP24**	9	8.5	8	7.7	**SP74**	7.7	9.2	8.7	8
SP25	7.4	9	7	7.5	**SP75**	7	8.4	8.5	8	**SP25**	7.3	7.2	9	8	**SP75**	8.1	8.4	8.9	8
SP26	8.9	6.5	7	7.5	**SP76**	8	8	9	7	**SP26**	9.2	7.3	7.8	8	**SP76**	7.6	9	9	7
SP27	8	9.5	9.5	8.3	**SP77**	7	9.5	9.5	7.9	**SP27**	9	7	8.9	8.5	**SP77**	9	8	8.3	8.5

SP28	7.6	7.6	8	7.5	**SP78**	7.8	7.6	6.6	7.5	**SP28**	7.7	9.2	8	7.5	**SP78**	8.4	8	8	7.5
SP29	8.4	6.5	8.5	7.2	**SP79**	8.9	7.7	8.5	8	**SP29**	8	7.9	8.2	8	**SP79**	9	7.3	8.1	8
SP30	7.2	7	8	7	**SP80**	6.7	8	8	7	**SP30**	8	7	7.5	7	**SP80**	8.5	6.5	7.4	7
SP31	7.4	7.6	8	7.5	**SP81**	7	8	8.5	7.7	**SP31**	7.4	9	8.2	7.7	**SP81**	7.1	8.9	9	8
SP32	6.9	6.3	7.3	6.7	**SP82**	8	7.3	9	7	**SP32**	7	8	7	7	**SP82**	9	7	7	7
SP33	9.1	9.5	9.7	8.2	**SP83**	8	8	9.5	7.2	**SP33**	8	9	9	8.5	**SP83**	8.2	8.9	8.8	8.7
SP34	9	9	9	8.1	**SP84**	8	8	8	7.5	**SP34**	9	9.3	9	8.2	**SP84**	9.1	9.5	8	7.7
SP35	8.2	8.5	8.5	8	**SP85**	9	8	7	7.6	**SP35**	8.2	9.1	8.7	8.1	**SP85**	7.5	7.3	9	8
SP36	7.6	8.4	8.5	8	**SP86**	9	6.5	7.5	7.5	**SP36**	8.5	8.3	7	8	**SP86**	9	8	7.8	8
SP37	6.7	7.3	9	7	**SP87**	7	9.5	9.5	8	**SP37**	8	7.3	8.9	7.1	**SP87**	8.6	7	9.5	8.5
SP38	8	9.5	9.5	8.5	**SP88**	9	7.6	8	7.5	**SP38**	9	8.8	9.5	8.5	**SP88**	7.8	9.6	8	7.5
SP39	8	7.6	8	7.5	**SP89**	8.5	6.5	8.5	8	**SP39**	7.8	8	8	7.6	**SP89**	9.2	7.9	8.1	8
SP40	8	7.7	8.5	8	**SP90**	7.2	6	7.2	6.5	**SP40**	9	7.9	8.1	8	**SP90**	8	7.1	7.5	6.7
SP41	8	7.5	8	7	**SP91**	7.3	7.6	8	7.5	**SP41**	8	9	6.4	7.7	**SP91**	8	8.9	8	7.7
SP42	7	8	7	7.8	**SP92**	6.6	7.8	8	6.8	**SP42**	7.7	7	9	8	**SP92**	7	8.2	7	7
SP43	7	7.8	9	7.5	**SP93**	7.6	9.5	9.5	8.2	**SP43**	9	7.6	9	8.3	**SP93**	7.8	9	9	8.5
SP44	8	8	8.5	8	**SP94**	9	9	9	8.8	**SP44**	7.8	9	8.8	8	**SP94**	9	9.4	9	8.6
SP45	9	8.4	8.5	8	**SP95**	8	8.2	8.5	8	**SP45**	7.9	8.4	9.1	8	**SP95**	8	9	8.4	8.5
SP46	7.6	8	9	7	**SP96**	8.5	8.6	8.3	8	**SP46**	7.5	9	9	7.1	**SP96**	8.5	7.6	8.1	8
SP47	8.8	9.5	8.9	8.5	**SP97**	7	7.3	9	7	**SP47**	9	7	7.5	8.9	**SP97**	6.8	7.3	8.9	7
SP48	7.9	7.6	7	7.5	**SP98**	9	9.5	7.2	8.5	**SP48**	7.4	7	8	7.5	**SP98**	8	7.5	9.5	8.6
SP49	9	7.7	8.5	8	**SP99**	7.8	8	8	7.5	**SP49**	9	7.9	8.4	8	**SP99**	8	7.4	8	7.5
SP50	7.6	8	7.3	8	**SP100**	8	7.7	8.5	8	**SP50**	7	7	7.5	8	**SP100**	8.2	7.9	8.4	8.1

Notes:

SP: Student paper
P1: Peer 1
Inst.: Instructor
P2: Peer 2
P3: Peer 3

Table 17 shows Writing Assignment 3's Grades given by student peers and instructors. As can be seen on Table 17 below, in 18 students' Draft 1 (18%), all of the three peers gave the same grades within a certain range of grade (i.e. grade A, B, C, or D). For example, SP74 and SP80 had all three student peers giving grade B (8.0 – 8.9) to the same paper they all graded; SP75 had all three peers giving grade C (7.0-7.9) to the same paper they all assessed. In addition, 67 students' Draft 1 (67%) had two student peers give the same grade range while the other gave a different grade range (either lower or higher). For example, in SP90 peer 1 and peer 3 gave grade B (8.0 – 8.9) whereas peer 2 gave grade C (7.0-7.9. Fifteen students' Draft 1(15%) had peers give all different grade range. For example, in SP5, peer 1 gave grade D (6.0-6.9), peer 2 gave grade C (7.0-7.9), and peer 3 gave grade B (8.0-8.9).

As to comparison between grades given by student peers and grades given by instructors on the same papers (see Table 17 below), in 11 students' Draft 1 (11%), all three peers' grades were in the same range with the instructors' grades; for example, peers and instructors gave grade C to SP45, SP49, SP90. In 43 students' Draft 1 (43%), two peers gave the same grade range with the instructors' while the other peer gave a grade in a range that differed from the instructors' (either lower or higher). For example, in SP63, grades given by peer 1 and peer 2 were in the same range with the instructor's grade - grade C (7.0-7.9) – whereas peer 3 gave a higher grade – grade A (9.0-10.0). In addition, there were 41 students' Draft 1 (41%) with one peer whose grade range given was the same to the instructor's and two peers whose grade range given were either higher or lower than the instructors' grade range. For example, in SP69, peer 2 gave grade C (7.0-7.9), which was in the same range as the instructor's while peer 1 and peer 3 gave grade B(8.0-8.9) – higher than the instructor's grade range. Also, there were 5 students' Draft 1(5%) in which none of the student peers gave the same grade range with instructors. For

instance, in SP41, peer 1 and peer 3 gave grade B (8.0-8.9), peer 2 gave grade D (6.0-6.9) and instructor gave grade C (7.0-7.9).

Regarding the grades of Final Draft of Writing Assignment 3, Table 17 below shows that 12 students' Final Draft (12%) had all three peers give the same grades within a certain range of grades (i.e. grade A, B, C, or D). For example, SP59 and SP95 had all three student peers giving grade B (8.0 – 7.9) to the same paper they all graded; Besides, 72 students' Final Draft (72%) had two student peers give the same grade range while the other gave a grade in a different grade range (either lower or higher). For example, in SP27 peer 1 and peer 3 gave grade A (9.0 – 10.0) whereas peer 2 gave grade B (8.0-8.9). Sixteen students' Final Draft (16%) had each of the peers with a different grade range. For example, in SP79, peer 1 gave grade A (9.0-10.0), peer 2 gave grade B (8.0-8.9), and peer 3 gave grade C (7.0-7.9).

As to comparison between grades given by student peers and grades given by instructors on the same papers, Table 17 below reveals that in 7 students' Final Draft (7%), all three peers' grades were in the same range with the instructors' grades; for example, peers and instructors gave grade A to SP57; grade B to SP59. In 24 students' Final Draft (24%), two peers gave the same grade range with the instructors' while the other peer gave grades in the range that differed from the instructors' (either lower or higher). For example, in SP42, grades given by peer 2 and peer 3 were in the same range with the instructor's grade - grade C (7.0-7.9) whereas peer 1 gave a higher grade – grade A (9.0-10.0). In addition, there were 43 students' Final Draft (43%) with one peer whose grade range given was the same to the instructor's and two peers whose grade range given were either higher or lower the instructors' grade range. For example, in SP35, peer 2 gave grade B (8.0-8.9) that was in the same range to the instructor while peer 1 gave grade C(7.0-7.9)- lower than the instructor's grade range and peer 3 gave grade A(9.0-10.0) – higher

than the instructor's grade range. Also, 26 students' Final Draft (26%) had none of the student peers giving the same grade range with instructors. For example, in SP36, peer 1 gave grade D (6.0-6.9), peer 2 gave grade B (8.0-8.9), peer 3 gave grade A (9.0-10.0) and the instructor gave grade C (7.0-7.9).

Table 17

Writing Assignment 3's grades given by student peers and instructors.

Draft 1's Grades						Final Draft's Grades									
SP	**P1**	**P2**	**P3**	**Inst.**		**P1**	**P2**	**P3**	**Inst.**	**SP**	**P1**	**P2**	**P3**	**Inst.**		**P1**	**P2**	**P3**	**Inst.**
SP1	7	7.5	8	7	**SP51**	7.2	8.9	8.7	7.6	**SP1**	8	9	6.6	7.2	**SP51**	8.6	9.1	9.6	8
SP2	9.1	8	7	8	**SP52**	7.5	7.3	8.2	7.1	**SP2**	9	9	9	8.5	**SP52**	7.8	8.7	7	7.5
SP3	7	7.8	9	7.5	**SP53**	8	8	9	8.1	**SP3**	9	8	9	7.7	**SP53**	7	8.9	8.8	8
SP4	8	7.2	8.5	8	**SP54**	8.6	8	8	8	**SP4**	8	9	8.7	8	**SP54**	8	7.5	7	8
SP5	6	7.4	8.5	7.5	**SP55**	6.8	7	7	7.5	**SP5**	8	7.4	8.5	8	**SP55**	8	8.6	9.2	8.3
SP6	8	8	9	8	**SP56**	8.5	7.3	7.5	7.5	**SP6**	9	9.2	9	9	**SP56**	8.1	8	7.8	8
SP7	8	8	9.5	8.5	**SP57**	7	8	9	8.5	**SP7**	8.4	7.1	9.5	9	**SP57**	9	9	9	8.6
SP8	7.8	7.6	7	7.5	**SP58**	8	8.1	8	7.5	**SP8**	7.4	8	8.4	7.5	**SP58**	7.3	8.7	9	7.7
SP9	8.7	7.6	8.5	8.3	**SP59**	7.7	7.5	8.4	8	**SP9**	9	7.9	8.1	8.9	**SP59**	8.5	8	8.1	8
SP10	6.8	8	7.1	7.5	**SP60**	8.2	7.1	8.1	8	**SP10**	7.6	6.7	7.5	8.8	**SP60**	9	7.4	7.5	8.5
SP11	7.6	7.9	8.5	7.7	**SP61**	7	7.5	7.6	7	**SP11**	7.2	8.8	9	8	**SP61**	8	7.3	6.7	7.2
SP12	8	8.3	9	8	**SP62**	8.6	8	7	7.8	**SP12**	9	7.4	8.1	8	**SP62**	9	9	8.1	8.5
SP13	8	9	9.2	8.5	**SP63**	7	7.8	9	7.5	**SP13**	8.8	9	9	8.7	**SP63**	9	8.2	9	7.8
SP14	8.1	8	8	7.5	**SP64**	8	7.3	8.5	7.5	**SP14**	9	9.6	8	7.7	**SP64**	8.3	9	8.7	8
SP15	7.5	6.7	7.6	7.5	**SP65**	6.9	8.4	8.5	8	**SP15**	7.5	7.4	9.1	8	**SP65**	8	8.4	9	8.1
SP16	7	8	8	8	**SP66**	7	8	8	7	**SP16**	7.7	9	9	8.1	**SP66**	8.4	9	9	8
SP17	8.7	8.6	8.8	8.5	**SP67**	9	9	9.1	8.7	**SP17**	9.2	9	9.4	9	**SP67**	9	8.5	9.5	8.5
SP18	8.1	7.4	7.2	7.5	**SP68**	7.8	7.6	7	7.5	**SP18**	8	9.2	8.2	7.7	**SP68**	8.6	8	8	8
SP19	9.5	9	8.8	8.5	**SP69**	8	7.7	8.5	8	**SP19**	8.6	7.9	8.7	9	**SP69**	8.7	7.9	8.1	8
SP20	6.8	8.1	6.9	7	**SP70**	8	8	6.5	7	**SP20**	8.1	6.8	7.5	7	**SP70**	7.5	8.8	7.5	7.3
SP21	7.2	8.9	8.7	7.5	**SP71**	7	8	8.5	7.7	**SP21**	8.3	9.2	9.6	8	**SP71**	7	8.9	9.6	8
SP22	7.5	7.3	8.2	8	**SP72**	7	7.3	9	7.5	**SP22**	8	8.7	7.8	8	**SP72**	9	8	8.1	8.1
SP23	7	8.1	9.5	8.1	**SP73**	7.9	8	9.5	8.5	**SP23**	7.3	8.7	9	8.5	**SP73**	7.6	9	9	8.7
SP24	7.9	8.7	7.8	7.6	**SP74**	8	8.4	8	7.5	**SP24**	8	9	8	8	**SP74**	9	9.1	8	8
SP25	7	7	7.2	7	**SP75**	7.5	7.5	7	7.5	**SP25**	8	8	9.2	8.3	**SP75**	7.8	7.2	9	8.2
SP26	8.5	7.8	7	7.5	**SP76**	7	8	7	8	**SP26**	8	9	7.8	8	**SP76**	6.6	9	9	8

SP27	7	9.5	9.1	8.5	**SP77**	7.7	8.6	8.8	8.5	**SP27**	9	8.6	9	8.5	**SP77**	9.2	9.8	9.5	8.6
SP28	8.2	7.6	8	7.5	**SP78**	8.1	7.6	6.8	7.5	**SP28**	8.7	9.5	9	7.7	**SP78**	8.7	9.2	8.1	7.7
SP29	8.4	7.5	8	8	**SP79**	9.5	8.7	8.3	8.1	**SP29**	7.6	7.8	8.1	8	**SP79**	9.2	8.7	7.5	9
SP30	8.2	8	8	8	**SP80**	8	8.2	8	7.9	**SP30**	9	9	7.5	8.6	**SP80**	8	8	7.6	8
SP31	7	7.5	8	7	**SP81**	7.2	8.9	8.7	7.5	**SP31**	8	9	8.9	7.2	**SP81**	8.4	9.1	9.6	7.5
SP32	9	8	7	8.5	**SP82**	7.5	7.3	8.2	7.1	**SP32**	9.5	8.9	9	8.5	**SP82**	8.8	8.7	7.3	8
SP33	7.4	7.8	8.9	7.5	**SP83**	7	8	9.3	8.1	**SP33**	9	8.1	9	7.7	**SP83**	7	8.9	9	8.2
SP34	8	9	8.5	8	**SP84**	8.8	8	7.8	7.6	**SP34**	7	8	8.7	8	**SP84**	8	9.5	8.5	8
SP35	7.8	8.4	8.5	8	**SP85**	7	6	7	7	**SP35**	7.9	8.3	9	8	**SP85**	7.9	7.3	9.2	8.3
SP36	8	8	8.2	7	**SP86**	8.5	8	7	7.5	**SP36**	6.9	8.8	9	7.8	**SP86**	8.1	8	8	8
SP37	8	7.6	8.5	8.1	**SP87**	7	9.5	8	8.1	**SP37**	9	7.4	9.6	8.5	**SP87**	9	9.3	9.4	8.5
SP38	7.4	7.6	8.1	7.5	**SP88**	8.2	7.7	8	7.5	**SP38**	7.5	7.6	7.3	7.5	**SP88**	8	9.5	8.9	7.7
SP39	9	7.7	8.5	8	**SP89**	8	7.8	8.5	8	**SP39**	8	7.9	8.1	8	**SP89**	8.5	7.9	8.1	8
SP40	7.1	8	6.9	7	**SP90**	7.2	7.1	7.5	7.2	**SP40**	8	8.4	7.5	7.2	**SP90**	9	8.2	7.5	8
SP41	8	6.9	8.5	7.7	**SP91**	7	7.5	8	7	**SP41**	9	9	9.3	8.1	**SP91**	8	8.8	8.4	8.3
SP42	8	8.3	9	7.3	**SP92**	9.3	8	7	8.3	**SP42**	9	7	7	7	**SP92**	9	8.3	9	8.3
SP43	8	7.8	8.7	8.5	**SP93**	7	8.8	9	7.5	**SP43**	8.3	8	9	8.7	**SP93**	9	8	9	7.7
SP44	8.2	8.1	8	7.5	**SP94**	8	7	8.5	8	**SP44**	9	9.9	8	7.7	**SP94**	7.1	9	8.7	8
SP45	7.8	7	7	7.5	**SP95**	9	8.2	8.5	8	**SP45**	7.5	8.2	9	8	**SP95**	8.1	8.4	8.9	9
SP46	7	8	8	8	**SP96**	8	7	9	7.6	**SP46**	7.7	9	9	8.1	**SP96**	7.9	9	9.3	8.1
SP47	7.9	8.6	8.3	8.5	**SP97**	8	7.5	8.7	8	**SP47**	8.3	8.8	9.5	8.6	**SP97**	9	8	9.1	8.5
SP48	7.9	7.6	7.1	7.5	**SP98**	7.8	8.6	7	7.5	**SP48**	8.7	8.2	7.5	7.7	**SP98**	7.8	7.2	8.3	7.5
SP49	9.2	8.7	9.4	8.4	**SP99**	9	7.7	8.5	8	**SP49**	9.2	7.5	8.1	8.3	**SP99**	9	7.9	8.1	8
SP50	8	8.2	7.6	7.2	**SP100**	7.6	8	7.2	7.5	**SP50**	8	8.5	7.5	7.7	**SP100**	8.7	9	7.5	9

Notes:

SP: Student paper
P1: Peer 1
Inst.: Instructor
P2: Peer 2
P3: Peer 3

From the analysis of the letter grades, Table 18 below provides an overview of the grades that student writers received from peers and instructors. Throughout three of the writing assignments, F grades were given by neither peers nor instructors. As to the comparison between the grades assigned by student peers with those assigned by instructors on the same writing assignment, there were examples that showed great similarities and examples that showed differences. For example with similarities in the number of grades given, students and instructors assigned the same number of grade Bs - 30Bs (30%) for students' graded draft 1 – Writing Assignment 1 and close to the number of grade Bs for students' graded draft 1 – Writing Assignment 3. For example with even greater difference in the number of letter grades given, grade As assigned by student peers outnumbered grade As assigned by instructors: students' draft 1 – writing assignment 1 had 16As (16%) by peers and 3As (3%) by instructors; students' final draft – writing assignment 2 had 10 As (10%) by student peers but zero As (0%) by instructors. In addition, instructors on the whole gave more low grades (i.e. Cs and Ds) on first drafts than student peers: for draft 1 – writing assignment 1, student peers gave 47 Cs (47%) and 7Ds (7%) while instructors gave 57Cs (57%) and 10 Ds (10%); for draft 1 – writing assignment 2, student peers gave 42 Cs (42%) and 4Ds (4%) while instructors gave 51Cs (51%) and 6 Ds (6%).

Table 18

Letter grades given by student peers and instructors.

Grades	Writing Assignment 1				Writing Assignment 2				Writing Assignment 3			
	Draft 1		Final Draft		Draft 1		Final Draft		Draft 1		Final Draft	
	Peer	Instr.	Peer	Instr.	Peer	Instr.	Peer	Instr.	Peer	Instr.	Peer	Instr.
A	16	3	15	5	7	0	10	0	5	0	11	3
B	30	30	36	52	47	43	66	56	41	42	61	58
C	47	57	33	41	42	51	24	41	39	48	23	39
D	7	10	16	2	4	6	0	3	15	10	5	0
F	0	0	0	0	0	0	0	0	0	0	0	0

Since the grade each student actually received on each draft of a writing assignment was the average of the grades given by the three peers on the same paper, to find out whether there was a significant difference between grades actually given by student peers and instructors, I ran six paired t-tests (also known as dependent t-tests) on each of three writing assignments' grades given by the student peers and the instructors. These six paired t-tests were conducted to compare the mean of overall grades of the first drafts and those of the final drafts given by student peers and overall grades given by the instructors in three writing assignments (application essay, summary and evaluation essay, and argumentative essay). The following Tables 19, 20, and 21 show the results of the paired-t-tests.

As illustrated in Table 19 below, for grades of Writing Assignment 1 – Application Essay, the results of the paired t-tests showed extremely statistically significant differences between overall grades given by peer students and those by the instructors in both Draft 1 and

Final Draft. The two-tailed P value is less than 0.0001, so by conventional criteria, this difference is considered to be extremely statistically significant. As observed, the mean of grades given by peers tended to be higher than the mean of grades given by the instructors. This held true for both drafts. In terms of letter grades translated from the number grades, the mean of student grades for Draft 1 was grade B while the instructors' was grade C.

Table 19

Paired t-test analysis for overall grades of Writing Assignment 1 – Application Essay.

Drafts	N	Peers' grading		Instructors' grading		Diff. Mean	t	p	Cohen's d
		Mean	SD	Mean	SD				
Draft 1	100	8.04	0.90	7.57	0.82	0.47	7.44	0.000*	0.54
Final Draft	100	8.14	0.97	7.71	0.81	0.43	8.76	0.000*	0.48

* Significant at $p < .05$

Regarding the grades of Writing Assignment 2 – Summary and Evaluation Essay, the mean of peer grading in both Draft 1 and Final Draft was higher than the mean of the instructor grading. The two-tailed P value for Draft 1 was 0.017, showing that the difference is considered to be statistically significant while P value for Final Draft was less than 0.0001, which means this difference was considered to be extremely statistically significant. See Table 20 below:

Table 20

Paired t-test analysis for grades of Writing Assignment 2- the Summary and Evaluation Essay.

Drafts	N	Peers' grading		Instructors' grading		Diff. Mean	t	p	Cohen's d
		Mean	SD	Mean	SD				
Draft 1	100	7.83	1.01	7.66	0.75	0.17	2.43	0.017	0.19
Final Draft	100	8.18	0.69	7.80	0.07	0.38	7.88	0.000*	0.77

* Significant at $p < .05$

Similar to the paired t-test results for grades of Writing Assignments 1 and 2, as can be seen from Table 21 below, for Writing Assignment 3 – Argumentative Essay, statistically significant differences in grading between student peers and instructors existed in Draft 1 and extremely statistically significant differences existed in Final Draft. Specifically, the mean of peer grading in Draft 1 (7.83) and Final Draft (8.18) was higher than the mean of the instructor grading (7.66 and 7.80, respectively). The two-tailed P value for Draft 1 was 0.009, showing that the difference is considered to be statistically significant while the two-tailed P value for Final Draft was less than 0.0001, which means this difference was considered to be extremely statistically significant.

Table 21

Paired t-test analysis for grades of Writing Assignment 3- Argumentative Essay.

Drafts	N	Peers' grading		Instructors' grading		Diff. Mean	t	p	Cohen's d
		Mean	SD	Mean	SD				
Draft 1	100	7.78	0.67	7.64	0.61	0.15	2.67	0.009	0.22
Final Draft	100	8.29	0.58	7.85	0.51	0.43	10.7	0.000*	0.81

* Significant at $p < .05$

In conclusion, the analysis of the grade actually given by the three student peers in relation to the grade given by the instructors indicates that there was some consistency among the student peers' grades and instructor's grades given. However, there were still differences between student peers' grades as well as between student peers' grades and instructors' grades: students tended to give grades within a certain range that differs from the range of the grades given by peers and by the instructors. The statistics of paired t-tests showed that there were significant differences in overall grades given by student peers in response to student drafts and those given by the instructors. Specifically, for both Draft 1 and Final Draft in three Writing Assignments, although using the same rubrics, student peers gave higher grades to student drafts than the instructors did.

3. Results for Research Question 3

Research Question 3. To what degree and in what way do comments/feedback given by peers in response to student drafts differ from those given by the instructors?

Question three examined differences between comments given by student peers in response to drafts written by peers and those given by the instructors participating in the composition MOOC course in this study. To answer this question, I conducted an analysis of a selective sample of 20 essay first drafts out of the 100 drafts total (for assignment #3) to analyze comments made on these essays by the student peers and the instructors.

To report the results of this analysis, I look first at the total number of comments made by each group in relation to the other and also by each assessor individually in relation to other individual assessors, considering patterns of similarity and difference that might emerge, or not, from these comparisons. Results to be reported to answer this question also include the categories of WHAT (idea development, organization, or sentence level) and HOW (explanation, question, correction) that emerged from the coding of the comments by the individual peer responders and instructors. In reporting the results, I set forth first the number of comments provided by each group and then move on to set forth the issues—the WHAT—the individuals in each group—peers and instructor—focused on and what form the commenters used to communicate these issues—the HOW. The results include not just a comparison of the two groups' comments on average but also a comparison of how each individual peer compares to the instructor in the number of comments and in the kinds of WHAT and HOW were offered to the writer of the paper in question.

With the 20 essays selected for the analysis in this present study, the total number of comments made by three student peers per essay were 865 (the mean was 289), which resulted in an average of approximately 14 comments per paper. A total of five instructors graded and commented on the essays of student participants, but only one instructor commented on each essay (unlike peers, with three peers commenting on each essay). The total number of comments by instructors was 278 (also with the average of approximately 14 comments per paper (see Table 22 below). However, it is clear that, beyond the averages, there is a great range in the number of comments among the sets of three peers assigned to grade the same paper. For example, as can be seen in Table 22, in student paper 1 [SP1], Peer 1 had 21 comments whereas Peer 2 had 12 and Peer 3 fell somewhere in between those two extremes. The same type of divergence in total number of comments can be seen in the peer comments for SP7: in SP7, Peer 1 had 7 comments while Peer 2 had 20 comments and Peer 3 had 16. Another example shows a similar spread in numbers: for SP16, Peer 1 had only 10 comments as compared to Peer 2's 22 comments and Peer 3's 18 (see Table 22 below). In the case of each of these papers, the comment number for the instructor was very close in number to only one of the three peers. For SP1, Peer 1's 21 comments come close to the instructor's 24 comments; for SP7: Peer 1's 7 comments are close to the 9 comments by the instructor; for SP16: the 10 comments by Peer 1 comes very close to the 11 comments by the instructor. This degree of similarity in the number of comments from instructor to each of the three peers is far from always in evidence. The difference in comment numbers also becomes very clear when comparing the number of peer comments to the number of instructor comments, using a peer-by-peer comparison with the instructor on comment numbers rather than a comparison using aggregate of the three peers. One of the examples show the three peers closer together in the number of comments as

compared to the relatively low number of instructor comments: in SP10, the instructor made 6 comments whereas Peer 1 had 13; Peer 2, 13; Peer 3, 16. Another example shows the three peers with gain relatively low close together in the number of comments as compared to the relatively high number of instructor comments: in SP3, the instructors had 24 comments while Peer 1 had 13; Peer 2 had 16; Peer 3, 12. Overall, these two specific examples show the variety to the extent that the peers' comment numbers in being similar to each other relative to the instructors' number of comments differ from points presented earlier in this paragraph that peers differ greatly among themselves in the number of comments they offer.

Table 22

Number of comments by student peers and instructors

Students' papers (SP)	Number of peer students' comments				Number of instructors' comments
	Peer 1	Peer 2	Peer 3	Means	
SP1	21	17	12	17	24
SP2	15	13	15	14	25
SP3	13	16	12	14	24
SP4	17	12	14	14	8
SP5	20	13	11	15	14
SP6	18	15	10	14	11
SP7	7	20	16	14	9
SP8	14	16	21	17	9
SP9	15	11	6	11	10
SP10	13	13	16	14	6
SP11	17	12	11	13	16
SP12	12	11	12	12	18
SP13	22	13	9	15	12
SP14	8	15	15	13	7
SP15	16	17	23	19	16
SP16	10	22	18	17	11
SP17	12	15	13	13	17
SP18	20	14	14	16	10
SP19	11	11	18	13	13
SP20	15	10	17	14	18
Total				**289**	**278**

For the analysis of the comments by peers and instructors, I employed a coding scheme with two major categories – the WHAT and the HOW. The WHAT consisted of three criteria - Idea Development, Organization, and Sentence Level, and the HOW consisted of three criteria: Explanation, Question, and Correction (more details can be seen in Chapter 3 – section 4). The coding consisted of the coding of the comments that fell under each category of the three umbrella categories in the "WHAT" and that revealed patterns within each category in the "HOW". For the data analysis, each comment was given one point in accordance with the criteria in the "WHAT" and the "HOW". When more than one category in the "HOW" might seem relevant, one point would be given to each of the corresponding criteria. Table 23 below shows the number of comments on the WHAT by each of individual student peers on the same line as the instructor responding to the same student essay, draft 1 of writing assignment #3.

Table 23.

Number of comments on the WHAT by peer students and instructors

Student' papers (SP)	Number of comments on the WHAT												Odds*			
	Idea Development				Organization				Sentence Level							
	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.
SP1	1/21	4/17	4/12	2/24	1/21	1/17	2/12	2/24	19/21	12/17	6/12	20/24	0	1/17	0	0
SP2	2/15	1/13	3/15	1/25	0	2/13	1/15	0/25	12/15	9/13	11/15	24/25	1/15	1/13	0	0
SP3	4/13	4/16	5/12	7/24	1/13	1/16	1/12	1/24	5/13	11/16	6/12	15/24	3/13	1/16	0	1/24
SP4	1/17	0	2/14	4/8	0	0	0	0/8	16/17	12/12	11/14	4/8	0	0	1/14	0
SP5	5/20	1/13	2/11	4/14	2/20	1/13	1/11	1/14	13/20	11/13	8/11	9/14	0	0	0	0
SP6	5/18	3/15	6/10	1/11	1/18	3/15	1/10	1/11	11/18	7/15	4/10	9/11	1/18	2/15	0	0
SP7	1/7	4/20	4/16	2/9	1/7	2/20	2/16	4/9	5/7	14/20	10/16	3/9	0	0	0	0
SP8	5/14	3/16	5/21	4/9	1/14	1/16	2/21	1/9	7/14	10/16	12/21	4/9	1/14	2/16	2/21	1/9
SP9	2/15	2/11	2/6	6/10	1/15	1/11	1/6	1/10	12/15	8/11	3/6	3/10	0	0	0	0
SP10	3/13	4/13	4/16	5/6	1/13	0	0	0/6	7/13	9/13	9/16	1/6	2/13	0	3/16	0
SP11	3/17	3/12	2/11	1/16	0	0	1/11	1/16	14/17	9/12	8/11	14/16	0	0	0	0
SP12	0	0	1/12	0/18	0	0	0	0/18	12/12	10/11	11/11	18/18	0	1/11	0	0
SP13	4/22	3/13	3/9	4/12	0	0	0	1/12	18/22	10/13	6/9	7/12	0	0	0	0
SP14	2/8	4/15	2/15	5/7	2/8	1/15	1/15	0/7	2/8	10/15	11/15	2/7	2/8	0	1/15	0
SP15	3/16	3/17	4/23	4/16	0	0	0	2/16	13/16	10/17	17/23	10/16	0	4/17	2/23	0
SP16	1/10	5/22	4/18	4/11	1/10	0	1/18	2/11	8/10	17/22	13/18	5/11	0	0	0	0
SP17	0	2/15	2/13	4/17	0	0	0	0	12/12	13/15	8/13	11/17	0	0	3/13	2/17
SP18	5/18	5/14	3/14	4/10	1/18	0	0	1/10	10/18	9/14	9/14	5/10	2/18	0	1/14	0
SP19	0/11	3/11	5/18	4/13	2/11	1/11	2/18	2/13	9/11	7/11	10/18	7/13	0	0	1/18	0
SP20	3/15	3/10	4/17	6/18	0	1/10	0	2/18	10/15	6/10	13/17	10/18	2/15	0	0	0

*****: *comments that indicate praises or general reader responses or general/vague points*

According to Table 23 above, the number of comments on Idea Development made by student peers varied from paper to paper, with comments in this category falling into a range of approximately 30-40% of each peer's total comments on each of the papers by individual student peers, excluding peers who made zero or one or two comments as in SP4, SP12, SP17, etc. For most papers, instructors had up to 37% of their total comments on Idea Development (except for a few cases in which they far surpassed 37%: SP4 – 50%, SP9 – 60%, and SP14 -70%). However, in a few other papers, instructors made only one, or two, or even zero comments on Idea Development, as in SP1 (2/24), SP2 (1/25), SP6 (1/11), SP11 (1/16), SP12 (0/18).

As to the comparison between one peer with other peers and between peers with instructors commenting on the same paper, there could be examples that showed great similarities and great differences. For example, for SP1, Peer 1 had 1 comment out of 21 total (4.76%) on idea development; Peer 2 had 4 comments out of 17 (23.5%), and Peer 3 had 4 out of 12 total (33.35); and the instructor, out of 24 comments total, had 2 on idea development (8.3%). For an example with even greater difference in the number and type of comment relative to the total number of comments given, in SP6 Peer 1 had 5 comments out of 18 (27.8%) on Idea Development while Peer 2 had 3 out of 15 (20%), Peer 3 had 6 out of 10 (60%), and the instructor had 1 out of 25 (4%). This same difference in number of comments on Idea Development was also observed in some other papers such as SP4, SP5, SP7, SP14, SP17, and SP19. In a few papers such as SP8, SP13, SP15, SP18, student peers had similar number of comments on Idea Development, all of which were also close to the instructors' number of comments on the same issue. There were five student peers who did not make any comments on Idea Development (0%); they were Peer 2 (SP4), Peer 1 and Peer 2 (SP12), Peer 1 (SP17), and

Peer 1 (SP19). The instructor made no comment (0%) on Idea Development on one paper - SP12.

In terms of Organization category in the WHAT, as can be seen in Table 23 above, the majority of student peers made comments in this category from zero to approximately 15% of the overall total comments (1 to 2 comments on Organization on each paper). Instructors had a slightly lower percentage of comments on organization up to only 10% of their total comments on Organization (1-2 comments) on all of the papers, except SP7, which had about 45%. In a few instances (5 papers), peers and instructors had the same number of comments on Organization; they were SP3 (1 comment each), SP4 (0 comment), SP9 (1 comment), SP12 (0 comment), SP17 (0 comment). For other papers of the 20, the number of comments by peers and instructors coded for the Organization category was, though not identical, very close, only one comment difference either between peers or peers and instructors as in SP1, SP5, SP8, SP10, SP11, SP19. On SP6, there was one peer (Peer 2) who had 3 comments on Organization while Peer 1, Peer 3 and the instructor had 1 comment.

Regarding Sentence Level category in the WHAT, Table 23 above reveals that the majority of comments made by peer students focused on Sentence Level issues. All of the student peers (except P1 in SP3 and P1 in SP14) had more than 50% of their comments on sentence level issues such as grammar, spelling, punctuation, mechanics, and word use. A few student peers had 100% of the comments on sentence level issues, for example Peer 2 - SP4 (12/12), Peer 1 - SP12 (12/12), and Peer 3- SP12 (11/11). Unlike student peers, not all instructors devoted the majority of their comments to Sentence Level issues; however, most of them did. Fourteen out of 20 papers graded by the instructors had at least 50% of the comments on issues of sentence level. For SP12, the instructor had 100% (18/18) comments on sentence level issues.

In 10 out of 20 papers, quite number of student peers had the same (or almost the same) number of comments on the Sentence Level as the instructors, for example Peer 1 – SP1, Peer 3 – SP5, Peer 3 – SP9, Peer 1 – SP11, Peer 3 –SP13, Peer 1 – SP14, Peer 2 – SP15, Peer 1 – SP17, Peer 2 – SP19, and Peer 1 – SP20. There were three papers (SP2, SP3, SP12) in which the instructors' number of sentence level comments much exceeded all of the student peers'; however, the opposite trend was observed in quite a few papers, as in SP4, SP7, SP8, SP10, SP16, SP18. In some cases of the same papers, peers' number of sentence level comments much varied from each other and also from the instructors'. For instance, in SP1, Peer 3 had 6 Sentence Level comments while Peer 1 had 19 and Peer 2 had 12; in SP13 Peer 3 had 6 comments while Peer 1 had 18 and Peer 2 had 10; in SP16, Peer 1 had 8 while Peer 2 had 17 and Peer 3 had 13.

Also shown in Table 23 above, a few comments made by peer students and instructors did not belong to any categories in the WHAT and thus were put in the Odd category. These comments mainly expressed either praise (i.e. "*Nicely said*"; "*Nice job*"; "*Great*") or personal opinions (i.e. "*You're right"; "It's true!!!! I experienced the same*"). The instructors made such comments on three papers, for example (1 comment on SP3 "*I agree"*; 1 comment on SP8 *"Good body paragraph"* and 2 comments on SP17 *"Nice job!", "Well done"*). Six papers (SP5, SP7, SP9, SP11, SP13, SP16) had no praise or personal opinion from either student peers or instructors. On the other papers, 23 student peers made Odd comments, and the number of the Odd comments varied from 1 to 4. For example, on SP8 Peer 1 had 1 comment ("*Right*"), Peer 2 had 2 comments ("*You did a good job!";"Agreed"*) and Peer 3 had 2 comments ("*Nice!*"; "*Well done*"); on SP15, Peer 1 had no Odd comment while Peer 2 had 4 comments (e.g. *"?"; "Good ☺"; "great"; "Nice job"*) and Peer 3 had 2 comments (e.g. *"well done", "awkward"*).

Overall, in terms of the WHAT, among comments made by student peers, 20% of student peers' comments focused on idea development, 5.3% of the comments on organization, 70% of the comments on sentence level, and approximately 4.7% of the comments expressed praise or personal opinions or general/vague points. In regard to instructors' comments, the total of comments about Idea Development made up approximately 25.9%; Organization, 8%; Sentence Level, 64.7%; others (i.e. praises, personal opinions), 1.4%. Figure 8 below shows the percentages of the WHAT criteria displayed through comments made by student peers and instructors.

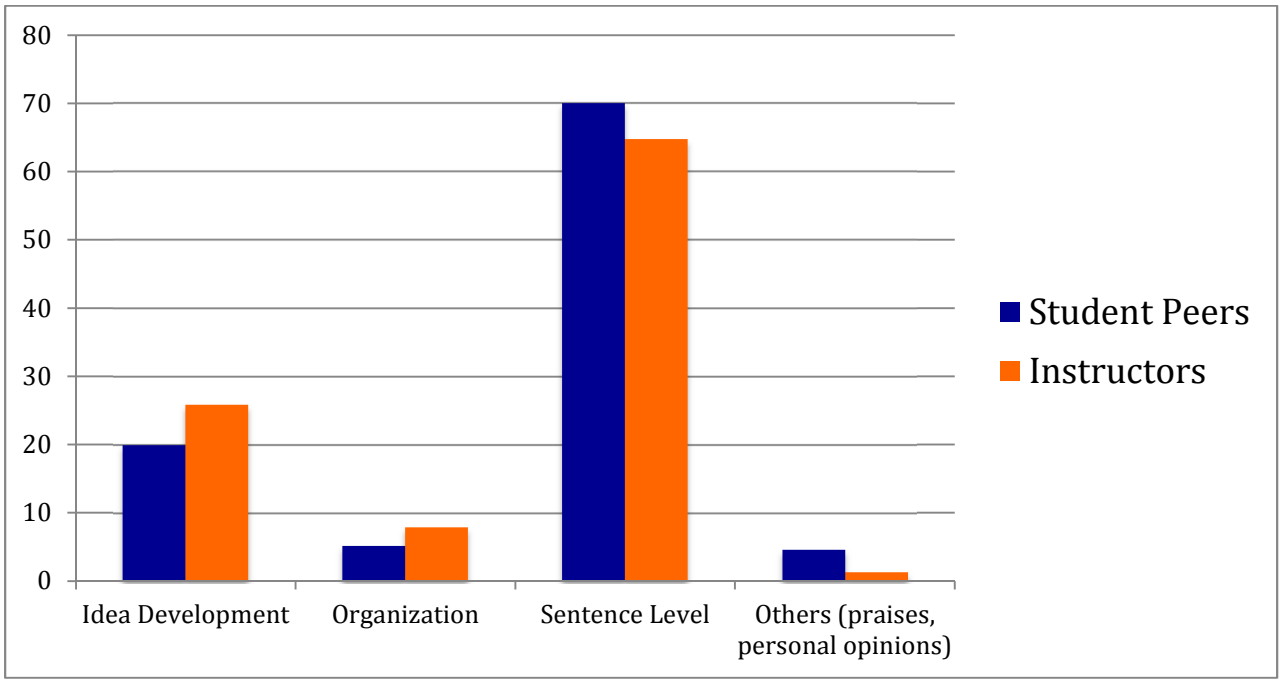

Figure 8. *WHAT categories displayed through comments made by student peers and instructors*

As noted in the methodology chapter (section 6.4), comments that expressed concerns about thesis, focus, development of ideas and clarity of ideas were marked as "Idea Development." Comments that expressed concerns about the ordering of and relationship between parts of the essay (i.e. introduction, body, and conclusion), text structures, transitions

between the ideas in writing, and logic of the ideas were marked as "Organization." Comments that expressed concerns about grammar, punctuation, spelling, mechanics, and vocabulary were marked as "Sentence Level." Table 24 below illustrates a few examples of comments categorized as Idea Development, Organization, and Sentence Level.

Table 24

Example comments made by student peers and instructors on the WHAT

The "WHAT"	Comments by student peers	Comments by instructors
Idea Development	• "*You need to add more examples to support the main idea in this paragraph.*" (SP3 – Peer 1) • "*Your main points are present but they lack details. You can give some examples to support the ideas*." (SP4 – Peer 3)	• "*You have a nice hook to open your introduction and you do a good job of stating your thesis*." (SP4) • "*Add some arguments to strengthen your refutation—your argument for having lots of acquaintances is pretty convincing, so you'll need to counter it more*." (SP5)
Organization	• "*Is the organization readily processed, aided by transitional elements that clearly express relationships between ideas?*" (SP19, Peer 3) • "*You should use a signalizing phrase to help reader be aware that this is your concluding paragraph. Pay attention to the transition from this paragraph to another*." (SP14 – Peer 1)	• "*This is a good body paragraph, but since the structure of the paper is designed to be 5 paragraphs, you should consider moving this up into your introduction or using parts of it in other paragraphs*." (SP7) • "*Overall the organization of the paper is highly effective and helps you succeed at presenting a highly emotion and descriptive argument*." (SP9)
Sentence Level	• *"More than one relationship means you need the plural 'relationships' "* (SP1 – Peer 1) • "*There's a difference between "few" and "a few". Few indicates negative and a few indicates positive. Which one do you mean?"* (SP6 – Peer 3)	• "*In standard English, and in most disciplines you write in, you will want to include the comma which separates the final two items in a series. For instance, "Loyalty, confidentiality, and unconditional support.*" (SP4) • "*Remember that a verb like "discussing" is a transitive verb meaning you actually perform that verb. In this case you are "discussing an unethical decision." You don't need to use "that" in between them*." (SP7)

In light of how the WHAT of comments were communicated and the importance of this HOW to the clarity of the comments, the HOW of the comments from these 20 essays was coded. The HOW included three categories: (1) Explanation, (2) Question, and (3) Correction (more details in Chapter 3, section 6 and also below, in the sections presenting each of these HOW categories). Table 25, Table 26, and Table 27 record the results of the HOW coding in relationship to the WHAT that the HOW communicated.

Table 25 shows the number of comments reflecting the HOW under the Idea Development category of the WHAT. As can be seen on Table 25 below, more than 67% of the student peers' comments and instructors' on Idea Development described or explained an aspect of what was written and/or included a "why" pointing the source of the problem. On five papers, SP5, SP7, SP9, SP13, SP16, both student peers and instructors provided explanations in all of their Idea Development comments. Many other papers had two peers with 100% of their Idea Development comments in the form of explanations, which were higher or quite close to the instructors', SP4, SP8, SP10, SP15, SP17, SP18, SP20. Regarding the Question category of the HOW, there were few Idea Development comments made by student peers and instructors that raised questions or issues for the writer to consider. Instructors raised questions about Idea Development issues on only three papers (SP10 – 1/5, SP13 – 2/4, SP20 – 1/6). On a few papers one or two peers raised questions while the others and the instructor had no questions; for example SP1 – peer 1 (0/4), peer 2 (1/4), peer 3 (1/4) and instructor (0/2); SP2 – peer 1 (1/2), peer 2 (0/1), peer 3 (1/3), and instructor (0/1); SP3 – peer 1 (0/4), peer 2 (0/4), peer 3 (3/5) and instructor (0/7); SP4 – peer 1 (0/1), peer 2 (0/0), peer 3 (1/2), and instructor (0/4); SP5 – peer 1 (1/5), peer 2 (1/1), peer 3 (0/2), and instructor (0/4); SP7 – peer 1 (0/1), peer 2 (0/4), peer 3 (2/4) and instructor (0/2); SP14 - peer 1 (0/2), peer 2 (1/4), peer 3 (2/2) and instructor (0/5). On SP6

and SP8, all of three peers raised questions about Idea Development issues, although the number of questions varied, whereas instructors had no questions on these two papers. In terms of the Correction category in Idea Development comments, there were 8 papers (SP2, SP4, SP6, SP7, SP8, SP9, SP10, SP12) on which instructors did not comment using Correction on Idea Development whereas one or two student peers did (although the number of corrections given by peers in idea development comment was low). On all of the other papers (except SP3 and SP14), instructors offered corrections in at least 80% of the comments on idea development while the majority of student peers did not provide correction in comments on idea development. On SP3 and SP18, all of the students peers suggested corrections in a few idea development comments, one of which was quite close to the instructor and the other of which was not [e.g. SP3: peer 1 (3/4), peer 2 (2/4), peer 3 (2/5), instructor (3/7); SP18: peer 1 (2/5), peer 2 (1/5), peer 3 (1/3), instructor (4/4)]. Other than providing explanations, questions, and corrections in comments on Idea Development, student peers and instructors expressed only praise or general reader responses when making comments on Idea Development on three papers (SP1, SP3, and SP18); for example, *"nice counter argument" (peer 1, SP1), "I like your counter argument*" (peer 3, SP1), "*Good topic sentence*" (peer 1 – SP3), "*well-said thesis statement"* (peer 2, SP3)*; "good explanation of the quote*" (peer 3, SP3); "*Good example*" (instructor – SP3).

Table 25.

Number of comments reflecting the HOW under the WHAT – Idea Development.

Student' papers (SP)	Number of comments reflecting the HOW under the WHAT - Idea Development												Odds*			
	Explanation				Question				Correction							
	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.
SP1	0	3/4	2/4	2/2	0	1/4	1/4	0	0	1/4	0	2/2	1/1	0	1/4	0
SP2	1/2	1/1	2/3	1/1	1/2	0	1/3	0	1/2	1/1	0	0	0	0	0	0
SP3	3/4	1/4	1/5	6/7	0	0	3/5	0	3/4	2/4	2/5	3/7	1/4	1/4	0	1/7
SP4	1/1	0	2/2	4/4	0	0	1/2	0	0	0	1/2	0	0	0	0	0
SP5	5/5	1/1	2/2	4/4	1/5	1/1	0	0	0	0	0	3/4	0	0	0	0
SP6	4/5	2/3	4/6	1/1	1/5	1/3	2/6	0	1/5	0	2/6	0	0	0	0	0
SP7	1/1	4/4	4/4	2/2	0	0	2/4	0	1/1	0	0	0	0	0	0	0
SP8	4/5	3/3	5/5	4/4	1/5	3/3	2/5	0	2/5	0	1/5	0	0	0	0	0
SP9	2/2	2/2	2/2	6/6	0	0	0	0	0	1/2	0	0	0	0	0	0
SP10	2/3	4/4	4/4	5/5	1/3	0	0	1/5	0	0	1/4	0	0	0	0	0
SP11	3/3	3/3	2/2	1/1	0	0	0	0	1/3	0	0	1/1	0	0	0	0
SP12	0	0	1/1	0	0	0	0	0	0	0	0	0	0	0	0	0
SP13	4/4	3/3	3/3	4/4	0	0	0	2/4	0	0	0	4/4	0	0	0	0
SP14	2/2	4/4	0	4/5	0	1/4	2/2	0	0	1/4	0	1/5	0	0	0	0
SP15	3/3	3/3	3/4	4/4	0	0	0	0	3/3	0	0	3/4	0	0	0	0
SP16	1/1	5/5	4/4	4/4	1/1	0	0	0	0	0	0	4/4	0	0	0	0
SP17	0	2/2	2/2	4/4	0	1/2	0	0	0	0	0	3/4	0	0	0	0
SP18	4/5	5/5	3/3	4/4	0	1/5	0	0	2/5	1/5	1/3	4/4	1/5	0	0	0
SP19	0	3/3	4/5	4/4	0	0	0	0	0	1/3	1/5	3/4	0	0	0	0
SP20	3/3	2/3	4/4	5/6	0	0	0	1/6	0	1/3	0	4/6	0	0	0	0

**: comments that indicate praises or general reader responses or general/vague points*

Table 26 shows comments reflecting the HOW under the WHAT category, Organization. According to Table 26 below, almost all of the comments on Organization made by student peers and instructors were provided as explanations. Nine of peers' comments (in 9 papers SP1, SP2, SP3, SP5, SP6, SP9, SP14, SP16, SP19) raised questions about organization such as SP1 (peer 3, 1/2), SP2 (peer 2, 1/2), SP3 (peer 2, 1/1), and SP5 (peer 3, 1/1). Instructors did not raise questions about Organization; there were only two comments by instructors (one in SP5 and the other in SP6) that began with a question followed by an explanation of the problem (in these instances the comments were categorized as both questions and explanations). The majority of comments on Organization by student peers lacked suggestions for correction; only SP1, SP2, SP14, and SP20 had a few peers' comments with suggested corrections combined with explanations (e.g. SP1: peer 1 (1/1), peer 2 (1/1), peer 3 (1/2); SP2: peer 1 (0), peer 2 (2/2), peer 3 (1/1); SP14: peer 1 (2/2)). On 13 papers that had instructors' comments on organization (22 comments in total), instructors suggested corrections in 11 comments on organization on 8 papers, and most of the corrections were followed or preceded by explanations (in these instances the comments were categorized as both corrections and explanations). Instructors had one comment on organization that express a praise only (e.g. "*well organized essay*"), and one peer had such a comment (e.g. "*well-done organization*" – SP11 – peers 3).

Table 26

Number of comments reflecting the HOW under the WHAT - Organization

Student' papers (SP)	Number of comments reflecting the HOW under the WHAT - Organization												Odds*			
	Explanation				Question				Correction							
	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.
SP1	1/1	0	2/2	2/2	0	0	1/2	0	1/1	1/1	1/2	2/2	0	0	0	0
SP2	0	1/2	1/1	0	0	1/2	0	0	0	2/2	1/1	0	0	0	0	0
SP3	1/1	0	1/1	0	0	1/1	0	0	0	0	0	0	0	0	0	1/1
SP4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SP5	2/2	1/1	1/1	1/1	0	0	1/1	1/1	0	0	0	0	0	0	0	0
SP6	1/1	3/3	1/1	1/1	1/1	0	0	1/1	0	0	0	1/1	0	0	0	0
SP7	1/1	2/2	2/2	4/4	0	0	0	0	0	0	0	0	0	0	0	0
SP8	1/1	1/1	2/2	1/1	0	0	0	0	0	0	0	0	0	0	0	0
SP9	0	1/1	1/1	1/1	1/1	0	0	0	0	0	0	0	0	0	0	0
SP10	1/1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SP11	0	0	0	1/1	0	0	0	0	0	0	0	1/1	0	0	1/1	0
SP12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SP13	0	0	0	1/1	0	0	0	0	0	0	0	0	0	0	0	0
SP14	2/2	1/1	1/1	0	0	1/1	0	0	2/2	0	0	0	0	0	0	0
SP15	0	0	0	2/2	0	0	0	0	0	0	0	2/2	0	0	0	0
SP16	0	0	1/1	2/2	1/1	0	0	0	0	0	0	1/2	0	0	0	0
SP17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SP18	1/1	0	0	1/1	0	0	0	0	0	0	0	1/1	0	0	0	0
SP19	2/2	1/1	1/2	2/2	0	0	1/2	0	0	0	0	1/2	0	0	0	0
SP20	0	0	0	2/2	0	0	0	0	0	1/1	0	2/2	0	0	0	0

**: comments that indicate praises or general reader responses or general/vague points*

Table 27 shows comments reflecting the HOW under Sentence Level of the WHAT. According to Table 27 below, more than 70% of comments on Sentence Level by student peers and instructors offered corrections in the comments (except Peer 2 -SP1, Peer 2 and Peer 3 -SP2, and the instructor -SP8). In many papers, instructors offered corrections to 100% of their comments on the sentence level, which can also be seen in many student peers' comments (as in SP1 (peer 3), SP4 (peer 2), SP6 (peer 2, peer 3), SP7 (peer 1, peer 3), SP9 (peer 3), SP10 (peer 2), SP12 (peer 2), SP13 (peer 3), SP15 (peer 2), SP20 (peer 2). Very few student peers raised questions in their comments on sentence level issues, and when they did, there were one or two comments that included questions. For example, peer 1, peer 2, and peer 3 had 1 sentence level comment with a question on SP3; peer 1 and peer 3 had 1 comment on SP4; SP14 had 1 sentence level comment from peer 3; SP16 had 1 sentence level comment from peer 2; and SP18 had 1 sentence level comment from peer 1. However, peer 2 (SP1) and peer 3 (SP2) each had 3 sentence level comments with questions raised. The instructor had only 1 Sentence Level comment raising a question (SP14). As also seen on Table 27, many student peers had the same or close percentages of comments on the sentence level on the same papers that described or explained the problems, varying from 35% to 50% of the total sentence level comments. Instructors provided more explanations to sentence level issues than student peers, ranging from 40%-80% of the total sentence level comments (except 10% - 2/20 - on SP1). Some peers gave no explanations or very few explanations in sentence level comments, compared with the other peers and the instructors on the same papers. For example in SP7 peer 3 had 0/10 sentence level comments that explained while peer 1 had 2/7, peer 2 had 4/14, and the instructor had 3/3; in SP5 peer 3 had 1/8 of sentence level comments that explained while peer 1 had 8/13, peer 2 had 4/11 and the instructor had 6/14; in SP15 peer 1 had 1/13 whereas peer 4 had 4/10, peer 3 had 9/17,

and the instructor had 8/10. Overall, as seen on Table 27, student peers and instructors provided explanations and corrections and raising questions in sentence level comments. However, in several instances, while instructors had specific comments on sentence level issues on all of the papers, 7 peers had a few of comments on sentence level issues that were too general and vague, for example, SP1 (peer 1 – 3/19 and peer 2 – 1/12), SP2 (peer 1 – 2/12), SP5 (peer 1 – 1/13), SP8 (peer 1 – 1/7), SP11 (peer 3 – 3/8), and SP14 (peer 2 -1/10). The following are some examples of such comments:

"tense": *SP14 – peer 2*
"awkward word choice": *SP 1 – peer 1;*
"punctuation": *SP 1 – peer 2*
"vocabulary", "grammar": *SP2 - peer 1*
"*awkward language use*": SP5 – peer 1
"*bad grammar*": SP8 – peer 1
"*word choice*" "t*ense*" "*punctuation*" : SP11 peer 3

Table 27

Number of comments reflecting the HOW under the WHAT – Sentence Level

Student' papers (SP)	Number of comments reflecting the HOW under the WHAT – Sentence Level												Odds*			
	Explanation				Question				Correction							
	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.	P1	P2	P3	Instr.
SP1	7/19	6/12	4/6	2/20	0	3/12	1/6	0	16/19	4/12	6/6	20/20	3/19	1/12	0	0
SP2	5/12	4/9	5/11	14/24	0	2/9	3/11	0	9/12	5/9	5/11	24/24	2/12	0	0	0
SP3	3/5	4/11	3/6	6/15	1/5	1/11	1/6	0	4/5	10/11	5/6	15/15	0	0	0	0
SP4	5/16	4/12	3/11	4/4	1/16	0	1/11	0	13/16	12/12	9/11	3/4	0	0	0	0
SP5	8/13	4/11	1/8	6/14	0	0	0	0	11/13	7/11	7/8	8/14	1/13	0	0	0
SP6	4/11	2/6	3/4	4/9	0	0	1/4	0	10/11	7/7	4/4	8/9	0	0	0	0
SP7	2/7	4/14	0	3/3	0	0	0	0	7/7	11/14	10/10	3/3	0	0	0	0
SP8	3/7	3/10	2/12	4/4	0	0	0	0	6/7	8/10	11/12	2/4	1/7	0	0	0
SP9	8/12	4/8	0	2/3	0	0	0	0	12/12	6/8	3/3	3/3	0	0	0	0
SP10	3/7	5/9	4/9	1/1	0	0	0	0	5/7	9/9	7/9	1/1	0	0	0	0
SP11	7/14	4/9	2/8	6/14	0	0	0	0	12/14	8/9	6/8	14/14	0	0	3/8	0
SP12	9/12	7/10	7/11	8/18	0	0	0	0	11/12	10/10	10/11	18/18	0	0	0	0
SP13	6/18	3/10	5/6	6/7	0	0	0	0	14/18	9/10	6/6	7/7	0	0	0	0
SP14	1/2	4/10	7/11	2/2	0	0	1/11	1/2	2/2	8/10	9/11	1/2	0	1/10	0	0
SP15	1/13	4/10	9/17	8/10	0	0	0	0	12/13	10/10	12/17	10/10	0	0	0	0
SP16	3/8	5/17	5/13	5/5	0	1/17	0	0	7/8	15/17	11/13	5/5	0	0	0	0
SP17	2/12	4/13	4/8	9/11	0	0	0	0	11/12	9/13	7/8	11/11	0	0	0	0
SP18	6/10	3/9	7/9	4/5	1/10	0	0	0	8/10	7/9	9/9	4/5	0	0	0	0
SP19	2/9	7/7	7/10	7/7	0	0	0	0	7/9	7/7	8/10	6/7	0	0	0	0
SP20	8/10	4/6	6/13	8/10	0	0	0	0	9/10	6/6	9/13	10/10	0	0	0	0

**: comments that indicate praises or general reader responses or general/vague points*

In general, in terms of the HOW, among comments on Idea Development on all of the 20 students' papers in the selective sample, student peers' comments (173 comments on Idea Development) included 148 explanations (85.5%), 29 questions (16.8%), 32 corrections (18.5%), and 5 odds (2.9%) such as praise, personal opinions, and general/vague things. All of instructors' comments on Idea Development (72 comments) consisted of 69 explanations (95.8%), 4 questions (5.6%), 34 corrections (47.2%), and 1.4% odds. For comments on organization, student peers had a total of 46 comments on organization, which broke down into the following categories: 38 explanations (82.6%), 9 questions (19.6%), 9 corrections (19.6%), and 1 odd (2.2%). Instructors made a total of 22 comments on organization, in which there were 21 explanations (95.5%), 2 questions (9.1%), 11 corrections (50%) and 1 odd (4.5%). Regarding comments on sentence level, student peers made a total of 605 comments, in which 262 comments offered explanations (43.3%), 18 comments raised questions (3%), 508 comments suggested corrections (84%), and 12 comments were general/vague (2%). Instructors made a total of 180 comments on sentence level issues, which consisted of 109 explanations (60.6%), 1 question (0.6%), 173 corrections (96.1%) and no odds (no comments that indicate praises or general reader responses or general/vague things). Figure 9 below shows the percentages of the HOW reflected under the WHAT through comments made by student peers and instructors, and Table 28 below illustrates a few examples of comments described as Explanations, Questions, and Corrections.

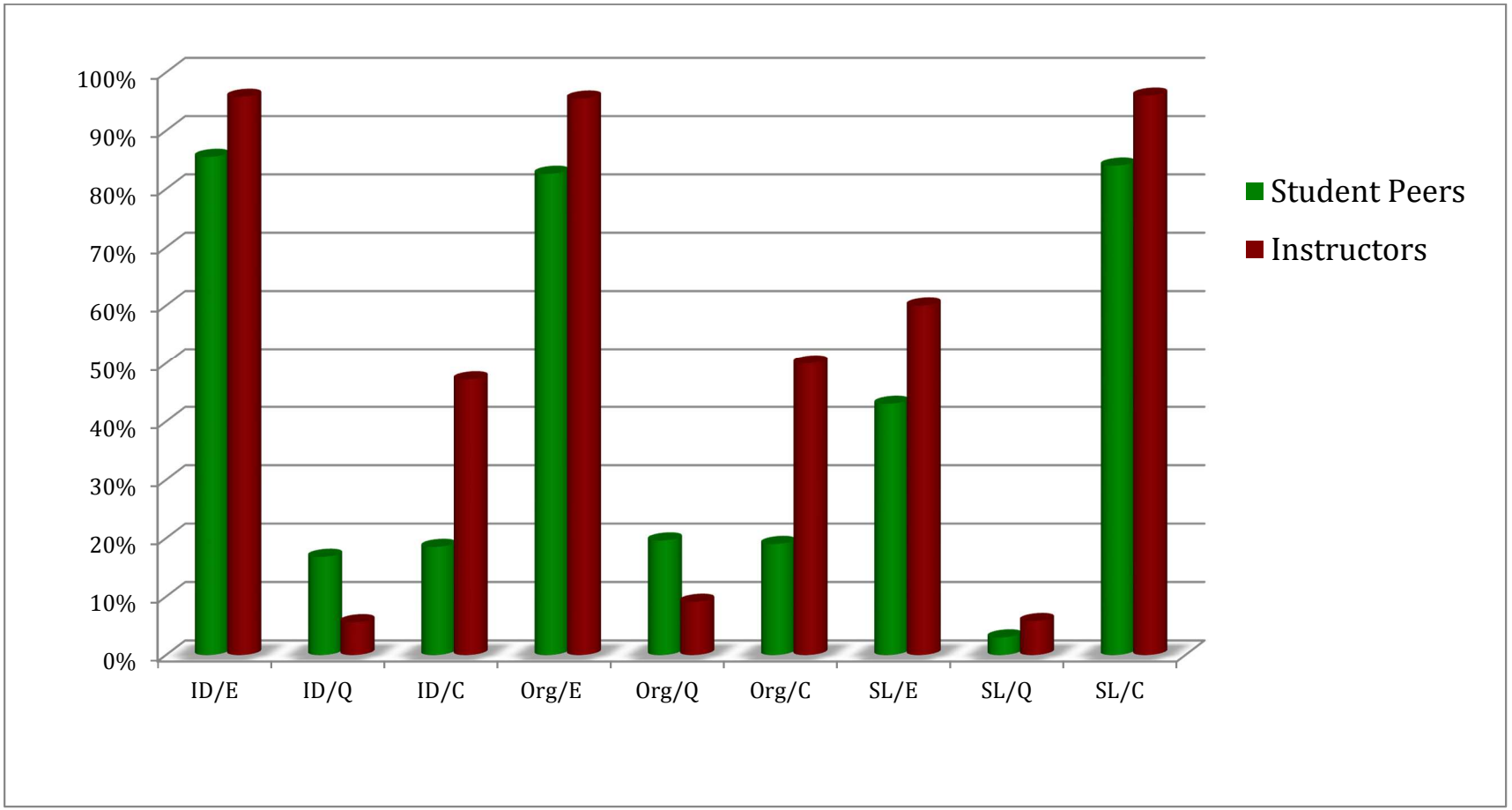

Figure 9. *HOW criteria displayed through comments made by student peers and instructors*

Notes:
Idea Development (ID)
Organization (Org)
Sentence Level (SL)
Explanation (E)
Question (Q)
Correction (C)

Table 28.

Example comments made by student peers and instructors described in the HOW

The "HOW"	Comments by student peers	Comments by instructors
Explanation	*The counter argument paragraph acknowledges the opposing view but doesn't summarize points*." (SP7 – Peer 2)	"*You have a nice hook to open your introduction and you do a good job of stating your thesis*." (SP4)
Question	"*Is the organization readily processed, aided by transitional elements that clearly express relationships between ideas?*" (SP19, Peer 3)	"*You make some very convincing points, though for the assignment are you to select one of the choices and argue for it?"* (SP10)
Correction	"*Use a comma after "especially"*" (SP6 – Peer 2)	"*Remember to maintain tense agreement in sentence construction. For example, "remember that you will take what you gave" is a disconnect between present and past."* (SP8)

From the analysis of all of the comments made by student peers, key words that connected with words and emphasis in the Writing Assignment #3 rubric and the assignment guidelines (see Appendix A and B) were looked into. This helped bring out patterns that revealed student peers were referring to specifics within the rubric, the assignment guidelines, and instructions in the peer assessment training. Table 29 below provides students' peer comments in 20 students' papers that were relevant to terms/phrases in the rubric and the assignment guidelines.

Table 29.

Peers' comments using key words/terms in the rubric and the assignment guidelines

Students' papers (SP)		**Peers' comments**	**Relevance to terms/phrases in Rubric (R) and Assignment Guidelines (AG)**
SP1	*P1*	--	
	P2	• This is a good **example** to back up your **claim** in the previous sentence	• AG
	P3	• Your **thesis statement** looks **strong** when you make a clear point to your **stand** • How could you know human can't connect at deep lever? Do you give any research showing this? • You need to join this paragraph to the previous one and use a linking adverb to connect the ideas in the previous paragraph because this paragraph and the previous on seem to talk about the same argument	• *R & AG* • *AG* • *R*
SP2	*P1*	--	•
	P2	• This **thesis statement** looks good, but you should spell out "the second idea" • If you could add the **counter argument** before jumping into the conclusion, your essay would be much better.	• *R & AG* • *R & AG*
	P3	• The **thesis statement** is not **clear** enough. • Nice **examples**. However, I think that the flow of ideas is not good enough. • The overall essay does not look like an argumentative essay. You need to discuss somewhere the opposite side – I mean you need a **counter argument**. • Are you contrasting yourself? The way you **support** your **main points** seems opposite. Do you mean "quality" here?	• *R* • *R & AG* • *R & AG* • *R & AG*
SP3	*P1*	• You need to add more **examples** to **support** the main idea in this paragraph. • I like this quote. However, you need to elaborate this quote to **support** your ideas. • This is a very nice **introduction**. You have a **strong thesis**.	• *R & AG* • *AG* • *R*
	P2	• Is this your **counter argument** paragraph of the argumentative essay? Then, more **support** for the counter argument is needed.	• *R & AG*
	P3	• Do you really develop this idea in the body paragraphs? • Good friendships with a lot of people or with one ore two friend? Expand the ideas and make the ideas more specific. • This essay has a good **organization** with strong supporting paragraphs. However, you need a stronger **counter argument** paragraph before jumping into the conclusion.	• *R & AG* • *AG* • *R & AG*
SP4	*P1*	• Your **introduction offers a context and a clear thesis statement.** • **Precise vocabulary usage** with **appropriate register**	• *R* • *R*
	P2	• --	

	P3	• Your **main points** are **present** but they **lack details**. You can give some **examples** to **support** the ideas.	• *R & AG*
SP5	*P1*	• **Conclusion** needs to cover the **arguments** presented above with some short words!	• *R & AG*
	P2	• You have a few **surface grammatical errors**, but these do not **impede the understanding**. • This is a good paragraph. Your **claims** are **well supported** with **logical reasoning and relevant evidence**.	• *R* • *R*
	P3	• Good **reasoning and relevance evidence** to **support** the **main points**. • Does the **transitional element** use **connect the various components of the essay** and **clearly express relationships between ideas**.? If here you use "First", the next paragraph is supposed the follow-up of this argument. However, it's not. The next should be the contrast. Suggested word: "On the one hand,"	• *R* • *R*
SP6	*P1*	• The **thesis** is **clear** enough to state **your stand**. • The way you **support** the **main point** in this paragraph is fine, but if you want to **support** your essay **thesis**, I think you are off the **topic**. • Did you try to follow the **organization** format because the essay lacks the **counter argument** paragraph?	• *R & AG* • *R & AG* • *R & AG*
	P2	• You made a few main **claims** in this same paragraph but you failed to **support the claims sufficiently**. • Again, too many transition words in such a short paragraph. • This paragraph does not **support** the **position** you stated in the **introduction**, though **comprehensible**. • There should be 5 paragraphs with the 4th one presenting **opposition**	• *R* • *R* • *R & AG* • *AG*
	P3	• Add the "because" clause to this sentence to make the **thesis statement** more **explicit** and stronger. • **Claims** are **well supported with relevant evidence**, but supporting ideas are not **sufficient** enough. • This would be a good paragraph with **well supported ideas** if the focus was about your preference of having one or two friends. • Prior the conclusion, you should have **refutation** one or two paragraphs that **acknowledge the opposing view**. An argumentative essay should have counter arguments.	• *R* • *R & AG* • *R & AG* • *AG*
SP7	*P1*	• The main ideas **lack detailed support** which **may include personal experience, statistics, examples, facts, or experts' opinions.** • You have good use of **transition element** to enhance the comprehension and **logics** of the essay.	• *R & AG* • *R*
	P2	• The **counter argument** paragraph **acknowledges the opposing view** but doesn't **summarize points**. • Pay attention to the use of **vocabulary** to make it more **appropriate**.	• *R & AG* • *R*
	P3	• **Clear progression of ideas, aided by sufficient transitional elements**	• *R*
SP8	*P1*	• I like this introduction – it **provides a strong opening and a clear thesis.** • You used a variety of sentences and vocabulary.	• *R & AG* • *R*
	P2	• --	•
	P3	• There's some wrong **use of English writing conventions** in this essay – pay attention to **indentation** and **punctuation**	• *R*
SP9	*P1*	• --	•
	P2	• --	•
	P3	• --	•
	P1	• Do you think your reader would **be able to read only the introduction and know your stand**?	• *R & AG*

SP10	*P2*	• Using **transitional elements** is good for a smooth **organization**, but you used too many transition words in a short paragraph.	• *R & AG*
	P3	• --	•
SP11	*P1*	• --	•
	P2	• Good **conclusion** with **accurate summary of main points of the argument**. • There are a few grammatical, and **usage errors** in the introduction such as tenses and fragments.	• *R & AG* • *R*
	P3	• You **successfully introduced the essay with a clear thesis statement**.	• *R*
SP12	*P1*	• --	•
	P2	• --	•
	P3	• Try to use a variety of sentence structures and avoid **wordy** sentences. • Your **word choice** seems not **appropriate** in this context	• *R* • *R*
SP13	*P1*	• There are a few **stigmatizing errors that impede understanding** in the second and third paragraph – run on and fragments. • This paragraph is not **well supported** with **relevance evidence**.	• *R* • *R & AG*
	P2	• **Claims** are not **supported** with **relevant evidence** • Your **refutation paragraph** is **vague**.	• *R & AG* • *R*
	P3	• The **conclusion** should **summarize** major **points of the argument**	• *R & AG*
SP14	*P1*	• If you have **specific examples from life** to **support** the **main point** in this paragraph, it will be more convincing. • Need a transition to make the **organization** look more **connected**. • You should you a signalizing phrase to help reader be aware that this is your concluding paragraph. Pay attention to the **transition** from this paragraph to another.	• *R & AG* • *R & AG* • *R & AG*
	P2	• You have a good hook here – **explicit thesis statement**. • What life evidence? You should make it clearer by giving one or two **examples**. • Is this your conclusion? Do you think it's harder to predict what is going on with the essay **organization** because of the lack of **transition** between paragraphs?	• *R & AG* • *AG* • *R & AG*
	P3	• What option? Do you think what is written here is not clear enough? • Is the **language use** effective enough to convey the ideas clearly? • Did you **support** this **claim** enough? • There's a lack of **transition** between paragraphs, so this makes the essay **organization** not coherent.	• *R & AG* • *R* • *R & AG* • *R & AG*
SP15	*P1*	• You could make the **thesis statement** stronger and more **explicit** by stating your position and adding the reason	• *R & AG*
	P2	• --	•
	P3	• --	•
SP16	*P1*	• It's a little bit difficult to follow what you are saying here. Using some **transition elements** would help **connect the ideas.**	• *R*
	P2	• Not really **relevant evidence**s to **support** the **claims** • Problems on using punctuations are repeated. These are not accepted in formal writing	• *R* • *R*
	P3	• Claims on this issue are somewhat unclear and off the focus.	• *R & AG*

SP17	*P1*	• --	•
	P2	• The introduction is not engaging, and the **thesis statement** is not **strong** and **clear** enough.	• *R & AG*
	P3	• This sentence is **wordy** although the **vocabulary** use is **sophisticated**.	• *R*
SP18	*P1*	• This is a fragment, which makes it hard to understand the conveyed ideas. • A problems with **English writing convention** happens in this paragraph: punctuation	• *R* • *R*
	P2	• You need to support your **counter argument** with more **relevant evidences**. • The **range of vocabulary** is **limited**. You need to you more **sophisticated vocabulary** to improve the quality of the essay.	• *R* • *R*
	P3	• The **organization** among the body paragraphs was not well **connected** due to the lack of **transitional elements.**	• *R & AG*
SP19	*P1*	• The **relationship between ideas** is not **clearly expressed** because it is not **aided by transitional elements.**	• *R*
	P2	• The **organization** of the ideas is problematic because of the **insufficient use of transitional elements**. • Poor **idea development** – you should use **specific examples or evidences** to **support** your **claims**.	• *R* • *R & AG*
	P3	• Is the **organization readily processed**, **aided by transitional elements that clearly express relationships between ideas?**	• *R*
SP20	*P1*	• Good job in supporting your **claims** with **logical reasoning and relevant evidence** • Fix errors in spacing between words and punctuations to not distract the reader. • Revise this part to avoid wordiness.	• *R* • *R* • *R*
	P2	• You showed a good understanding of the topic, **aided by relevant** supporting ideas. • Good **conclusion** with **sufficient summaries** of **main points** discussed in the body paragraphs	• *R* • *R & AG*
	P3	• Good **examples** and **details** to **support** what is being said • I like it when you use a great **range of vocabulary**	• *R & AG* • *R*

As can be seen in Table 29 above, through 20 students' papers graded by 60 student peers, 46 student peers (approximately 76.7%) used terms/phrases in the rubric and the assignment guidelines or used wording in some of their comments (89 comments). The bold words/phrases in the comments shown in Table 29 were precise words/phrases taken from the rubric and the assignment guidelines, for example "**Clear progression of ideas, aided by sufficient transitional elements**" (Peer 3, SP7), "**Precise vocabulary usage** with **appropriate register"** (Peer 1, SP4), and "You have a few **surface grammatical errors**, but these do not **impede the understanding**" (Peer 2, SP5). In some comments, although student peers didn't have the precise words from the rubric and the guidelines, they used wording that in essence connects the comment to the rubric and the guidelines. For example, the comment by Peer 3 (SP1) "*You need to join this paragraph to the previous one and use a linking adverb to connect the ideas in the previous paragraph because this paragraph and the previous on seem to talk about the same argument*" seems related to the description of Organization criterion in the rubric "Organization is readily processed, aided by transitional elements that connect the various components of the essay and clearly express relationships between ideas." Through 20 papers, 39 student peers' comments out of 89 comments (43.8%) included precise terms/phrases in the rubric and guidelines, and 56.2% of the comments had the wording that in essence connects the comment to the rubric and the guidelines.

The number of student peers' comments using the precise terms/phrases in the rubric and assignment guidelines or using the wording connected to the rubric and the guidelines varied among peers and papers. For instance, on SP2, peer 1 had no such comment, peer 2 had two, and peer 3 had four. SP9 did not have any comments that peers used precise terms/phrases or used wording whereas SP6 had 11 comments from three peers (peer 1 had 3, peer 2 had 4, and peer 3

had 4). Also as shown in Table 29 above, 13 student peers' comments out of 89 comments (14.6%) were about sentence level issues whereas 85.4% of the comments addressed global issues of idea development and organization.

In summary, this whole Results chapter describes the results of the data analyses targeted to answer the three research questions of the study. Section 1 of the chapter results from the survey and interviews targeted to answer Research Question 1. Section 2 provides results from the comparison of peer and instructor grading of student essay drafts for answering Research Question 2. Section 3 shows results from the coding of peer and instructor comments on the student essay drafts to answer Research Question 3. The following chapter – Chapter 5 - presents major conclusions and discussions withdrawn from the findings of the study.

CHAPTER 5

CONCLUSIONS AND DISCUSSIONS

As composition MOOCs have begun to be offered in more settings and peer assessment has been looked to as a means of supplying feedback and grading for class sizes impractical for instructors alone to handle, questions about the quality of the use of peer assessment have arisen. This study strove to investigate students' perceptions of peer assessment used in a MOOC-based composition course provided by E-Center for Professional Development and to identify if there was a significant difference between grades and comments given by student peers and those given by the instructors. The following sections summarize the results of the study and provide discussions related to the findings. This chapter also presents the limitations of the study, recommendations for future research, and contributions of the study. At points, the limitations will be integrated into the section discussing the conclusions and summarized later. The decision to follow this order was made for several reasons: first, because the study's two largest sets of data—the survey data and the grading data—were collected anonymously, thus unfortunately limiting interesting triangulations for interpreting the three sets of data (the two just mentioned and the peer comments on peers' drafts) and, second, because the number of interesting questions about these data sets interrelationships unable to be triangulated relate closely to understanding the conclusions that the study was able to draw from the single data sets alone.

1. Student Perceptions of Peer Assessment in a Composition MOOC

Regarding students' perception of peer assessment in a Composition MOOC, the empirical findings of this study provided evidence that peer assessment was well received by the majority of students in the ESL/EFL Writing for Academic and Professional Success MOOC-based Course. Specifically, 85.6% of the participants took part in peer assessment training and referred to the training instructions and rubric when commenting on and grading their peers' papers (survey question #5). Approximately 88% of the participants either agreed or strongly agreed that they followed the guidelines provided during training for commenting on and grading peers' writing and followed the rubrics for grading peers' writing (survey questions # 9 and #10). Approximately 62.8% of the participants believed the feedback they gave their peers on peers' writing was useful (survey question #7), and the majority of the participants (76.7%) believed that their feedback on peers' writing was thorough and constructive (survey question #11). Approximately 89% of the survey participants cited that they used peers' feedback when they revised their writing (survey question #17). Most of the survey particpants (96%) perceived that peer assessemnt helped improve their writing performance (survey question #16). Regarding what students perceived toward peers' qualifications for grading and providing feedback, 66.3% of the participants agreed or strongly agreed that peers were qualified in giving feedback and grades.

Overall, the positive attitudes and perceptions of students towards peer assessment as observed in this study confirmed previous studies in that peer assessment was perceived as helpful in improving students' performance. However, the findings from these studies addressed student perceptions in different disciplines: Simkin and Ramarapu (1997) looked at peer-reviewed grading in computer science courses; Nelson and Carson (1998) and Mangelsdorf's

(1992) investigated students' perceptions of peer review in ESL composition classes; Lou et al. (2014) investigated students' perceptions of peer assessment in a non-composition MOOC. Given the lack of research in student perceptions of peer assessment in composition MOOCs, this present study added to the existing literature the fact that students had positive perceptions towards online peer assessment in a composition MOOC context. In addition, the study revealed an interesting finding in students' perception of peer assessment – that is, students' attitudes and perception towards peer assessment might have changed during the course of the class. For example, interview responses on the qualifications needed for peer assessment indicates an instance in which respondents disagreed on feeling qualified might have been closer to saying they agreed later in the course after training and experience with peer assessment. An interview respondent stated "*To tell the truth, I have never given feedback to peers before, neither comments nor grades, and I did not really feel confident in the peer assessment activity at all. However, after I watched posted videos on how to do the peer assessment, followed instructions on peer assessment and tried grading and commenting on two writing samples, I felt just a little bit more confident in doing the job, but I don't think I was well prepared"*(I – 4). Such perceptions, though from a minority of respondents, suggests that students' positive attitudes and perceptions towards peer assessment in the context of a Composition MOOC might be affected by the training on peer assessment and the experience with peer assessment.

Despite the positive attitudes and perceptions towards online peer assessment reflected in the survey results, a portion of the students in the ESL/EFL Writing for Academic and Professional Success MOOC-based Course also expressed some negative feelings about this activity. Approximately 12.7% of the survey participants perceived that the feedback they got from peers was negative and critical (survey question # 13), and 38.5% of the survey participants

believed that peers did not give fair grades on their writing (survey question #15). Around 61.5% of the survey participants felt that they themselves were not qualified to give feedback and grades on peers' writing (survey question # 18), and 34.7% felt that peers were not qualified in assessing others' work (survey question #19). A few numbers of students (at least 3%) claimed that peer assessment did not help improve their writing at all (survey question # 16). Seventy three percent (73%) of the survey participants found peers' comments and grades were different and not consitent (survey question #14). Generally, these findings on the students' negative attitudes towards peer assessemnt in this study were similar to what Kaufman and Schunn (2011) found in that students sometimes thought peer assessment unfair and that students believed they were not qualified to grade peers' work. However, Kaufman and Schunn's study focused on a small number of students' whereas this study sought to test out students' perception of online peer assessment in an as yet unexplored context of MOOCs with their massive enrollment. Plausible interpretations for negative perceptions of the qualifications of peers for assessment might include (1) students are not perceived as qualified because they are not teachers, (2) students may not receive sufficient training on peer assessment, (3) students may not have quality practice in peer assessment necessary to providing quality feedback and grading. While this study was not designed to produce data capable of determining which interpretations might be the correct one(s), the data that was collected suggest questions raised by student perceptions on their qualifications as peer responders should be considered in determining the training peer assessors are to receive and also should be pursued in future research, as I'll note later in this chapter.

While findings in other studies in the literature (i.e. non-composition scholarship, non-composition MOOC scholarship, non-MOOC composition scholarship) supported the usefulness

of peer assessment in helping students improve their writing performance (e.g. Simkin & Ramarapu ,1997; Nelson & Carson,1998; Mangelsdorf, 1992; Brammer & Rees, 2007; Lou et al., 2014), one finding in this study is different in that a few students (approxiamtely 3%) claimed that peer assessment did not help improve their writing at all. This percentage of 3% seems small, but with the large scale of MOOC student population – thousands of students - the same 3% would include a good many students.

The students also expressed difficulties and challenges that they had with the online peer assessemnt used in the ESL/EFL Writing for Academic and Professional Success MOOC-based Course. Responses to the open-ended survey question (20a) show that as both commentators and graders, the majority of the students spoke about their lack of confidence, lack of experience, and lack of qualification. Others' difficulties and challenges included technical issues (27%), peer assessment tools i.e. the rubrics (15%), the anonymity of peer assessment (13%), the types of writing assignments (9%), and other difficulties and challenges like one's English competences (5%), emotional factors (4%), different rhetorical strategies (3%) and so forth. As a writer being commented on and graded by their peers, 68% of the students said that they had difficulty interpreting and understanding the comments given by peers. Explaining reasons for their difficulties/challenges, students blamed for the comment and grade contradictions (34%), fairness (23%), peers' qualification (19%), the anonymity of peer assessment (15%) and others (10%) like trust, conflicting comments, cultures. Difficulties and challenges students in the ESL/EFL Writing for Academic and Professional Success MOOC-based Course were found similar to what Watters (2012) noted about peer assessment in non-composition MOOCs, for example the anonymity of feedback, the lack of feedback on feedback. What this study has added to the existing literature on peer assessment is findings that speak in greater detail than

previously to the difficulties and challenges students face in peer assessment in a composition MOOC and from the perspectives of both writers commenting on and grading other writers and writers being commented on and graded.

2. Students' Grades Compared to Instructors' Grades

The analysis of grades given by student peers reveals that there was consistency among student peers' grades on quite many student papers. For Writing assignment 1, in a total of 45 out of 200 student papers – 22.5% (both Draft 1 and Final Draft), all of the three peers gave grades within a certain same letter grade; Writing Assignment 2 had 29 student papers (14.5%); and Writing Assignment 3 had 30 student papers (15%). In addition, 134 Writing Assignment 1's student papers (72%) had two student peers giving the same grade while the other gave different grade range (either lower or higher); Writing Assignment 2 had 124 student papers (62%); and Writing Assignment 3 had 137 student papers (68.5%). A level of consistency among student peers' grades was also found in a perception of consistency in certain aspects of peers' assessment registered in the survey's findings. Approximately 19% of the survey participants responded that there were similarities among peers' comments and peers' grades (survey question #20d). The combination in this survey question of grades and comments might cause some loss in interpreting the similarities among peers' grades separate from peers' comments; however, it was evident from the survey that students perceived the similarities among peers' grades: one survey participant responded "*"The grades I got from peers were quite close, so I think peers' grades were similar" (Female #714, China).* Generally, the findings of this present study on grades given by student peers were in certain respects similar to previous studies on

peer grading in non-composition, non MOOC contexts (e.g. Cho et al., 2006; Sadler & Good, 2006; Bouzidi & Jaillet, 2009) as well as non-composition MOOC contexts (e.g. Lou et al., 2014) in that grades given by student peers were found to be fairly consistent. Since there has been, however, a lack of research on consistency in peer grading in composition MOOCs, these findings, though limited to the context of this study have added to the existing literature the point that there is consistency among peers' grades through peer assessment in a composition MOOC context. Although the study was not designed to produce data capable of determining what factors might help enable this consistency, plausible interpretations for the consistency among student grades might include (1) students having received some training on peer assessment so they might have known what to do and how to do so when grading a paper, (2) students having used the same rubric with the same grading criteria for assessment.

In addition to a degree of consistency among student peers' grades worth noting, this study found that student peers also assigned grades within a certain range that differed from the range of the grades given by the other peers when they graded the same papers. Writing Assignment 1 had 21 student papers (10.5%) with three peers giving grades in three different grade ranges; Writing Assignment 2 had 47 student papers (23.5%); and Writing Assignment 3 had 31 student papers (15.5%). This inconsistency partly supports the findings from the survey: approximately 73% of the survey participants responded that there were differences among peers' grades and peers' comments (survey question #20d). (Again, the combination of two aspects of peer assessment –peers' grades and peers' comments in one survey question—limits the conclusion). Possible interpretations for the inconsistency among peers' grades will be discussed later in this section, together with possible interpretations for the inconsistency between peers' grades and instructors' grades.

As to the comparison between grades given by student peers and grades given by instructors on the same papers, Writing Assignment 1 had 20 student papers (10%) in which all three peers' grades were in the same letter grade with the instructors' grades. Writing Assignment 2 had 14 student papers (7%); and Writing Assignment 3 had 18 student papers (9%). Plus, Writing Assignment 1 had 52 student papers (26%) in which two peers gave the same grade with the instructors' while the other peer gave grades in the letter grade that differed from the instructors' (either lower or higher); Writing Assignment 2 had 55 student papers (27.5%); and Writing Assignment 3 had 67 student papers (33.5%). In addition, there were 99 Writing Assignment 1's student papers (44.5%), 104 Writing Assignment 2's student papers (52%), and 84 Writing Assignment 3's student papers (42%), which all had one peer whose letter grade given was the same to the instructor's and two peers whose letter grade range given were either higher or lower the instructors' letter grade.

From the analysis of the letter grades through out student drafts' of three writing assignments, the comparison between student peers with instructors on the same writing assignment showed great similarities and differences. Neither student peers nor instructors gave grade F to students' drafts. Student peers and instructors assigned the same number of grade Bs - 30Bs (30%) for students' graded draft 1 – Writing Assignment 1 and close to the number of grade Bs for students' graded draft 1 – Writing Assignment 3. Despite similarities, there was even greater difference in the number of letter grades given by student peers and instructors. For example, grade As given by student peers outnumbered grade As given by instructors; instructors tended to give more low grades (i.e. Cs and Ds) on first drafts than student peers. In addition, statistical analysis in this study shows that there were significant differences between the grades given on student papers by student peers and those given by the instructors. For both Draft 1 and

Final Draft of all three writing assignments, the descriptive analysis revealed that the means of peer-grading scores were higher than the means of the instructor grading scores. This indicates that the grades the students awarded to their peers tended to be higher in comparison to the instructor-assigned grades.

Overall, the findings on student peers' grades in comparison with instructors' grades show that there was a low consistency among grades given by student peers and grades given by instructors. While studies by Cho et. al (2006), Sadler & Good (2006), and Bouzidi & Jaillet (2009) show a high correlation between peer grading and teacher grading, the findings of this study are opposite to previous studies in the literature (i.e. Cho et al., 2006; Sadler & Good, 2006; Bouzidi & Jaillet, 2009). Possible explanations for this low consistency among peers' grades and instructors' could be that previous studies' findings (i.e. Cho et al., 2006; Sadler & Good, 2006; Bouzidi & Jaillet, 2009) are generally based on the context of non-composition courses with small or moderate enrollments, not a large scale of MOOC student population like this study. However, the present study's findings on the consistency among peer and instructor grades were also different from other studies in a MOOC, though non-composition context. The most comparable of these studies is by Lou et al., (2014); it is the only empirical study to examine peer grading in a MOOC (at least to the best of the research's knowledge). Their findings provided evidence that peer-grading scores were highly similar to instructors grading scores, which is contrary to the finding of this study.

While this study was not designed to look at factors that might affect the consistency among peers' grades and instructors', the analysis of this study's data on grading and of its relationship to the survey data and to comments by peers and instructors suggests that several other possible factors besides size of the class enrollment might help account for differences

among peer grades and between peer grading and instructor grading. In some instances, the impact of the factors noted here are difficult to determine given. In the context of this study, differences between student grading scores and instructor grading scores might result from different factors as follows:

Students' Attitudes: Among these factors would be the students' perception of themselves as peer assessors in their survey responses, for instance in responses that relate to the level of confidence they see themselves as bringing to their assessment. As reported in the Results chapter, approximately 65% felt that they were not confident in their ability to assess their peers' writing. Another finding interesting to consider in relationship to a lack of confidence and the higher grades given by peers as compared to instructors, is peers' greater likelihood than instructors to respond with comments that praise characteristics in their peers' papers. However, aspects of how the mixed methods data for this study was collected stand in the way of conclusions about whether either of these factors contributed to the majority of peers having given higher grades than the instructors. With the anonymity of the surveys and the grades, examining a potential connection between a lack of confidence and higher grades given or between praise used in comments and higher grades is not possible. Yet the combination of these findings collected from this study's use of mixed methods suggests the value of future research that focuses on the role of such factors as lack of confidence and praise offered might play in the grading differential.

Diversity of MOOC Students: This study is also limited in conclusions that can be drawn on how some of characteristic of the students studied might have influenced some others. Given the

lack of research on the quality and viability of peer review use in MOOC composition courses, this study focused on this gap in the scholarship for students in general and not on a design to reach findings relevant to the diverse and mainly ESL/EFL students that have tended to make up the majority of students taking composition classes in the E-Center for Professional Development. Nonetheless, the diversity of the students in this study's MOOC composition course is important to highlight in the discussion of the findings and their limits. Students in this study's MOOC-based composition course - ESL/EFL Writing for Academic and Professional Success were diverse, coming from many different countries (mostly non-native English speaking countries), at all ages and from a variety of culture backgrounds. More importantly, the students had greatly varied levels of English proficiency. 5% of the survey participants expressed that their difficulties/challenges with peer assessment were due to low English competences (survey question #20a). For example, a respondent stated "*One of the challenges was that my English is not good enough. It's my low proficiency in English that limited my ability in evaluating others' work. I hardly knew if peers had good word choices or used correct grammar, hence I could neither point out peers' errors nor made suggestions"* (Male 315, Indonesia).). However, as is the case with student perception of their own lack of confidence discussed above, triangulation of the students' survey responses on their perception of their low language proficiency with their grading and their comments to better interpret the role of language proficiency was not possible, again given the anonymity of the grading and the survey responses. This inability to use this study's mix of methods to shed light on the language proficiency issue is another limitation of the current study. In future research, a research design able to take full advantage of the mixed methods employed would be helpful so as to test out what role a low level of English proficiency might play in grading patterns.

Peer Assessment Tools: From the survey responses, approximately 15% of students claimed that one of the challenges for them as a commentator and grader was the peer assessment tools i.e. the rubrics (survey question #20a). Many students did not understand standard criteria for peer assessment and the rubrics thoroughly. The lack of understanding of the rubric and peer assessment criteria may be the source of the difference of peer grades and instructors.

Students' Qualification in Peer Assessment: Many students in the survey (around 61.5%) believed that they themselves were not qualified to grade peers' writing, giving reasons like lack of confidence, lack of experience, low English competences, and so forth. The course provided students with some training, within a limited time of one week, on how to write comments and how to grade the papers based on the assigned rubrics before peer assessment. The majority of students (at least 85.6%) responded that they took part in peer assessment training, followed the guidelines provided during training for commenting on and grading peers' writing, and followed the rubrics for grading peers' writing. However, that students following the peer assessment instructions and rubrics does not guarantee that students will assess peers' writing accurately, given that students received the training over only a short period of time. Students should have quality follow-up practice after receiving the initial training on peer assessment because good grading technique is difficult to learn.

Number of student graders: The number of peers grading each peer essay in this study may also be a factor that can help account for the inconsistency between grades assigned by peer graders and those given by instructors. At least the findings of a study by Cho et al. (2006) would suggest that (the study by Cho et al. was not conducted in a setting, like the present's study, of a

composition MOOC). In this study's composition MOOC - ESL/EFL Writing for Academic and Professional Success, each submitted paper was graded by three student graders based on the same rubric. The final peer-grading score for a paper was calculated by the mean of all the scores given by peers. With more graders, the consistency of student-grading scores might be more significant (as found in Cho et al., 2006). For more significant consistency in peer-grading scores, the suggested number of student graders should be at least five or seven.

Using peer-grading scores as final scores: Similar to other MOOCs, the ESL/EFL Writing for Academic and Professional Success course in this study utilized peer assessment due to the large enrollments of thousands of students. That peer-grading scores are used as final scores for students' submitted papers and as a replacement for instructor grading scores may cloud the accuracy of peer grading, especially when students tend to inflate peers' writing and award peers high scores.

3. Students' Comments Compared to Instructors' Comments

In a selective sample of 20 essay first drafts of Writing Assignment 3 – Argumentative Essay - of the ESL/EFL Writing for Academic and Professional Success course, the total number of comments made by student peers (three peers per essay) was 865 and that by the instructors (one per essay) was 278. Based on individual papers, the number of comments made by each of individual student peers varied among peers and also differed from the number of comments made by instructors on the same paper.

The analysis of the comments by student peers and instructors was done through a coding scheme with two major categories: the WHAT consisting of three criteria - Idea Development, Organization, and Sentence Level, and the HOW consisting of three criteria: Explanation, Question, and Correction. In terms of the WHAT, among comments made by student peers, 20% of the comments were on idea development, 5.3% of the comments on organization and 70% of the comments on sentence level issues. In regard to instructors' comments, approximately 25.9% of the comments focused on idea development, 8% of the comments on organization, and 64.7% of the comments on sentence level. Since most textbooks and even reading materials used during the training on peer assessment in the ESL/EFL Writing for Academic and Professional Success course put considerable stress on global issues of idea development and organization, it might be expected that instructors would emphasize these issues. It was surprising, then, to find that although instructors commented more on average on idea development and organization than did student peers, both peers and instructors had the majority of comments on sentence level issues but very few comments on organization. For example, instructors had 24/25 (SP2), 14/16 (SP11), and 18/18 (SP12) comments on sentence level issues; some student peers had 100% of the comments on only sentence level issues (e.g. peer 2 – SP4 (12/12); peer 1 – SP12 (12/12), peer 3 – SP12 (11/11)). That student peers in this study had more comments on sentence level issues than did instructors supports what Keh (1990) observed – "students have a tendency to read for surface, mechanical errors, usually referred to as 'lower order concerns' (LOCs). Students tend not to read for 'higher order concerns' (HOCs) such as the development of ideas, organization, and the overall focus of what they are writing" (Keh, 1990, p. 296). This study was not designed to focus on the reasons that might be behind this tendency for students—either native speaker or non-native speaker—to focus on sentence-level issues and cannot use its

findings to help explain them. However, given that all of the students of the MOOC-based course I studied– ESL/EFL Writing for Academic and Professional Development – were English non-native speakers, I can report anecdotally on observations from my experience as an ESL learner and then as an ESL teacher that might explain this tendency and also point toward a potentially productive focus for further research on reasons behind the sentence-level focus of comments by my study's ESL student participants. First, ESL students, who did not acquire English through immersion but rather learned it through its rules, might more often do grammar exercises and thus know English grammar and English structures better than other aspects of English language. Second, ESL students might not be much aware of various factors that are involved in structuring the text as well as rhetorical choices of English composition and thus be less able to comment on writing issues on that level. With that said, ESL students might be more confident to give feedback on sentence level issues like grammar and sentence structure rather than larger issues of focus, idea development, and organization. Regarding the fact that instructors had more comments on lower order concerns, not more comments on higher order concerns as expected, though both higher order and lower order concerns are important, it is interesting to find that this tendency was noted by previous researchers such as Sommers (1982), Zamel (1985), and Connors and Lunsford (1993) although these researchers did not point out reasons for this tendency. Connors and Lunsford, though, suggestively stated:

> "*Most composition teachers know what the field says is important-our public "tropes," so to speak. We talk and write often of purpose, of audience, of organization, of proof, of process and invention and revision and so on. These words fill our journals, our professional books, our conferences, and especially our textbooks. But do we really follow through? Do comments on papers show us acting on these public tropes, giving them more than lip service? Or do we have more genuine and less overt agenda*."
>
> (Connors and Lunsford ,1993, p. 209)

The plausible interpretations of the instructors' tendency to give feedback on lower order concerns can be varied. First, the writers of the graded papers were ESL students, and the instructors of the course ESL/EFL Writing for Academic and Professional Development were all experienced in teaching ESL students. According to Leki (1990a), ESL students are likely to violate much more basic categories of English, for example placement of adjectives: *a day beautiful;* they may not be familiar with certain forms and uses of English for example the distinction between formal and informal registered English. ESL instructors might be aware of language difficulties and linguistic complications that ESL students might have. Therefore, instructors might feel that their major task is to promote language learning and thus provide more feedback on sentence level issues such as language use and grammar. Second, the papers graded were first drafts of the writing assignment and students would revise their first drafts before submitting the final ones. The instructors might view feedback on the first drafts as more developmental by focusing on correcting and editing. Third, Leki (1990b) reported that when instructors provide comments on higher order concerns like idea development and organization, students react in different ways: they might not read the comments; they might read the comments but not understand the comments; or they might understand the comments but do not know how to response to the comments. Therefore, it might make more sense for instructors to focus more on the sentence level issues.

In terms of how each of the criteria of the WHAT should be communicated, the HOW of the comments was analyzed based on three criteria: (1) Explanation, (2) Question, and (3) Correction. The findings show that student peers' comments on Idea Development (a total of 173 comments by three peers/paper) included 85.5% explanations, 16.8% questions and 18.5% corrections while instructors' comments on Idea Development (with a total of 72 comments) had

95.8% explanations, 5.6% questions, and 47.2% corrections. It is not surprising that instructors provided more explanations and corrections in their comments on Idea Development than the student peers. This might be because (1) the instructors are considered more experienced and knowledgeable, and (2) their comments are considered to be of better quality (i.e. in terms of clarity and specificity). Regarding comments on Organization, student peers made a total of 46 comments including 82.6% explanations, 19.6% questions, and 19.6% corrections; instructors had a total of 22 comments with 95% explanations, 9.1% questions and 50% corrections. Although both student peers and instructors tended to explain the problems and offered suggestions for corrections rather than raising questions in their comments on Organization, the instructors provided more explanations and corrections than the student peers. Again, the plausible interpretation is instructors are considered to be more experienced and knowledgeable than students. For comments on Sentence Level, among 605 student peers' comments were 262 comments (43.3%) with explanations, 18 comments (3%) with questions, and 508 comments (84%) with suggestions for correction. Instructors made a total of 180 comments on sentence level issues, which consisted of 109 explanations (60.6%), 1 question (0.6%), and 173 corrections (96.1%). Not mentioning whether the explanations and suggested corrections were right or wrong, it is interesting to find that student peers tended to provide explanations and suggestions for correction in most of their comments on Sentence Level. Again, this tendency might be explained by the previous interpretation, that, given the focus on rules and correctness in their learning of English, ESL students might be more confident to give their feedback on sentence level issues like grammar rather than larger issues of focus, idea development, organization, etc. In addition, it is also interesting to notice that instructors provided way more explanations and corrections in their comments on Sentence Level than students. This tendency

seems to be recommended by Mina Shaughnessy – isolating the problems/errors and correcting them. With 96.1% of the comments on Sentence Level offered corrections, it can be understood that the instructors might want to offer students (ESL students in this study) as much help as they can. Sometimes there seems to be an obvious temptation to get control over the student's paper, correct every single error a writer makes, make the writing in higher register compared with the actual language proficiency of the writer, and even change the meaning of what the writer is actually saying. It seems that the instructors' appropriation of second language writing arises not in relation to control of topic or content, or cultural knowledge, but to control of language. As Sommers (1982) pointed, that instructors appropriate student writing especially when they mark up errors in style, diction, and usage in a first draft would make students understand that these errors need to be addressed before the meaning of the text is attended to.

4. Students' Comments in Connection with Assignment Guidelines and Rubrics

Prior to peer assessment, instructors had explained to students the assignment guidelines and guided through a training process for understanding and using the peer assessment tool (e.g. the rubric) so that students would be then able to use this tool to assess the work of their peers. From the analysis of the comments made by student peers on 20 first drafts of Writing Assignment #3, the study looked into words/phrases connected with words and emphasis in the rubric and assignment guidelines to bring out patterns that revealed student peers following reference to specifics within the rubric, the assignment guidelines, and instructions in the peer assessment training. The findings show that 46 student peers (out of 60) used certain terms/phrases in the rubric and the assignment guidelines (43.8%) or used wording that in

essence connects the comment to the rubric and the guidelines (56.2%) in 89 comments. These findings support what was found from the survey and the interviews. Survey responses revealed that 85.6% of the participants took part in peer assessment training and referred to the training instructions and rubric when commenting on and grading their peers' papers. Because the survey question combined two aspects of taking part in the peer assessment training and referring to the training instructions when commenting in one single question, there might be some overlap in interpreting the result. However, this result indicates that students perceived their comments as connected with the guidelines and rubrics in some way. In addition, more than 88% of the survey participants agreed or strongly agreed that for each of the assignments they followed the guidelines provided during training for commenting on and grading peers' writing and followed the rubrics for grading peers' writing. The interviews corroborated the survey findings. From the interview findings, the majority of interviewees expressed that they followed the guidelines and the rubrics when assessing peers' work; for example, one interviewee responded that *"I followed the rubrics and used some terms from the rubrics when I wrote the comments. I tried to make my comments clear to my peers."*

There are some hypotheses for the reasons students used terms/phrases connected to the guidelines and rubrics when commenting. One of the hypotheses is students applied what they learned from peer assessment training – i.e. following the instructions, using the rubrics as powerful tools to reflect, unpack, and critique. If this hypothesis holds true, training on peer assessment in the MOOC-based composition course is considered to be helpful for students, at least in assisting students to write relevant commentary on peers' writing. The second hypothesis for the reason students used terms/phrases from the rubrics or the guidelines is students had difficulty writing comments, or did not feel confident in writing comments, or did

not know how to write commentary. Responses from the survey support this hypothesis – more than 72% of the survey respondents expressed that they had difficulty writing comments. A survey participant expressed:

> *"Writing comments was the most challenging to me. I took a lot of time to practice from the training how to make comments. Many times I got stuck at not being able to think of what comments should be made. I found it very hard to write relevant and constructive comments."*
>
> (Female 69, Japan)

Therefore, it might be students' difficulties/challenges in writing comments that made students depend on details in the guidelines/rubrics. If the second hypothesis holds true, it can be understood that the guidelines and the rubrics plays an important role in helping students to write commentary on peers' writing.

The findings from the analysis of student peers' comments also show that 85.4% of the comments that had terms/phrases or wording connected with the rubric and assignment guidelines addressed global issues of idea development and organization. The plausible interpretation for this is student peers in the ESL/EFL Writing for Academic and Professional Success MOOC-based Course might not be fully aware of various factors that are involved in structuring the text as well as rhetorical choices of English composition, and they might not be confident to give their feedback on global issues of focus, idea development, organization (as explained in section 3 of this chapter). Given that, student peers tended to rely on the guidelines and rubrics when they made comments on such global issues as idea development and organization. Although assignment guidelines and rubrics cannot completely substitute for student peers' commentary, detailed and clearly articulated assignment guidelines and rubrics can be helpful for student peers to reflect, unpack, and critique.

5. Limitations of the Study

Earlier in this chapter, especially in the section comparing peers' and instructors' grades, limitations of this study related to the lack of triangulation possible in the separate data sets have been explained and the implications for future research related to these limitations forwarded. This section will forward other limitations, ones that may not be as readily identified.

First it seems important to note the potential impact on qualitative data that has been suggested when qualitative data has been quantified (Bazeley 2004). The current study employed a mixed methods approach, combining paradigms of quantitative and qualitative research to look into connections and disconnections among the different data sets and provide the differing perspectives to answer the research questions. A few portions of the qualitative such as the survey's open-ended questions, and students' graded papers (including the students' comments, and the instructors' comments) were analyzed through a coding process and then calculated in numbers and percentages. According to Bazeley (2004), loss of depth and flexibility occurs when qualitative data are quantified. In addition, the need to collect and analyze qualitative data forced the research to reduce from the original plan the sample size (i.e. the number of the students' papers used for the analysis of peer and instructor comments). A greater sample size including comments from more of the commented on drafts collected for data analysis would have made the results more convincing than was possible with results from analyzing only 20 papers of those most commented on. Plus, because the study was conducted based on a context of small-sized composition MOOC, the findings of this study of a small-sized MOOC may not be predictive of the findings of research on the use of peer assessment in the context of large-sized composition MOOCs.

Second, one of the data sources used in this study consisted of an online survey. Surveys are self-report measurement techniques designed to question people about themselves, their attitudes and behaviors (Creswell, 2003). This type of measurement can be potential sources of unreliable answers because respondents may exaggerate. Participants may be embarrassed to state their true response, or they may simply forget the true account. With both Likert scale questions and open-ended questions in the online survey, it is possible that some participants might not have wanted to spend sufficient time giving detailed and true responses, which might make the results of the study not convey the true perceptions of the participants. In addition, a few questions in the survey had a combination of the two questions, for example "*I participated in peer assessment training and referred to the training instructions and rubric when commenting on and grading my peers' papers*" (survey question #5). Participants' responses to "Did you take part in the training?" could be quite different than their responses to "Did you refer to the training instructions and rubric when commenting?" As a consequence of combining them, there could have been some loss in interpreting the results of this item -- there could be high agreement responses on participating that could mask lower agreement with the degree to which the respondents referenced the guidelines and rubrics and vice versa.

Third, there is the lack of research and scarcity of document archives on issues the study discussed. The concept and practice of MOOCs is still new to researchers and educators, and there have been very limited studies about this area. Therefore, the literature review chapter of the study might lack significant discussions about this scholarship, especially scholarship of composition MOOCs. Besides, the concept of "peer assessment" discussed in this study is used as an umbrella concept to capture the diversity of definitions and involves any activity in which students make only comments on peers' work, or students only assign a grade to peers' work, or

students both make comments on peers' work and assign a grade. Although in the literature review chapter this study tried to make a distinction by using "peer review/peer response" to refer to an activity in which students make commentary on peers' work and "peer grading" to indicate an activity in which students assign a grade to peers' work, when reporting the results regarding students' perception of peer assessment, the generalization was made for the combination of peer review and peer grading. This broad conclusion drawn from the results of the study might cause certain loss in interpreting students' perception of peer review and that of peer grading, for example students' perception of commenting and grading could be different from each other.

Fourth, the participants of the study were diverse in terms of genders and cultures. Gender and culture bias should have been reflected in the study's results because of the inclusion of a greater proportion of participants from one language or cultural group or from one gender more than another; however, the study's design wasn't set up with a way to account for or compensate for such a bias. In addition, the study did not examine other factors such as level of education, English language proficiency and cultural background to see whether these factors might affect the participants' perception of peer assessment and their assigning grades to peers' writing and their commentary of peers' writing.

Last but not least, the finding that the majority of the participants perceived that peer assessment was useful and feedback from peers helped improve their writing. However, the study did not examine the students' revisions on the papers they wrote for the course to see if and how students' writing improved after peers' feedback. Moreover, the study strived to compare student peers' comments with instructors' comments, but the study just looked into marginal comments made by student peers and the instructors, not the head notes or end notes. The

reasons that the study did not look at head and end notes are (1) it was optional for students to make head and end notes; (2) it was noticed that only one paper in the sample had an end note, and analysis of this end note in the context of this single paper would not be significant enough for any conclusion to be drawn. Analysis of the head or end notes could have added to the analysis of marginal comments connections between the big picture (head/end notes) and the details (marginal comments). Also, regarding the analysis of the comments, the study did not look into the length of the comments (how many words in the comments) made by student peers and instructors. The reason that the study did not look at this aspect is that the study focused on marginal comments rather than on end comments, where length, given the specificity it allows, would be more likely to be an issue. And in analyzing student peers' comments in terms of the WHAT and the HOW, the WHAT and the HOW could explain the clarity though still lack of the analysis of the length of the marginal comments might cause some loss in reporting the relative specificity of comments. Finally, it is noted in the study that student peers made very vague and general comments at certain points like "*?*", "*Nice job!",* "*Grammar*" but the study was not designed to examine possible reasons why students may not have commented more than instructors using this general type of comment.

6. Viability of Peer Assessment in Composition MOOCs

The large enrollments of multiple thousands of students in MOOCs seem to exceed the assessment capacity (i.e. evaluating and grading) of instructors; therefore, the inability for instructors to grade so many papers is likely responsible for MOOCs in general and composition MOOCs in particular turning to peer assessment. One question raised is "Does peer assessment work in MOOCs in composition? The answer is both YES and NO.

On one hand, peer assessment in ESL/EFL Writing for Academic and Professional Success MOOC-based Course was well received by students. Students had positive perceptions and attitudes towards peer assessment, especially the usefulness of peer assessment; the majority of students used peers' feedback when they revised their writing and peer feedback helped improve students' writing performance. The improvement of students' writing performance can be evidenced by scores that students got. For all three writing assignments, the mean of students' scores for Final Draft was higher than those for Draft 1, no matter who graded students' writing - peers or instructors.

On the other hand, the challenges of peer assessment noted in the survey responses of the 1215 survey partcipants – such as students' lack of confidence, students' lack of experience, students' lack of qualification, diversity of MOOC students, anonymity of peer assessment, peer assessment tools, and accuracy of peer assessment - may have affected the quality of peer assessment in this composition MOOC. Although given the anonymity of the survey responses, I could not triangulate these responses with other datasets to analyze how these challenges might have affected the assessments' quality. The challenges, including 72% of the respondents noting difficulty writing comments and assigning grades, seem important to keep in mind in designing MOOCs in composition. This is particularly true since, due to the large scale of MOOC student enrollment peer assessment is likely to continue to be considered the most feasible assessment method and peer assessment does not work by magic. For optimal peer assessment and reduction of challenges, providing students with sufficient training on peer assessment and creating high-quality peer assessment tools i.e. rubrics, peer assessment instructions, and guideline sheets are of utmost importance.

In terms of training on peer assessment, findings of the study on students' perception of peer assessment may indicate that instructors' preparations play an important role in the success of peer assessment. To account for the importance of this role, prior to the peer assessment session, instructors need to determine learning goals and effective peer assessment strategies based on students' writing proficiency, feedback skills, and experience in collaborative work. By doing so, instructors are more likely to know what criteria to set for the peer assessment and what strategies can help to maximize the activity. Moreover, since composition MOOCs have a great diversity of students and issues related to this diversity were noted by a few survey participants, it would be advisable to include some cultural awareness training, so that students begin to appreciate the subtle differences in peer responses. Last but not least, practice makes perfect. Students need sufficient time to practice for quality peer feedback.

In addition to the findings that seem to suggest the value of sufficient training on peer assessment, findings on student perception of rubrics as confusing could indicate that a high-quality rubric might be also an important factor for the success of peer assessment. To be clear on the meaning of the rubrics, peer assessment tools should state a number of specific questions or present a list of areas for the readers to focus on when assessing peers' writing. If guided questions and instructions are too general, students might have difficulty identifying what to respond.

Moreover, demographic information of the course participants shows that composition MOOCs have diverse students including ESL students; survey participants perceived different levels of English proficiency (5%) and different rhetorical strategies (3%) as challenges/difficulties they encountered. Therefore, it is important to keep in mind in designing composition MOOCs that factors such as how linguistic, personal history with writing, and

cultural backgrounds might influence the organizational structures of ESL student texts and how the students perceive peers' writing. Clearly defined words and phrases in assignment rubrics and guidelines (e.g. 'thesis,' 'thesis statement,' 'topic sentence,' 'audience,' and the like) are of great importance for ESL students since they can help them develop an awareness of English rhetorical patterns and structure of writing. Moreover, peer assessment tools (i.e., peer assessment instructions and peer assessment guided questions) explore the questions of what causes writers to revise. According to Sommers (1980) in her study on revision strategies, one of major reasons that make students revise is "incongruities between intention and execution" (p. 385). Plus, the social interaction perspective on writing (i.e. the relationship between the writer and reader) and the mismatch possible between a writer's intention and the reader's comprehension do offer insights about revision, such as what aspects of the writer's work the peer reviewer might be trying to get the writer to rethink and accomplish the next draft. Nold (1981) added that in order to have successful revisions, writers need to both realize mis-constraints and have the capacity to produce a clearer alternative to the current text.

7. Recommendations for Future Research

With limitations of the study discussed above, several recommendations can be made for future research. First of all, to broaden the understanding of peer assessment's use in MOOCs, future research might investigate aspects of peer assessment in larger size composition MOOCs and increase the sample size for more convincing results and conclusions; the need for a larger sample size would be particularly worthwhile for the analysis comparing peer and instructor comments, limited in the current study to a sample of 20 student essays. Second, since students

in composition MOOCs tend to be diverse, future research would be useful on whether such factors as gender, culture, levels of education, and English language proficiencies affect students' perceptions and attitudes toward peer assessment as well as their assessment of peers' writing in composition MOOCs. Regarding students' commentary on peers' writing in composition MOOCs, future studies examining different aspects of different kinds of commentary would be worthwhile. For example, the nature and specificity of head-notes, end-notes and marginal comments would be useful to analyze according to the commentary length and also on the extent to which student writers incorporate in their revisions-the advice in their peers' comments. From analysis of these sorts of components and consequences of peer comments, future research can add useful findings that can potentially advance the understanding of peer assessment's usefulness and viability in composition MOOCs and the degree and kind of training needed to maximize its viability and usefulness. Future studies might usefully consider both the extent to which the writers' perception of comments might be deemed worthwhile for deciding how viable peer assessment is or whether extra training to avoid vague/general comments would be worthwhile. Finally, future studies would be useful to look into peer assessment as a learning assessment tool, emphasizing student learning from participating in peer assessment as the assessor as well as the assessed.

8. Contributions of the Study

Despite its limitations, this study has made certain contributions to the existing literature on peer assessment, especially when employed in a composition MOOC, a type of MOOC for which little research yet exists. One contribution this study has made to the research on student

perceptions of peer assessment is its finding on the positive attitudes and perceptions of students have towards peer assessment. Although similarly positive student perceptions of peer assessment have been found in previous studies, these studies were carried out in different disciplines and contexts, such as non-MOOC composition contexts and non-composition MOOC contexts (e.g. Simkin and Ramarapu ,1997; Nelson and Carson,1998; Mangelsdorf, 1992; Brammer and Rees, 2007; Lou et al., 2014). Since there are no studies in the literature looking at student perceptions of peer assessment in a composition MOOC context, the findings of this study adds to the existing literature the fact that students - at least those in this study - have positive perceptions towards online peer assessment in a composition MOOC context. It also adds to the literature findings that training on peer assessment in composition MOOCs may play a role in helping students do peer assessment more effectively, given the comments of peers on their increased confidence in carrying out peer assessment as the course and its training progressed. In addition, the analysis of the grades given by peers and instructors in the study provides a mix of evidence-based information (1) about whether online peer assessment should be used in MOOCs, especially MOOCs in composition—for example, the fairly high consistency among grades given by peers might be considered evidence of the viability of peer grading in this composition MOOC while the low consistency between peers' grades and instructors' grades might be considered potentially against its viability—and (2) about what factors might affect the applicability and consistency of peer grading in MOOCs. Analysis of the data also provided insights into types of comments that students in a composition MOOC made on their peers' essays and what the comments looked like, such as their greater percentage of comments on sentence-level errors as compared to instructors' comments or the greater percentage of their comments that take the form of corrections. These insights have the potential to usefully inform

the quality of the training on peer assessment and the practice of peer assessment in composition MOOCs. Besides, this study may also provide information for writing program administrators, writing centers and university administrators that would help them establish more appropriate and efficient writing services for their students, especially for MOOC students.

REFERENCES

Bain, K. (2006). What the best college teachers do. *Perspectives: Teaching Legal Research and Writing*, *15*(1).

Baird, J. R., & Northfield, J. R. (1992). Learning from the PEEL experience. Melbourne, Australia: Monash University.

Baker, S., Gersten, R., & Graham, S. (2003). Teaching Expressive Writing to Students with Learning Disabilities Research-Based Applications and Examples. *Journal of Learning Disabilities*, *36*(2), 109-123.

Balfour, S. P. (2013). Assessing writing in MOOCs: automated essay scoring and calibrated peer review. *Research & Practice in Assessment*, *8*(1), 40-48.

Bazeley, P. (2004). Issues in mixing qualitative and quantitative approaches to research. *Applying qualitative methods to marketing management research*, 141-156.

Brammer, C., & Rees, M. (2007). Peer review from the students'perspective: invaluable or invalid?. *Composition Studies*, *35*(2), 71.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (1999). *How people learn: Brain, mind, experience, and school*. National Academy Press.

Bennett, R. E. (2006). Technology and writing assessment: Lessons learned from the US National Assessment of Educational Progress. Annual Conference of the International Association for Educational Assessment, Singapore, IAEA. http://www.iaea.info/documents/paper_1162a26d7.pdf.

Bill & Malinda Gates Foundation. Retrieved at http://www.gatesfoundation.org/What-We-Do/US-Program/Postsecondary-Success

Boston, W. & Helm, J. S. (2012). Why student learning outcomes assessment is key to the future of MOOCs. *National Institute for Learning Outcomes Assessment.* Retrieved at http://illinois.edu/blog/view/915/84723?displayType=month&displayMonth=201212

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

Bridgeman, B., Trapani, C., & Yigal, A. (2012). Comparison of human and machine scoring of essays: Differences by gender, ethnicity, and country. Applied Measurement in Education, 25(1), 27-40.

Bruffee, K. A. (1984). Collaborative learning and the" Conversation of Mankind". *College English*, *46*(7), 635-652.

Byrne, R., Tang, M., Truduc, J., & Tang, M. (2010). eGrader, a software application that automatically scores student essays: With a postscript on the ethical complexities. Journal of Systemics, Cybernetics & Informatics, 8(6), 30-35.

Carson, J.G. (1992). Becoming bi-literate: First language Influences. *Journal of Second Language Writing, 1*, 37-60.

Caulk, N. (1994). Comparing teacher and student responses to written work. *TESOL Quarterly*, *28*(1), 181-188.

Chen, C-F. E., & Cheng, W-Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94-112.

Cheville, J. (2004). Automated scoring technologies and the rising influence of error. *English Journal*, 93(4), 47-52.

Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, *98*(4), 891.

Chodorow, M., & Burstein, J. (2004). Beyond essay length: Evaluating e-rater's performance on TOEFL essays (TOEFL research report, No. RR-04-73). Princeton, NJ: Educational Testing Service.

Cindy, J. (2007). Validating a computerized scoring system for assessing writing and placing students in composition courses. *Assessing Writing*, 11(3), 167-178.

Colman, D. 2013, MOOC Interrupted: Top 10 Reasons Our Readers Didn't Finish a Massive Open Online Course. 5 April 2013. *Dan Colman: Open Culture*. Available from: < http://www.openculture.com/2013/04/10_reasons_you_didnt_complete_a_mooc.html/ >. [30 April 2013]

Condon, W. (2013). Large-scale assessment, locally-developed measures, and automated scoring of essays: Fishing for red herrings? *Assessing Writing*, 18(1), 100-108.

Connor, U. & Asenavage, K. (1994). Peer response groups in ESL writing classes: How much impact on revision? *Journal of Second Language Writing, 3*, 257-276.

Connors, R. (1997). *Composition-rhetoric: Backgrounds, theory, and pedagogy.* Pittsburg: University of Pittsburg Press.

Coursera. (2012). Peer Assessment.

Coursera. Duke University. Retrieved at https://www.coursera.org/course/composition

Coursera. Georgia Institute of Technology. Retrieved at https://www.coursera.org/course/gtcomp

Coursera. The Ohio State University. Retrieved at https://www.coursera.org/course/writing2

Creswell, J. W. (2003). Research design: Qualitative, quantitative, and mixed methods approaches (2nd ed). Thousand Oaks, CA: Sage Publications.

Creswell, J.W. & Clark, P. V.L. (2007). *Designing and conducting mixed methods research.* Thousand Oaks, CA: Sage.

Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). *Advanced mixed methods research designs.* In A. Tashakkori & C. Teddlie (Eds.)*, Handbook of mixed methods in social and behavioral research (pp. 209-240).* Thousand Oaks, CA: Sage.

Deane, P. (2013). On the relationship between automated essay scoring and modern views of the writing construct. *Assessing Writing*, 18(1), 7-24.

DiPardo, A., & Freedman, S. (1988). Peer response groups in the writing classroom: Theoretic foundations and new directions. *Review of Educational Research, 58,* 119-150.

Downes, S. et al. (2011). *The MOOC guide.* Retrieved at https://sites.google.com/site/themoocguide.

Educause. (2012). Retrieved from http://net.educause.edu/ir/library/pdf/eli7078.pdf

Educause (2013a). Peer Assessment in MOOCs. Retrieved at https://net.educause.edu/ir/library/pdf/ELI139_OL14.pdf

Educause. (2013b). Writing II: Rhetorical Composing. Retrieved at http://www.educause.edu/sites/default/files/library/presentations/E13/SESS008/Writing2-Final-Report.pdf

Elbow, P. (1981). *Writing with power: Techniques for mastering the writing process*. Oxford University Press.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of educational research*, *83*(1), 70-120.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Innovations in Education and Training International,* 32(2), 175-187.

Falchikov, N. (2001). Learning together: Peer tutoring in higher education. London: RoutledgeFalmer.

Falchikov, N. (2005). Improving assessment through student involvement. New York: Routledge Falmer

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.

Ferenstein, G. (2014). Massive Online Courses enroll an average of 43,000 students, 10% completion. Retrieved from http://techcrunch.com/2014/03/03/study-massive-online-courses-enroll-an-average-of-43000-students-10-completion/

Fry, S. A. (1990) Implementation and evaluation of peer marking in higher education. *Assessment and evaluation in higher education, 15,* 177-189.

Fosmire, M. (2010). Calibrated Peer Review: A New Tool for Integrating Information Literacy Skills in Writing-Intensive Large Classroom Settings. Portal: Libraries & The Academy, 10(2), 147.

Gardiner, L. F. (1994). *Redesigning Higher Education: Producing Dramatic Gains in Student Learning. ASHE-ERIC Higher Education Report No. 7*. ERIC Clearinghouse on Higher Education, One Dupont Circle, NW, Suite 630, Washington, DC 20036-1183.

George, D., & Mallery, P. (2009). *SPSS for Windows step by step: a simple guide and reference 16.0 update*. Boston, MA: Allyn and Bacon.

Gielen, S., Peeters, E , Dochy, F , Onghena, P , & Struyven, K (2010) Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20(4),* 304-315.

Graesser, A. C., & McNamara, D. S. (2012). Automated analysis of essays and open-ended verbal responses.

Grosseck, G. (2009). To use or not to use web 2.0 in higher education? *Procedia-Social and Behavioral Sciences*, *1*(1), 478-482.

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

Head, K. (2013). Lessons Learned From a Freshman-Composition MOOC. *The Chronicle of Higher Education.* Retrieved at http://chronicle.com/blogs/wiredcampus/lessons-learned-from-a-freshman-composition-mooc/46337

Holmes, L. (2013, October 8). An Early Report Card on Massive Open Online Courses. The Wall Street Journal. Retrieved from http://online.wsj.com/news/articles/SB10001424052702303759604579093400834738972

Holt, M. (1992). The value of written peer criticism. *College Composition and Communication,*

43(2), 384-392.

Huff, R., & Kline, C. R. (1987). *The Contemporary Writing Curriculum: Rehearsing, Composing, and Valuing*. New York: Teachers College Press.

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.

Keh, C. L. (1990). Feedback in the writing process: A model and methods for implementation. *ELT journal*, *44*(4), 294-304.

Kehl, D. G. (1970). The art of writing evaluative comments on student themes. *English Journal*, 972-980.

Kolowich, S. (2013). George Siemens Gets Connected. The Chronicles of Higher Education. Retrieved from http://chronicle.com/article/George-Siemens-Gets- Connected/143959/

Krashen, Stephen. (1982). *Principles and practice in second language learning and acquisition*. Oxford: Pergamon.

Krause, S. (2013). The end of the Duke Composition MOOC: again, what did we learn here? Retrieved at http://stevendkrause.com/2013/06/21/the-end-of-the-duke-composition-mooc-again-what-did-we-learn-here/comment-page-1/

Kuh, G. (2012). *What matters to student success*. Keynote address to the annual National Symposium on Student Retention, New Orleans, LA. In Boston, W. & Helm, J. S. (2012). Why student learning outcomes assessment is key to the future of MOOCs. *National Institute for Learning Outcomes Assessment.* Retrieved at http://illinois.edu/blog/view/915/84723?displayType=month&displayMonth=201212

Leal, F. Report: U.S students lack writing skills. Retrieved at http://www.ocregister.com/articles/students-371409-writing-graders.html

Lee, H.-J., & Lim, C. (2012). Peer Evaluation in Blended Team Project-Based Learning: What Do Students Find Important? *Educational Technology & Society, 15* (4), 214–224.

Lewin, T. (2012). Instruction for Masses Knocks Down Campus Walls. *The New York Times.* Retrieved at http://www.nytimes.com/2012/03/05/education/moocs-large-courses-open-to-all-topple-campus-walls.html?_r=4&hpw&

Leki, I. (1990a). Potential problems with peer responding in ESL writing classes. *CATESOL Journal, 3*, 5-19.

Leki, I. (1990b). Coaching from the margins: Issues in written response.*Second language writing: Research insights for the classroom*, 57-68.

Likkel, L. (2012). Calibrated Peer Review™ essays increase student confidence in assessing their own writing. *Journal of College Science Teaching, 41*(3), 42-47

Liu, E. Z. F., Lin, S. S. J., & Yuan, S. M. (2002). Alternatives to instructor assessment: A case study of comparing self and peer assessment with instructor assessment under a networked innovative assessment procedures. *International Journal of Instructional Media, 29* (4), 10.

Luo, H., Robinson, A. C., & Park, J. Y. (2014). Peer grading in a MOOC: Reliability, validity, and perceived effects. *Online Learning: Official Journal of the Online Learning Consortium*, *18*(2).

Mackness, J., Mak, J., Williams, R. (2010). The ideals and reality of participating in a Mooc", proceedings of the 7th International Conference on Networked Learning 2010. Retrieved from http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?_r=0

Magin, D. (1993). Should student peer ratings be used as part of summative assessment? *Research and Development in Higher Education, 16,* 537- 42.

Mangelsdorf, K. (1992). Peer reviews in the ESL composition classroom: What do the students think?. *ELT journal*, *46*(3), 274-284.

Masters, K. (2011). A brief guide to understanding Moocs. *The Internet Journal of Medical Education*. *1* (2).

Marchionda, D. (2004). Peer review times two. *National Writing Project.* Retrieved from http://www.nwp.org/cs/public/print/resource/1984

Mc Auley, A., Stewart, B., Siemens, G., & Cormier, D. (2010). The Mooc model for digital practice. Retrieved from http://www.elearnspace.org/articles/mooc_final.pdf.

McLoughlin, C. & Lee, M. J. W. (2007). Social software and participatory learning: Pedagogical choices with technology affordances in the Web 2.0 era. In ICT: Providing choices for learners and learning. Proceedings ascilite Singapore 2007. http://www.ascilite.org.au/conferences/singapore07/procs/mcloughlin.pdf.

Min, H. T. (2005). Training students to become successful peer reviewers *System*, *33*(2), 293-308.

Mousesite. Retrieved at http://sloan.stanford.edu/mousesite/

Murau, A. M. (1993). Shared Writing: Students' Perceptions and Attitudes of Peer Review. *Working Papers in Educational Linguistics*, *9*(2), 71-79.

National Council of Teachers of English. (2013). Machine scoring fails the test. NCTE Position Statement on Machine Scoring. Retrieved from http://www.ncte.org/positions/statements/machine_scoring

Nelson, G. L., & Carson, J. G. (1998). ESL students' perceptions of effectiveness in peer response groups. *Journal of second language writing*,*7*(2), 113-131.

Nelson, G. L., & Murphy, J. M. (1992). An L2 writing group: Talk and social dimension. *Journal of Second Language Writing*, 1, 171–193.

Nelson, G., & Murphy, J. (1993). Peer response groups: Do L2 writers use peer comments in revising their drafts? *Journal of Second Language Writing, 27*, 135-142.

New Media Horizon and EDUCAUSE. NMC Horizon Report: 2013 Higher Education Edition. New Media Consortium, 2013.

Nicol D. (2008). *Transforming assessment and feedback: Enhancing integration and empowerment in the first year*. Scotland, UK: Quality Assurance Agency

Onwuegbuzie, A. J., and C. Teddlie 2003 A Framework for Analyzing Data in Mixed Methods Research. In Handbook of Mixed Methods in Social and Behavioral Research. A. Tashakkori and C. Teddlie, eds. Pp. 351-383. Thousand Oaks, CA: Sage.

Pappano, L. (2012, November 2). The Year of the MOOC. *The New York Times*. Retrieved from http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html?pagewanted=all

Porter, J. E. (2013). MOOCs,'Courses', and the Question of Faculty and Student Copyrights. *The CCCCC-IP Annual: Top Intellectual Property Developments of 2012*, *8*, 2-21.

Rada, R., Michailidis, A., & Wang, W. (1994). Collaborative hypermedia in a classroom setting. *Journal of Educational Multimedia and Hypermedia, 3,* 21-36.

Rees, J. (2013). The MOOC Racket. Retrieved from http://www.slate.com/articles/technology/future_tense/2013/07/moocs_could_be_disastrous_for_students_and_professors.html

Rothschild, D., & Klingenberg, F. (1990). Self and peer evaluation of writing in the interactive ESL classroom: An exploratory study. *TESL Canada Journal*,*8*(1), 52-65.

Rushton, C., Ramsey, P., & Rada, R. (1993). Peer assessment in a collanorative hypermedia environment: A case study. *Journal of Computer- Based Instruction, 20* , 75-80.

Sadler, P. M., & Good, E. (2006). The impact of self and peer-grading on student learning. *Educational assessment*, *11*(1), 1-31.

Shermis, M. D., Burstein, J., Higgins, D., & Zechner, K. (2010). Automated essay scoring: Writing assessment and instruction. In E. Baker, B. McGaw, & N.S. Petersen (Eds.), International encyclopedia of education (3rd ed., pp. 75–80). Oxford, England: Elsevier.

Simkin, M. G., & Ramarapu, N. K. (1997). Student perceptions of the peer review process in student writing projects. *Journal of Technical Writing and Communication*, *27*(3), 249-263.

Sloan, C. (2013). *10th Annual Survey of Online Learning: Changing Course: Ten Years of Tracking Online Education in the United States. Retrieved from http://sloanconsortium.org/node/384451*

Swales, J. M., & Feak, C. B. (2004). *Academic writing for graduate students: Essential tasks and skills* (Vol. 1). Ann Arbor, MI: University of Michigan Press.

Sommers, N. (1982). Responding to student writing. *College composition and communication*, 148-156.

Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing*, *1*(3), 217-233.

Straub, R. (1997). Students' reactions to teacher comments: An exploratory study. *Research in the Teaching of English*, 91-119.

Thompson, R. F. (1981). Peer grading: some promising advantages for composition research and the classroom. *Research in the Teaching of English*, 172-174.

Topping, K. J.; Smith, E. F.; Swanson, I.; Elliot, A. (2000). Formative peer assessment of academic writing between post students. *Assessment & Evaluation in Higher Education, 25* (2), 151-169.

Topping, Keith J. "Peer assessment." *Theory into Practice* 48.1 (2009): 20-27.

Villamil, O. S., & Gurrerro, M. C. M. (1996). Peer revision in the L2 classroom: Social cognitive activities, mediating strategies, and aspects of social behaviors. *Journal of Second Language Writing, 5* (1), 51-75.

Vygotsky, L. S. (1978). Thought and language, 1962. *Mind and Society*.

Watters, A. (2012, August 27). The problems with peer grading in Coursera. *Inside Higher Ed*. Retrieved from: http://www.insidehighered.com/blogs/hack-higher-education/problems-peer-grading-coursera

Wen, M. L., & Tsai, C.-C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education, 27*(18), 27–44

Zamel, V. (1985). Responding to student writing. *Tesol Quarterly*, *19*(1), 79-101.

Zhang, S. (1995). Reexamining the affective advantage of peer feedback in the ESL writing class. *Journal of Second Language Writing, 4,* 209-222.

APPENDICES

APPENDIX A

WRITING ASSIGNMENT 3

ARGUMENTATIVE (PERSUASIVE) ESSAY

Writing Prompt

Friendship is important in one's life. Do you believe that it is better to have one or two close friends than to have a large number of casual acquaintances? Write an essay that persuades your readers to agree with your point of view. Use specific reasons and details to support your argument.

Suggested page-length for draft: 1-2 double-spaced pages (5 paragraphs). 12pt. font, 1'' margins

Genre: Academic Argumentative (Persuasive) Essay

Guidelines:

A. The completed Argumentative (Persuasive) essay must address the topic and the task effectively.

B. The introduction of the Argumentative (Persuasive) essay should introduce the topic and explain your opinion or proposition to the audience. Your reader should be able to read only the introduction and know your stand.

C. The body paragraphs of the Argumentative (Persuasive) essay should state your reasons (why do you feel the way you do about the topic) and anticipate the strongest argument your opposition might have (the counter argument). The body paragraphs should provide specific support. These supports may include personal experience, statistics, examples, facts, or experts' opinions. They may be garnered from television shows, magazines, newspapers, textbooks, studies, or interviews.

D. The conclusion of the Argumentative (Persuasive) essay should summarize main arguments and make an appeal for readers to act on your suggestion.

E. The Argumentative (Persuasive) essay should display unity, progression, and coherence.

F. The Argumentative (Persuasive) essay should be free of mechanical, grammatical, and usage errors.

APPENDIX B

RUBRIC FOR WRITING AN ARGUMENTATIVE (PERSUASIVE) ESSAY

	Excellent (4 points)	Good (3 points)	Fair (2 points)	Unacceptable (1 point)
Focus	The response is fully sustained and consistently and purposefully focused: • claims are clearly stated, focused and strongly maintained • alternate or opposing claims are clearly addressed • claims are introduced and communicated clearly within the context	The response is adequately sustained and generally focused: • claims are clear and for the most part maintained, though some loosely related material may be present • context provided for the claims is adequate	The response is somewhat sustained and may have a minor drift in focus: • may be clearly focused on the claims but is insufficiently sustained • claims on the issue may be somewhat unclear and unfocused	The response may be related to the purpose but may offer little relevant detail: • may be very brief • may have a major drift • claims may be confusing or ambiguous
Introduction	Introductory section provides a strong opening, adequate context and a clear thesis statement	Introductory section offers context and a clear thesis statement.	Introductory section contains some context and an unclear thesis statement	Introductory section contains neither context nor a clear thesis statement
Main points (body paragraphs – refutation)	Three or more main points are well developed with supporting details. Refutation paragraph(s) acknowledges the opposing view and summarizes their main points	Three or more main points are present but may lack detail and development in one or two. Refutation paragraph(s) acknowledges the opposing view but doesn't summarize points.	Three or more main points, but all lack development. Refutation paragraph(s) missing and/or vague	Less than three main points, with poor development of ideas. Refutation missing or vague.
Conclusion	Essay provides a concluding statement that summarizes the major points of the argument and explains their significance fully	Essay provides a concluding statement that accurately summarizes the major points of the argument.	Essay provides a concluding statement that is vague or unclear in its summary of the major points of the argument	Essay provides a concluding statement that is unrelated to the argument OR essay does not include a conclusion.
	Claims are well supported with logical reasoning and relevant	Claims are supported with reasoning and evidence,	Evidences used to support the claims are not relevant	Claims are not supported with relevant evidence or

Idea Development	evidence, demonstrating a complete understanding of the topic	demonstrating some understanding of the topic	or not completely thought out	lack evidences.
Organization	organization is readily processed, aided by transitional elements that connect the various components of the essay and clearly express relationships between ideas.	a clear progression of ideas, aided by sufficient transitional elements that connect components of essay and express relationships though there might be a few minor lapses in surface "flow"	an organizational structure that, though largely comprehensible, is difficult to process in places, either through some minor lapses in logic or insufficient transitional elements	multiple lapses in various aspects of organization that undermine comprehension; or a random organization of ideas that defies comprehension
Grammar	absence of stigmatizing errors and virtual absence of surface errors	a few isolated surface errors; possibly a stigmatizing error or two	possibly a stigmatizing error pattern or two and/or some recurring surface errors that prove distracting in places	a few stigmatizing error patterns and a diversity of surface errors that impede understanding in places; or multiple stigmatizing errors and a diversity of surface errors that render the essay exceedingly difficult to process
Punctuation, spelling, and mechanics	Mastery of English writing conventions: margins, capitals, indentation, punctuation and spelling; correct citations	Correct use of English writing conventions: margins, capitals, indentation, punctuation, spelling, citations, with one or two errors	Frequent problems with English writing conventions: margins, capitals, indentation, punctuation, spelling, citations	Serious problems with English writing conventions: margins, capitals, indentation, punctuation, spelling, citations, dominated by errors; unacceptable to educated readers.
Vocabulary/ Word usage	Precise, sophisticated vocabulary usage; concise; appropriate register.	Extensive vocabulary usage; one or two errors of word choice; not wordy; appropriate register.	Limited range of vocabulary; more consistent errors that may impede intelligibility; some evidence of inappropriate register	Very limited range of vocabulary; consistent and frequent errors with words choice or usage impede intelligibility; inappropriate register

APPENDIX C

SURVEY QUESTIONS FOR STUDENTS

Dear participant,

This survey asks your opinions and perceptions of peer assessment utilized in the ESL/EFL Writing for Academic and Professional Success course offered by E-Center. Your answers will be confidential and anonymous. There are no right or wrong answers, so please be open and honest in your responses.

Thank you for taking the time to complete this survey. It is very important for the quality of my research that you answer all the questions honestly and conscientiously.
I appreciate your time and effort.

Most gratefully,

Lan Vu

Please check—or write in, as appropriate--the relevant answer.

1. Your gender: ◦ M ◦ F
2. Your nationality and first language:……..
3. Your educational level:
   - ◦ High school diploma ◦ Some college, no degree
   - ◦ Bachelor's degree ◦ Master's or Doctoral degree
4. Have you taken any online writing courses other than the *ESL/EFL Writing for Academic & Professional Success* course offered by E-Center?
   ◦ Yes ◦ No

If yes, have you ever made comments and graded your peers' papers in your previous online writing classes? ◦ Yes ◦ No

Please rate the degree to which you agree or disagree with the following statements about the peer assessment used in the ESL/EFL Writing for Academic and Professional Success Course offered by E-Center.

5. I participated in peer assessment training and referred to the training instructions and rubric when commenting on and grading my peers' papers.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

6. The guidelines for peer assessment were clearly outlined as to how to grade and to make comments.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

7. The feedback I gave my peers on their writing for this class was useful.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

8. The feedback I gave my peers on their writing for this class was too negative or critical.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

9. For each of the assignments, I followed the guidelines provided during training for commenting on and grading peers' writing.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

10. For each of the assignments, I followed the rubrics for grading peers' writing.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

11. The feedback I gave my peers on their writing for this class was thorough and constructive.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

12. The feedback provided to me by my peers connected very clearly to the standards set forth in the assignment rubrics for the course assignment.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

13. The feedback my peers gave on my writing for this class was too negative and critical.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

14. The feedback I got from one peer was similar to the feedback I got from other peers on the same paper.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

15. Peers gave me fair grades on my writing.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

16. Feedback from peers in this class helped me improve my writing.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

17. I found feedback from my peers helpful, and I used their feedback when revising my writing for this class.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

18. I felt qualified to give feedback and grades on my peers' writing for this class.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

19. I think my peers were not qualified to give feedback and grades on my writing for this class.

	1	2	3	4	
Strongly disagree	◦	◦	◦	◦	Strongly agree

20. Please state in the space below your opinion regarding the following aspects of peer assessment (that is, peers providing the feedback and the grades) in the ESL/EFL Writing for Academic and Professional Success course offered by E-center:
(a): difficulties/challenges you experienced as commentator and grader for your peers
(b) difficulties/challenges you experienced as a writer being commented on and graded by your peers.
(c): peer assessment's usefulness in helping improve your writing performance
(d): similarities or differences among peers' comments and peers' grades

……
……

Thank you for your cooperation!

APPENDIX D
INTERVIEW QUESTIONS FOR STUDENTS

Interview Protocol

I. Interview introduction

The interviewer will start the interview by introducing herself and reading the following statement:

"Thank you for agreeing to participate in my study! The interview will be about your perception and opinions about peer assessment used in ESL/EFL Writing for Academic and Professional Success course. All of your answers will be confidential. There are no right or wrong answers, so please be open and honest in your responses. I will record your answers so that I can analyze them later. During the interview, if you don't feel comfortable, you can skip any individual questions or fully stop the interview at any time."

II. Main Interview Questions:

1. Before you and your peers started to assess each other's writing, you were given instructions on how to grade and how to make comments. Do you think you were well prepared for the peer assessment - with peers providing both the feedback and the grades?
2. When you graded your peers' writing, did you strictly follow the rubric?
3. What did you pay more attention to when grading your peers' writing? Organization? Idea development? Grammar and other sentence-level issues?
4. What do you think about the comments you made on peers' writing? Were they constructive and comprehensive or negative and limited? Could you give an example to illustrate why you see your comments as you do?
5. What do you think about the comments that your peers made on your writing? Were they constructive and comprehensive or negative and limited? Again, could you give an example why you see peers' comments as you do?
6. Do you think your peers gave fair grades?
7. Do you think your writing improved over the period of the ESL/EFL Writing for Academic and Professional Success course? If so, do you see peer assessment—your peers' comments and grading—as having contributed to that improvement? And if so, in what ways did the peers' comments and grades help improve your writing?
8. Now that the course is over, are you more aware of strategies for improving your writing than you were at the outset of the course?
9. If so, what aspects of the course, your own work, or the peer grading or comments do you think helped you become more aware? In other words, what do you think it was about these factors that contributed to a greater awareness? And could you elaborate on the strategies you became more aware of?
10. What difficulties/challenges did you encounter when you assessed your peers' writing – both grading and commenting on peers' papers? How did you deal with these difficulties/challenges?

III. Closing

"Thank you very much for your answers. Your answers will be used for our study purpose only and they all will be erased when the study is completed."

APPENDIX E

CONSENT FORM

Dear E-Center Student:

My name is Lan Vu, a doctoral student at the Department of English, Southern Illinois University, Carbondale. I am currently conducting a study on peer assessment used in an open online setting to complete my PhD degree in Rhetoric and Composition. In this study, I explore aspects of peer assessment utilized in a MOOC-based composition course offered by E-Center for Professional Development.

If you agree to participate in this study, you may be involved in any of its four parts and may withdraw from the study at any time:

1.) Complete the survey regarding your opinions and perceptions of aspects of peer assessment utilized in a MOOC-based composition course offered by E-Center for Professional Development
2.) Allow the researcher to access your writing assignments that you submitted for the course on composition.
3.) Allow the researcher to access the grades and comments you gave to your peers and those your peers gave to you.
4.) Participate in an interview by the researcher on aspects of peer assessment utilized in the MOOC-based composition course offered by E-Center.

By checking "yes" to any of the four levels of participation in the spaces provided in the Informed Consent form below, you will be giving your consent to participate in that portion of the study and consenting to possible use in the researcher's reports of the material collected, with your name removed.

Involvement, then, whatever the level, means agreeing that the researcher may use in research reports your words and comments from the material for the study that you agree to submit for any particular level of involvement. All reports based on this research will maintain the confidentiality of the study's participants through removing the participant's name and identifying each participant's words and comments through a pseudonym. The researcher and the researcher's advisor will be the only individuals to have access to the material with your name attached to your words. All reasonable steps will be taken to protect your identity, and the records of the names and pseudonyms will be destroyed once the study is completed. Participation at all levels of the study is voluntary, and subjects may withdraw from this study at any time.

After I obtain your confirmation for participation in my study, I will contact you two more times in the course of one month if you have not responded. This makes a total of three email contacts, including this one. If you do not respond to the three emails, you should be assured that there

will be no future emails sent to your address and your name will be removed from the pool of potential subject emails.

Be aware that participation in this study is VOLUNTARY and if you change your mind, you may withdraw at any time without hesitation. Moreover, all your responses will be anonymous and confidential. The people who will have access to the data will be myself, the researcher, and my dissertation advisor, Dr. Jane Cogie. After the study is completed, all the records on Quia will be destroyed. We will take all reasonable steps to protect your identity. In the dissertation itself, you will be identified only by the general name: E-Center student. Also, the website does not collect any information that could potentially identify you (such as machine identifiers). There are no known or anticipated risks in participating in this survey.

If you have any questions about this study, please contact the researcher, Lan Vu, Department of English, Southern Illinois University Carbondale, Carbondale, Illinois 62901; Phone: (618)6604386; Email: vulan@siu.edu. and the research's advisor, Dr. Jane Cogie, Associate Professor of English, Southern Illinois University Carbondale, Faner 2283, MC 4503, Carbondale, Illinois 62901; Phone (618) 453-6846; Email: jcogie@siu.edu.

Thank you for your valuable collaboration and assistance in this research.

Best regards,

Lan Vu

This project has been reviewed and approved by the SIUC Human Subjects Committee. Questions concerning your rights as a participant in this research may be addressed to the Committee Chairperson, Office of Sponsored Projects Administration, Southern Illinois University Carbondale, Carbondale, IL 62901-4709. Phone (618) 453-4533. Email: siuhsc@siu.edu

I agree to participate in the study being conducted by Lan Vu under the supervision of Dr. Jane Cogie, Department of English, Southern Illinois University Carbondale. I have made this decision based on the information I have read in the Information-Consent Letter. I realize that I may withdraw without prejudice at anytime.

1.) I agree to complete the survey regarding your opinions and perceptions of aspects of peer assessment utilized in a MOOC-based composition course offered by E-Center for Professional Development
…Yes/No……..

2.) I allow the researcher to access your writing assignments that you submitted for the course on composition.
…Yes/No……..

3.) I allow the researcher to access the grades and comments you gave to your peers and those your peers gave to you.
…Yes/No……..

4.) I agree to participate in an interview by the researcher on aspects of peer assessment utilized in the MOOC-based composition course offered by E-Center.
…Yes/No……..

Signature
……(Print Name)

Click HERE if you have read these statements, understand them, and agree to part

APPENDIX F

RECORDING CONSENT FORM

Dear E-Center learner,

Thank you for agreeing to participate in the survey portion of our study. Because the study aims to examine aspects of peer assessment used in the MOOC-based composition course offered by E-Center through collecting feedback from the students involved, I would like to interview you to find the answers to my research questions. The total time for the interview should be no longer than 30 minutes. If you agree to join my study, please send me your available schedule so that I can set up an appointment for the interview. The interview can be conducted online via Gmail chat or Skype. You can tell me which option you prefer to do. All of the data will be confidential. There are no right or wrong answers, so please be open and honest in your responses. I will record your answers in the interview so that they will be available later for analysis. During the interview, if at any time, you don't feel comfortable, you may stop the interview. Your answers will be used for my study purpose only, and they will all be erased when the study is completed.

Be aware that participation in this study is VOLUNTARY and if you change your mind, you may withdraw at any time without hesitation. Moreover, all your responses will be *kept confidential up to the limits possible.* The people who will have access to the data will be myself as the researcher, and my dissertation advisor, Dr. Jane Cogie. After the study is completed, all the recorded interviews will be destroyed. We will take all reasonable steps to protect your identity. In the dissertation itself, you will be identified only by the general name: E-Center student. There are no known or anticipated risks in participating in this survey.

If you have any questions about this study, please contact the researcher, Lan Vu, Department of English, Southern Illinois University Carbondale, Carbondale, Illinois 62901; Phone: (618)6604386; Email: vulan@siu.edu. and the researcher's advisor, Dr. Jane Cogie, Associate Professor of English, Southern Illinois University Carbondale, Faner 2283, MC 4503, Carbondale, Illinois 62901; Phone (618) 453-6846; Email: jcogie@siu.edu.

Thank you for your valuable collaboration and assistance in this research.

Best regards,

Lan Vu

This project has been reviewed and approved by the SIUC Human Subjects Committee. Questions concerning your rights as a participant in this research may be addressed to the

Committee Chairperson, Office of Sponsored Projects Administration, Southern Illinois University Carbondale, Carbondale, IL 62901-4709. Phone (618) 453-4533. Email: siuhsc@siu.edu

I agree to participate in the study being conducted by Lan Vu under the supervision of Dr. Jane Cogie, Department of English, Southern Illinois University Carbondale. I have made this decision based on the information I have read in the Information-Consent Letter. I realize that I may withdraw without prejudice at anytime.

I agree to participate in an interview by the researcher on aspects of peer assessment utilized in the MOOC-based composition course offered by E-Center. And I allow the researcher to record the interview for research analysis.

Signature
……..(Print Name)

Click HERE if you have read these statements, understand them, and agree to participate.

VITA

Graduate School
Southern Illinois University

Lan T. Vu

vulanlequydon@gmail.com

Quy Nhon University
Bachelor of Foreign Language Teaching, June 2001

Southern Illinois University Carbondale
Master of TESOL & Applied Linguistics, August 2011

Special Honors and Awards:

U.S. Department of State, *Fulbright Student Scholarship* (2009 – 2011)
Service of Education and Training, *Excellent Teaching Performance Award* (2004)
Quy Nhon University, *Excellent Academic Performance Award* (2000)

Dissertation Title:

A Case Study of Peer Assessment in a Mooc-based Composition Course: Students' Perceptions, Peers' Grading Scores versus Instructors' Grading Scores, and Peers' Commentary versus Instructors' Commentary

Major Professors: Lisa McClure and Jane Cogie

Publications:

Vu, P., Vu, L., Cao, V., & Fadde, P. (2014). An Exploratory Study in an Emerging approach to E-learning: MOOC. In M. Limbu & B. Gurung (Eds.), *Emerging pedagogy in the networked knowledge society: Practice integrating social media and globalization* (162-174). IGI Global.

Vu, P., Cao, V., Vu, L., & Capero, J. (2014). Factors driving learner success in online professional development. *International Review of Research in Open and Distance Learning. 15* (3).

Vu, L. (May 2013). Is the TOEFL score a reliable indicator of international graduate students' academic achievement in American higher education? *International Journal on Studies in English Language and Literature (IJSELL). 1*(1).

Vu, P., & Vu, L. (2013). Videotape and reflection -based feedback to student teachers: A case study. *Critical Issues in Teacher Education. 20,* 5-13.

Howard, J., Vu, P., & Vu, L. (2012). How do undergraduate students use their iPads? *Journal of Technology Integration in the Classroom. 4*(3), 5- 12.

Vu, P., & Vu, L. (2012). Teaching English as a foreign language major to gifted students. *The Southeast Asian Journal of English Language Studies. 18*(2), 57- 66.

Vu, P., & Vu, L. (2012). Techniques to bring humor and create a pleasant learning environment in ESL adult classroom. *Journal of Research and Practice for Adult Literacy, Secondary, and Basic Education, 1*(1), 50-53.