The Effects of Gamified Peer Feedback on Student Writing in High School English Language Arts

Kerise Amaris Broome

Follow this and additional works at: https://scholarcommons.sc.edu/etd

Part of the Curriculum and Instruction Commons

Recommended Citation

This Open Access Dissertation is brought to you by Scholar Commons. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact digres@mailbox.sc.edu.
DEDICATION

This work is dedicated to my family: most importantly, my husband Jason, without whom I could have never conquered such a goal, my mother who gave me the vision to pursue higher education, my daughter who inspires me every single day, and finally, my dog, Kaya, who was at my feet for every single word.
ACKNOWLEDGEMENTS

My first acknowledgement is to my alma mater, The University of South Carolina and the School of Education. My journey to obtain this degree has been the most challenging yet rewarding learning experience of my life due to the support and expertise of every professor in the program. Dr. Lucas Lima de Vasconcelos, my dissertation chair, was more than a professor; he was an inspiration, expert, and mentor. As an educator, I will always be inspired by his standard of professionalism and excellence. I would also like to thank my Cohort Qwerty for the camaraderie and Facebook posts that let me know I was not alone in the struggle. Finally, I acknowledge the sacrifices made by my husband, daughter, sister and parents to allow me to pursue this dream.
ABSTRACT

National assessments on writing show that student proficiency is low in part because many high school students lack revision skills. Compounding the problem is that as classroom sizes grow, teachers have less time to give the feedback that students need to improve. Peer feedback shows promise in being able to solve the feedback problem, but it must be carefully structured to be effective. The purpose of this action research was to evaluate the impact of gamified peer feedback in a high school English Language Arts classroom at Southern Charter High School. Three research questions guided this study: (1) What impact does gamified peer feedback have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school? (2) What impact does the type of peer feedback given and received have on their writing improvement? (3) What are the effects of gamified peer feedback on low versus high ability students’ writing improvement?

The study utilized a predominantly quantitative multi-method design approach with 28 student participants in a tenth grade Honors English Language Arts class. Students experienced 12 class periods of a gamified peer feedback innovation in conjunction with writing three essays during the semester. Qualitative data including peer feedback comments and feedback discussion videos were collected in addition to quantitative data from analytic scoring rubrics and analyzed to determine the effect on writing improvement. Results indicated that gamification improved the quality of peer feedback that students
gave and that led to improved writing. Students of lower ability demonstrated more writing improvement than students of higher ability. Implications and recommendations for future research are discussed.

Keywords: peer feedback, peer assessment, revision, writing, gamification.
# TABLE OF CONTENTS

DEDICATION ....................................................................................................................... iii

ACKNOWLEDGEMENTS ...................................................................................................... iv

ABSTRACT .......................................................................................................................... v

LIST OF TABLES .................................................................................................................... x

LIST OF FIGURES ................................................................................................................ xii

LIST OF ABBREVIATIONS ................................................................................................... xiii

CHAPTER 1: Introduction ..................................................................................................... 1

  National Context ................................................................................................................. 1

  Local Context ..................................................................................................................... 4

  Statement of Problem ........................................................................................................ 5

  Purpose Statement ............................................................................................................. 6

  Research Questions .......................................................................................................... 6

  Subjectivity and Positionality ............................................................................................ 8

  Operationalized Definitions ............................................................................................... 11

CHAPTER 2: Review of the Literature .................................................................................. 14

  Method ............................................................................................................................... 14

  Peer Feedback ................................................................................................................... 15

  Theoretical Basis for Peer Feedback .................................................................................. 17

  Benefits of Peer Feedback ............................................................................................... 22

  Structuring Peer Feedback in the Classroom ................................................................. 24

  Peer Feedback Study Methods ........................................................................................ 27
Impact of Technology to Support Peer Feedback ........................................ 29
Game-based Learning and Gamification .................................................... 31
Chapter Summary .................................................................................... 34
CHAPTER 3: Method .................................................................................. 36
Research Design ....................................................................................... 36
Setting and Participants ........................................................................... 37
Innovation .................................................................................................. 38
Data Collection Methods ........................................................................... 44
Data Analysis Methods ............................................................................. 47
Procedures and Timeline .......................................................................... 51
Rigor and Trustworthiness ......................................................................... 51
Plan for Sharing ........................................................................................ 53
CHAPTER 4: Analysis and Findings ............................................................ 55
Qualitative Coding .................................................................................... 55
Quantitative Data Analysis ......................................................................... 65
Chapter Summary ...................................................................................... 78
CHAPTER 5: Discussion, Implications, and Limitations ......................... 80
Discussion ................................................................................................. 80
Implications ............................................................................................. 93
Limitations ............................................................................................... 97
Closing Thoughts ....................................................................................... 99
REFERENCES .......................................................................................... 100
APPENDIX A: Consent Form and IRB approval ...................................... 115
APPENDIX B: Analytic Scoring Rubric for Phase One Technology Essay .... 118
APPENDIX C: Analytic Scoring Rubric for The Kite Runner Essay ............ 119
APPENDIX D: Analytic Scoring Rubric for Research Paper ............................ 120
LIST OF TABLES

Table 3.1 Phase Description of Writing Instruction and Activities ...................... 39
Table 3.2 Research Questions and Data Method Alignment .............................. 46
Table 3.3 Research Questions, Data Collection, and Data Analysis Alignment .. 48
Table 3.4 Phase Implementation and Timeline .............................................. 52
Table 4.1 Peer Feedback Codes by Category, Number, and Type ...................... 64
Table 4.2 Number of Peer Feedback Comments Given by Type ...................... 67
Table 4.3 Grade Means for Each Essay Cycle .............................................. 68
Table 4.4 Shapiro-Wilk Normality Tests—Total Given and Total Received ....... 71
Table 4.5 Shapiro-Wilk Normality Tests—Peer Feedback Type
Given and Received ...................................................................................... 72
Table 4.6 Shapiro-Wilk Normality Tests—Comments Given vs.
Received Correlation .................................................................................... 72
Table 4.7 Levene’s Equality of Variances—Comments Given vs.
Received Correlation .................................................................................... 72
Table 4.8 Shapiro-Wilk Normality Tests—Peer Feedback Type and Grade
Improvement ................................................................................................. 73
Table 4.9 Levene’s Equality of Variances—Peer Feedback Type and
Grade Improvement ...................................................................................... 73
Table 4.10 Shapiro-Wilk Normality Tests—Grade Improvement by
Ability Group ............................................................................................... 74
Table 4.11 Levene’s Equality of Variances—Grade Improvement by
Ability Group ............................................................................................... 74
Table 4.12 Correlation of Total Comments Given vs. Received
And Writing Improvement ............................................................................ 75
Table 4.13 Correlation Between Frequency of Comment Types and Writing Improvement ........................................................................................................................................75

Table 4.14 Grade Improvement by Highest Givers or Receivers .........................76

Table 4.15 Grade Improvement by Feedback Type Given and Received .............77

Table 4.16 Grade Improvement by Ability ................................................................79
LIST OF FIGURES

Figure 1.1 Root Cause Analysis Diagram of Students’ Lack of Revision and Poor Writing Quality ................................................................. 7

Figure 3.1 Teacher-Created Video for Phase Two: Thesis, Thailand ............. 42

Figure 3.2 Students Competing in a Team-Building Challenge ...................... 42

Figure 3.3 Leaderboard with GPF badges .................................................. 45

Figure 4.1 Feedback Discussion Video Transcript and Corresponding Codes ................................................................. 58

Figure 4.2 Post It Notes with Number and Type of Comments Each Student Received ........................................................................... 66

Figure 4.3 Mean Writing Improvement for Each Essay Cycle ......................... 68

Figure 4.4 Google Sheet Calculations of Quantitative Data ......................... 70

Figure 4.5 Grade Improvement Comparison by Group ................................ 78

Figure 4.6 Mean Grade Improvement Comparison Between High and Low Ability Groups ................................................................. 79
LIST OF ABBREVIATIONS

ACT ................................................................. American College Testing
ELA................................................................. English Language Arts
GPF ............................................................... Gamified Peer Feedback
NAEP......................................................... National Assessment of Educational Progress
CHAPTER 1
INTRODUCTION
National Context

In 2011, The National Assessment of Educational Progress (NAEP) administered the first computer-based assessment in writing to 24,000 eighth graders and just over 28,000 twelfth graders (U. S. Department of Education, 2011). The results were grim: only 27% of students, at both grade levels, performed proficient or higher, indicating a writing problem across all levels of high school. Additionally, students of color performed consistently lower than their white counterparts pointing to a clear opportunity gap. Unfortunately, preliminary results of a subsequent writing assessment of eighth graders in 2017 showed an even lower performance (U.S. Department of Education, 2017). Additional analyses are still being performed to determine the effect that using digital tablets versus laptops may have had, but problems surrounding student writing are evident.

Lack of revision amongst developing writers is at least part of the problem. Research indicates lack of revision amongst college freshmen (Dave & Russell, 2010), and a 2013 study (Witte) found that even amongst educators, knowledge about revision was lacking. Compounding the lack of knowledge about revision are the demands for time in the ELA classroom. Writing is a time-consuming process for students and teachers. Multiple cycles of writing and feedback are needed for growth. Research has also shown that the revision
process increases self-regulation and a growth mindset in student (Feltham & Sharen, 2015), but across the country, where student-to-teacher ratios are at a record high (Rich, 2013), teachers cannot provide the amount of feedback their students require for optimal revision.

Although the ratio of students to teachers can pose a workload problem for teachers, students may prove to be part of the solution. Outlining the work done by Vygotsky (Vygotsky & Cole, 1978) regarding the socio-cultural learning theory and zone of proximal development, researchers have argued that writing and revision are cultural learning tasks in which a novice learner can work with more capable learners to build skill and knowledge (Hanjani, 2019). In his work with peer feedback, Topping (2009) found peer feedback superior to teacher feedback in many ways. Subsequent studies have reiterated those benefits (Farrah, 2012; Gielen, Tops, Dochy, Onghena, & Smeets, 2010) and extended them to the digital environment (Kayacan & Razi, 2017) where students can more easily access share their writing and offer feedback.

What we know currently, however, is that peer feedback and peer assessment regarding writing is not intuitive and that students need to be guided through that process so that their feedback is more beneficial (Hanjani, 2019). Early and Saidy (2014) found a positive correlation between participation in revision instruction and peer feedback on the quality of revisions made to their students’ writing. Substantive feedback has been found far superior, however, to surface feedback critical in terms of its impact on writing quality (Saidy & Early, 2016). Substantive refers to writing concerns that are high-level in nature because they are content-related or structural, whereas surface indicates low-level
concerns, mostly grammatical changes to writing (Faigley & Witte, 1981; Wu & Schunn, 2020; Zhang et al., 2017).

Unfortunately, peer feedback can be ineffective in classrooms. If the effect on learning is not evident, Boud and Malloy (2013) argue that it has not occurred at all. Many studies have begun the work of grounding peer feedback in theory, but more needs to be done to make it an effective tool for writing improvement. Some studies have suggested that student attitudes play an important role in its effectiveness (van Zundert et al., 2010), but more research is needed (Farrah, 2012). Although peer feedback is frequently implemented with academic writing tasks in higher education, “a quantitative synthesis is still lacking for the impact that peer feedback has on students’ writing performance” (Huisman et al., 2019, p. 863). Indeed, most of what is known about peer feedback comes from studying it in college classrooms with very little being known about how it works with high school students.

In one of the few studies of peer feedback in the high school setting, Saidy & Early (2016) explain that listening and talking are integral elements of any relationship and therefore integral part to the revision process. Many studies have asserted the importance of fostering relationships among peer writing groups to improve the feedback they give each other’s writing. The digital environment provides the opportunity for a more gamified approach to peer feedback in conjunction with programs that can easily share writing and feedback between students and make revisions. This gamified framework that encourages group interaction for conducting peer feedback aligns with the social aspect emphasized by these studies and is well-suited to the developmental stage of high schoolers.
Local Context

Student writers’ lack of revision and its negative effects on writing quality are apparent at the local level of Southern Charter High School as well. In the context of the high school ELA classroom where revision should be an integral part of writing instruction, most of my sophomores claim minor surface changes as the extent of their revision efforts. Their lack of knowledge about what revision is leads to an absence of skill when asked to improve their first drafts of writing. Year after year, I strive to teach students how to make substantive changes to their essays, but the same pattern persists: students generally write their first draft the night before an essay is due, make minor edits, and then submit it as their final draft for a grade. While I know that more feedback is needed to teach students to make better revisions (Berg, 1999), I struggle to manage more than one evaluation of the 500 or so student essays I assign each year. Offering each student abundant feedback and multiple revision opportunities is simply untenable due to time constraints.

Students’ lack of writing achievement in our classrooms and on standardized assessments is an area of concern for all English teachers at our school. Three years ago, we began a student writing center to address writing deficiencies that we knew existed. After only one year, however, it failed. Teachers did not want to bring their classes during instructional time, and students did not want to come during lunch or after school. When it comes to teaching writing at our school, there is very little consensus among us about what does work. We all use the available technology for writing such as Google Classroom, Google Docs, and online rubrics, but in various ways with unclear results. Unfortunately, the only commonality across our department is an
opportunity gap; it affects our entire school and follows the same racial lines cited in the Nation’s Report Card on Writing (2011). Minority students at our school score persistently lower on college entrance tests such as the American College Testing (ACT) Test. On average, they earn lower writing grades in my classroom. In essence, the lack of writing skills affects many of the students at Southern Charter High School, but it impacts our minority groups more intensely.

Writing improvement is more difficult because it is not a problem that is openly discussed amongst teachers. It may be embarrassing for teachers to admit that our students are not writing to standard, and that we do not have the time to properly address it. The lack of research about peer feedback at the high school level and its absence from our state standards affects why it is not more widely utilized to improve student writing at our school.

**Statement of Problem**

Revision is an important skill reflected in most writing standards, but student-writers require effective feedback to accomplish it. Without it, they equate surface changes with actual revision. In traditional instruction where peer feedback is absent from the skills taught, the teachers is often the sole source of writing feedback. Lack of instructional time per student becomes a significant obstacle with large class and their respective number of essays that multiply with each revision cycle. Time is also a contributing factor on the student end, with most students underestimating the time and skill that good writing takes. Many teachers lament that their students see writing as a one-and-done event rather than seeing it as a recursive process with continual steps toward improvement.
Students need help to make substantive revisions that improve the quality of their writing (U. S. Department of Education, 2011; 2017), but teachers cannot always give the timely feedback that is needed (Topping, 2009). Peer feedback can help, but only if students are taught how to effectively engage in it so that they produce substantive revisions. Indeed, students may be more receptive to advice from their peers (Gielen et al., 2010; Ducasse & Hill, 2019; Topping, 2009), but peer feedback must move beyond surface-level suggestions to help student writers achieve their writing potential. Figure 1.1 below details these and other factors contributing to the students’ lack of revision and poor writing quality.

Effective peer feedback has shown it can improve student revision and writing (Berg, 1999; Huisman et. al., 2019; Landry et. al., 2015), both for the assessor and the assessed (Lundstrom & Baker, 2009), but it is often not intuitive for students. They need guidance through the process so that their feedback results in substantive revisions that are necessary for writing improvement (Hanjani, 2019).

Purpose Statement

The purpose of this action research was to investigate how a gamified peer feedback model in a digital writing environment affects the quality of peer feedback and subsequent writing improvement for tenth grade ELA students at Southern Charter High School.

Research Questions

RQ1: What impact does gamified peer feedback have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school?
Figure 1.1

Root Cause Analysis Diagram of Students’ Lack of Revision and Poor Writing Quality
RQ2: What impact does the type of peer feedback given and received have on their writing improvement?

RQ3: What are the effects of gamified peer feedback on low versus high ability students’ writing improvement?

**Subjectivity and Positionality**

Becoming an English teacher was a natural choice for me. My teachers have always told me that I was a good writer, and getting good grades and positive reinforcement from them helped me enjoy school. As a result, I went into teaching as a career and have taught high school English for 27 years. I have always been especially interested in using technology to engage students, even when I started teaching in 1997 and the technology was much less advanced. I enjoy making learning fun for my high school students and helping students feel successful like I did as a student.

I have always felt that technology made writing easier and more improved. It concerns me if students prefer hand-writing their drafts because I think using a word processor is more conducive to revision. Having continuous access to technology has most likely afforded me this belief. Students are expected to use technology to draft and collaborate on their writing in this study, but not all students feel as comfortable with technology as I do. Peshkin (1998) asserts that one must “seek out their subjectivity” (p. 17) prior to and throughout the research process. In doing this, I must acknowledge what he refers to as the first “I” in my subjectivity. It is the successful student “I” that leads me to assume that all students will enjoy technology, writing and learning like I do, which is a potential bias. Another potential bias stems from my self-confidence in classroom settings, a quality that I know some students lack. This lack of self-
confidence may affect their comfortability with group interactions like peer feedback. I will need to closely observe students and adapt instructional strategies to meet their needs throughout the action research process to counter these potential biases.

I must also be careful of what Herr and Anderson (2014) warn is a “common mistake in this type of research [which] is to treat one’s personal and professional self as an outside observer rather than as an insider committed to the success of the actions under study” (p. 33). My positionality as the teacher makes me uniquely able to report about the effectiveness of the peer feedback method in my context (which they acknowledge is a necessary addition to the body of research), but I must acknowledge my bias in wanting students to be successful. As Zeni (1998) asserts, “we aren’t outsiders peering from the shadows into the classroom, but insiders responsible to the students whose learning we document” (p. 10).

Therefore, like most action researchers, I consider my positionality to be that of an insider who is collaborating with other insiders (Herr & Anderson, 2014), the student participants. This is because I feel comfortable as a facilitator in my classroom, versus being the sole source of knowledge and expertise. My role as facilitator results in many students becoming comfortable as collaborators. Stringer (2013) validates this dynamic underlining participants’ role even in developing research inquiry: “All stakeholders—those whose lives are affected by the problem under study—should be engaged in the processes of investigation” (p. 15). I currently incorporate peer feedback activities and I see students teaching one another very effectively at times. This aligns with my interpretivist paradigm, a view that defines reality through that same lens of social
interactions, claiming that “social reality is different than natural reality because the subjects of social reality are human beings and their relations with each other.” (Tekin & Kotaman, 2013, p. 84). It defies standardized research techniques and allows the researcher and the subject to change through their interaction. It is through this interpretivist lens that I feel the truth about peer feedback—the interactions between my students and in their views about the process—will come into focus, rather than through a limited test of a prescribed hypothesis.

There are some obstacles to the goal of full collaboration between me and my participants. While my positionality as someone who believes in peer feedback seems beneficial to helping students become more comfortable with it, there is a paradox regarding trust between students and teachers that is important to acknowledge. Herr and Anderson (2014) refer to this saying, “power relations in a setting operate even when insiders think they are being collaborative” (p. 36). Honors students are typically very motivated by their grade, and it follows that they typically want to please the teacher, however, there is an inherent conflict with asking for participation in a study and then grading that behavior. I must be cognizant of how my positionality as the person who awards grades could be affecting the outcome of the study. Again, Zeni (1998) asks a critical question here, “Will your research strengthen this trust or perhaps abuse it?” I feel that by the conclusion of the study, students will agree, the trust has been strengthened because of what we learn together.

Nevertheless, I admit that all the positive experiences that I have had with school, writing and technology shape the expectations that I have for students. My pilot studies with peer feedback have shown me that not all students
experience these in the same way. Some students feel uncomfortable working with others, while some feel intimidated by technology or with writing in general. Because of those differences as well as racial, gender, and socio-economic differences, I operate as an outsider to some groups of students. Agee (2009) points to the difficulty that new researchers have with assessing the level of risk for participants, so I must be cautious of all these mitigating factors. Zeni (1998) questions, however, “Will your study attempt to read and interpret the experience of people who differ from you in race, class, gender, ethnicity, sexual orientation or other cultural dimensions” (p. 13)? This speaks to the value in highlighting the differences that students bring to the table.

Ultimately, as an educational technology researcher, I must work to be inclusive, encouraging, and understanding of students who come from different backgrounds and experiences than me. I must use my unique positionality as the manager of classroom technology access and student interactions to include them versus exclude them from the benefits that I see of using peer feedback and technology to become better writers.

**Operationalized Definitions**

In terms of writing instruction in the ELA classroom, **feedback** refers to any comments, assessments, suggestions, questions, or edits given to a student intended to “narrow the gap” between the actual level and the expected level of a performance (Ramprasad, 1983, p. 4). Traditionally, feedback is given by a teacher, however, peer feedback describes when students engage in this process to help each other.

In this research, **gamified peer feedback (GPF)** describes a set of instructional lessons mimicking the popular CBS TV show The Amazing Race
designed to teach students effective peer feedback methods. The teacher will structure fun, physical challenges to build team rapport, and then engage in instruction about how to provide peer feedback based on rubrics for each writing assignment.

**Gamification** refers to the application of game elements to other areas of activity, to increase motivation and engagement. **Game-based learning** occurs when these game characteristics are brought into the classroom with specific learning outcomes in mind (Deterding, Dixon, Klaed, & Nacke, 2011). Toda, Palomino, Oliveira, Rodrigues, Klock, Gasparini, Cristea, & Isotani, (2019) have classified 21 game elements that can be incorporated to make classroom environments more game-like.

The feedback will be classified as either **substantive**, which refers to the writing content based on rubric specifications. Substantive comments refer to high-level, whole document concerns like meaning and organization; they are in contrast to **surface comments**, which characterizes low-level, grammatical concerns (Early & Saidy, 2014). “Substantive revision can help improve the overall quality of a final piece of writing, improve the organization and presentation of ideas, and strengthen a line of argument” (Bridwell, 1980).

A **digital writing environment** (Witte, 2013) refers to the students’ use of computers and Internet applications such Google Classroom and Google Docs software to implement all stages of the writing process (e.g. drafting to revising, editing and finalizing).

**Writing improvement** is defined as the improvement in student writing quality measured by the writing rubrics. These online rubrics are digitally
formatted by the teacher-research using Google Sheets and based on National ACT and South Carolina State ELA Standards. See Appendices B-D.
CHAPTER TWO
REVIEW OF LITERATURE

This action research examined the effects of gamified peer feedback on the quality of peer feedback and subsequent writing for tenth grade ELA students at our school. The review of literature situates the following aspects of the study’s research questions: (1) the potential effect of gamification on the process of peer feedback, (2) how peer feedback can impact writing improvement, (3) how giving versus receiving peer feedback may affect writing improvement and (4) how peer feedback may impact the writing of low versus high ability students.

Method

The search for extant literature began in the electronic educational databases Education Source, ERIC, and JSTOR, and it consisted of several stages over the course of many months. After initial keyword searches for “revision and high school writing” and “high school and peer editing” produced negligible information, I honed the keywords to “peer feedback and writing,” which produced vastly more studies, although many were at the undergraduate rather than high school level. Another drawback with that search criteria, however, was that many of the studies were conducted internationally and with adult students who were learning English as a foreign language—a population far removed from that of high school students writing in their first language. I rejected many of these studies based on the population differences, but I also noticed that some of them referenced seminal researchers in the field. Mindfully, I have included
some of their findings, especially in the cases where theirs was the only study on a given aspect such as using peer editing software technology or writing in a digital environment. Throughout my search, the keyword “high school” as a limiter proved to be restrictive as well, only ever producing a handful of studies. To resource action research specifically, I conducted a specific search of Global Dissertation and Theses database, which yielded several important doctoral studies. The remaining half of my research was found in the literature reviews and bibliographies of these initial articles. These studies and books were all available in either previously mentioned databases or through Google Scholar, apart from two books I obtained with the help of the University’s reference desk. One was sent electronically, and another was mailed to my home address. I felt confident the review was nearing completion when everything I read began to reference the same research.

This literature review is organized into six sections. It begins with a definition of peer feedback. Section two offers theories that support peer feedback and give insight into its efficacy as part of writing instruction. The next section covers the research methods that have been employed by the researchers which will inform my study. In the fourth section, I discuss the literature regarding benefits of peer feedback, followed by the limitations. Structuring peer feedback is discussed in section five, with the last section encompassing the impacts of technology and the potential that gamification offers for peer feedback.

Peer Feedback

Peer feedback is closely tied to assessment. Brew (1999) describes peer assessment (and self-assessment) as students checking knowledge, performance
or skill level against a model answer using criteria provided by the teacher. The focus is on the action of assessing and what that can teach a student engaging in the process. Graesser, Person, and Magliano (1995) argue this well stating, “Students are engaging with criteria and standards, and applying them to make judgements (p. 497). Training in peer assessment seeks to develop this capability of asking these types of questions which has learning benefits for the student who is assessing.

Peer feedback rests on the seminal research of Richard Topping and Nancy Falchikov who began discussing it in the 1980s. Topping (2009) connected feedback to assessment stating that when assessment is shared with the author of the product, it becomes beneficial for the learner as well and thus becomes feedback. When it was formative rather than summative, he argued, its intent is to help students identify their strengths and weaknesses, target areas for remediation, and develop metacognitive and interpersonal skills. He uses the terms interchangeably: “Peer feedback is available in greater volume and with greater immediacy than teacher feedback. A peer assessor with less skill at assessment but more time in which to do it can produce an assessment of equal reliability and validity to that of a teacher” (p. 253).

Others have reiterated the connection between formative assessment and feedback stating the purpose of formative assessment is to provide feedback to the learner, informing them of strengths and weaknesses and illuminating areas for focus during their studies rather than at the end (Boud, 1990; Topping, 1998). In their meta-analysis, Hattie and Gann (2007) define feedback as information provided by a teacher or other agent regarding aspects of one’s performance. Falchikov (2005) echoes that it is a process in which learners obtain information
about their work so they may better compare the standards for the work to the qualities of the work itself. The last step, generating improved work, being what she coined closing the feedback loop.

In terms how feedback pertains to writing instruction, specifically, Ramaprasad (1983) defines it as any comments, assessments, suggestions, questions, or edits given to a student intended to narrow the gap between the actual level and the expected level of a writing performance. Again, peer assessment is included, and it is appropriately characterized as being formative. Because students may perceive peer feedback negatively if they feel their grade is being determined by a peer (Kaufman & Schunn, 2011), it is important to establish that peer assessment, by definition, should only occur for the sake of feedback rather than summative assessment (Boud, 2000). Reynolds and Trehan (2000) validate this arguing that it is based on granting or withholding qualifications—a a situation of power relations that should not exist between students (p. 298). Peer feedback, however, may include additional elements such as closing the loop which is echoed in Boud’s (2015) definition as well. For this study, peer feedback for writing will be defined as comments, questions, edits, and assessments given by a peer intended to improve the level of writing in relationship to a set of standards which results in improvement.

**Theoretical Basis for Peer Feedback**

Peer feedback research is grounded by writing and revision theory in addition to well-accepted learning theory established by Vygotsky (Vygotsky & Cole, 1978). Revision theory began to evolve in the 1970s from a view of making small editorial changes (a product-related view) to a more process-related one (Fitzgerald, 1987). Murray’s (1978) seminal work on revision established that it
was an important part of writing theory and that alterations in a text should involve not only minor editorial changes but also reconceptualization of idea and meaning. In 1983, Humes discredited writing as a linear process which was followed by Flower and Hayes (1981) recursive theory of writing. Others would reiterate that revision was indeed viewed as both surface and meaning-based, and both microstructure and macrostructure-related (macrostructure referring to something that would change the summary of a text) (Kintsch & van Dijk, 1978).

Bridwell (1980) argued that substantive revision can help improve the overall quality of a piece of writing such as organization and strengthening a line of argument. Bartlett (1982) defined substantive revision as diagnosing problems in writing and executing solutions. Fitzgerald (1987) echoed these larger revisions referencing how writers take an original draft through a series of changes to identify differences between what they intended to say and how it may be perceived. These deeper types of revisions that more experienced writers make contrast with what Faigley and Witte (1981) coined surface revisions. These two researchers also established the use of qualitative analysis of revision changes to examine the complex process of revision that later studies have followed.

Other researchers have classified feedback, similarly, labeling them high-level criticism comments and low-level criticism comments respectively (Wu & Schunn, 2020; Zhang et. al, 2017). “High-level comments were related to thesis, argument, rhetorical strategies, evidence for claims, explaining evidence, and organization, while low-level comments involved control of language and conventions” (Wu & Schunn, 2020, p. 6). Zhang et al. (2017) used similar terms: citing high-level and low-level writing goals. McGarrell and Verbeem (2007)
reiterate this demarcation in their argument that teachers should “encourage substantial revision” by emphasizing content over form (p. 229). Feedback in this study will be classified as either substantive, which refers to the writing content that is more global and content-relevant, versus surface, which characterizes errors, often grammatical, made at the sentence level (Early & Saidy, 2014).

Peer feedback has been found to increase the quantity and quality of substantive revisions, which aligns with writing theory in this and many other ways (DeMartino, 2017; Early & Saidy, 2014; McGarrell & Verbeem, 2007). Early and Saidy’s research is especially relevant because it sampled a class of sophomore students in high school, some of whom came from disadvantaged backgrounds, and found that a 3-day workshop on peer feedback clearly resulted in more substantive revisions. Later research has supported the idea that more is not always better when it comes to peer feedback, and that often brief use of it can have positive benefits.

It is also rooted in writing theory in the way that it helps students begin to see writing as the iterative, collaborative process that it is (Kayacan, & Razi, 2017; Saidy & Early, 2016; Santangelo, Harris, & Graham, 2007). For example, several studies highlight the dialogic and social nature of the most effective peer feedback, all of which point to the process-nature of good writing (as opposed to an isolated, one-and-done act (Espasa, Guasch, Mayordomo, Martinez-Melo, & Carless, 2018; Yim, Warschauer, Zheng, & Lawrence, 2014). Thus, it seems that involving students in peer feedback helps them understand the communicative purpose of their writing and that deep revision must occur for writing to effectively convey meaning.

Peer feedback is also validated by Vygotsky’s socio-cultural learning
theory (Muhammed, 2012; Saidy & Early, 2016) and zone of proximal development work (Lundstrom & Baker; 2009, Nassaji & Swain; 2000, Zhu, 1994; Vygotsky & Cole, 1978). Säljö (2010) also echoes the theory that learning is a social process, that social interaction plays an integral role in developing cognition, and that there are various learning communities in addition to that of an adult or teacher:

Vygotskian ideas have become important for approaches to teaching and learning that emphasize the situated nature of knowing and the necessity of actively engaging learners in the process of appropriating knowledge and skills in cooperation with adults and in various kinds of learning communities (Säljö, 2010, p. 502).

Studies indicate it is the social aspect of peer feedback, such as its dialogic nature and its superiority to self-assessment, that make it so effective (Ducasse & Hill, 2019; Espasa, et. al., 2018; Huisman et al., 2019). Gan and Hattie (2014) argued that we need to see peer feedback as negotiating meaning and connecting ideas—skills clearly tied to making substantive revisions.

The other tenet of Vygotskian theory, the zone of proximal development, may offer further insight into why peer feedback is effective. ZPD is defined as "the distance between the actual developmental level as determined by independent problem solving and the level of potential development as determined through problem-solving under adult guidance, or in collaboration with more capable peers" (Vygotsky & Cole, 1978, p. 86). With students’ potential development in a learning situation being connected to their ability level, the implication is that that they may learn more from a more capable peer than from someone who is too advanced in skill. Heard (2013)
echoes this in reference to peer feedback stating that studying writers with somewhat similar abilities invites students to see the potential for writing improvement that may not be evident when reading only expert texts. Cho, Schunn, & Wilson (2006) point out that teachers may be so intellectually distanced from low performing students that the validity of their assessment is reduced.

Indeed, research has found that students often perceive the feedback they receive from peers as more understandable and helpful than teacher feedback, because it is written in a more accessible language (Falchikov, 2005; Patchun & Schunn, 2016; Topping, 1998). Perhaps even more illuminating, while peer feedback has been found effective for various student groups, low ability writers demonstrate the most growth (Landry, Jacobs, & Newton, 2015; Lundstrom & Baker, 2009; Strijbos, Narciss, & Dünnebier, 2010).

Research about why peer feedback effectiveness may be connected to ability is incomplete. For example, we know that the quality of peer feedback and revisions is better for high ability writers than low ability writers (Faigley & Witte, 1981; Noorizi, Biemans, & Mulder, 2016; Patchun & Schunn, 2016), but still, high ability writers give equal credence to the suggestions from their low ability peers (Patchun & Schunn, 2016). Interestingly, and this harkens specifically to ZPD theory, that low ability writers were more likely to improve their essay using feedback from other low ability reviewers than from high ability reviewers (Patchun & Schunn, 2016, p. 251). Most studies cite writing improvement for all student groups whether they are high or low ability (DeMartino, 2017; Pham, Huyen, & Nguyen, 2020) and the lesser degree of improvement for high ability writers may simply be explained by the law of
diminishing returns. While these findings are intriguing, research about effectiveness of peer feedback for students of varying ability levels is ongoing.

**Benefits of Peer Feedback**

There is overwhelming evidence that peer feedback provides a wide range of benefits which is comparable to self-assessment and teacher feedback for several reasons. The most convincing evidence for this comes from meta-analyses conducted in the field. Hattie and Temperley (2007) cite 196 studies to determine that feedback (granted this includes all feedback, not just peer feedback) ranked in the top five influences on student learning alongside factors such as direct instruction and student’s prior cognitive ability.

Peer feedback may be especially valuable, however, because it provides feedback that teachers, due to time constraints, cannot (Boud, & Molloy, 2013; Chanski, & Ellis, 2017; Topping, 2009). Although they cited a lack of quantitative controlled studies overall, Huisman et al. (2019) synthesized 24 quantitative studies revealing that students’ involvement in peer feedback resulted in larger writing improvements compared to no-feedback controls and compared to self-assessment. They also found that peer feedback and teacher feedback resulted in similar writing improvements. Furthermore, students who engage in peer feedback consistently demonstrate better writing performance (Berg, 1999; O’Donovan et al., 2004) and reduce their language errors more than students who engaged in self-editing (Mawlawi, 2010).

Berg (1999) provides one of the few quantitative studies where a treatment group who received peer feedback was compared with a control group, although his participants were ESL students. Even in the face of limited writing instruction, frequent use of peer review resulted in academic writing growth
(DeMartino, 2017; Gielen et al., 2015). In their quantitative analysis of a large group of undergraduate student writers, Zhang et al. (2017) found most revision changes are attributable to peer feedback. In a comprehensive study using a large population of undergraduate writers, Zhang et al. (2017) employed a multi-method approach to investigate what affects peer feedback and whether peer feedback affects revision. Peer comments alone appeared to make the largest contributions to revision, and high-level revision improved the most with peer feedback (Zhang et al., 2017).

An added advantage of peer feedback is that it has clear benefits for the assessor, or individuals providing the feedback (Chanski & Ellis, 2017; Huisman, Saab, Van Driel & Van Den Broek, 2018; van Popta, Kral, Camp, Martens, & Simons, 2017). Involvement in peer feedback improves self-efficacy and fosters a growth mindset (Boud & Molloy, 2013; Feltham & Sharen, 2015; Moore & Teather, 2013). Bandura (1977) defined self-efficacy as people’s beliefs about their capabilities to produce designated levels of performance which he found greatly impacted their willingness to initiate and persist in the face of obstacles such as writing. Growth mindset has been linked to students’ willingness to see writing as an evolving rather than fixed skill as well as their ability to revise (Feltham & Sharen, 2015), a principle of revision theory. Boud (1995) emphasizes that self-assessment and peer-assessment are inextricably bound because we self-assess ourselves by the standards that others would use to assess us.

Also remarkable, students seem to base their peer feedback on their own writing product (Ohlsson, 1996; Siegler, 2002; van Popta et al., 2017), which initiates the critical thinking needed for writing improvement and it familiarizes them with the identification of standards and the criteria representing these
standards (Liu & Carless, 2006, Nicol, Thomson, and Breslin, 2014; Sluijsmans, Brand-Gruwel, & van Merriënboer, 2002). Liu and Carless (2006) argue that all actions which help to increase the transparency of assessment to students aid in their understanding. Therefore, it seems that having students provide peer feedback is a beneficial to their own writing growth.

It is important to point out, however, that these benefits have not been substantiated in every study. Peer feedback does not always improve writing quality among first year undergraduate writing students (Li, Liu, & Steckelberg, 2010; Zellner, 2017), although Zellner’s study, which was limited to asynchronous online interaction, indicated there was probably a need for more training. Even with extensive training with a narrative marking guide, secondary students in an international school did not manage to provide feedback pertaining to large-scale, substantive revision (Singh & Hoon, 2016). Additionally, the improvement associated with peer feedback diminishes for writing that is already of high quality (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010), and students may distrust the process if they feel their grade is being determined by a peer (Falchikov, 2005; Liu & Carless, 2006; Kaufman & Schunn 2011; Planas et al., 2014;). With these caveats in mind, the research on structuring peer feedback to maximize its rewards becomes especially salient.

**Structuring Peer Feedback in the Classroom**

Multiple studies have shown that structuring peer feedback and training students in how to give it can be critical to its success in the classroom. Training in the peer feedback process increases both confidence and writing improvement (McConlogue, 2015; Simmons, 2019; van Zundert et al., 2010) and can eliminate anxiety (Falchikov, 2005). In terms of structuring peer feedback, both the
quantity and quality of peer comments appear to be important (Wu & Schunn, 2020). Localization, connecting a particular comment to the relevant part of the text, has been found to be significant (Nelson & Schunn, 2009; Nguyen, Xiong, & Litman, 2017). Cognitive-based comments tied to meaning in the text versus affective comments such as praise have been shown to be more effective (Cheng, Liang, & Tsai, 2015; Patchun, Schunn, & Correnti, 2016). While some pointed to the value of praise (Landy, 2015; Patchun et al., 2016), especially for less mature writers who may improve in their understanding of good writing (Cho & Cho, 2011; Simmons, 2019), Others point to how it can also diminish revision quality (Cheng et al., 2015; Patchun et al., 2016). Involving students in the creation of assessment criteria may also be an important step in helping them to internalize and understand the standards upon which they are being judged (Carless, 2006; Falchikov, 2005). The assessment task should be familiar yet challenging to the participants and require critical analysis (Ballantyne, Hughes, & Mylonas, 2002). Boud (2015) reiterates these principles in stating that trust must exist between the giver and receiver, that there should be dialogue, so that the learner will invest the required time and effort to act on the information given and “appreciate the standards being applied” (p. 2).

Boud’s (2015) perspective relates a concept known as dialogic feedback, which has garnered much attention in peer feedback research. Dialogic feedback is constructed as an ongoing conversation between teacher, learners, and peers and holds the greatest potential for improving learning (Zhu & Carless, 2018; Espasa et al., 2018). Although dialogic interaction can occur asynchronously, it has been found especially critical in online environments with several studies pointing to the importance of face-to-face dialogue in the process of peer
feedback specifically (Crossman & Kite, 2012; Saidy & Early, 2016; Zhu & Carless, 2018). Highly dialogic feedback is obtained when feedback is incorporated throughout an iterative writing process and this has benefits for the quality of peer feedback and for quality revision (Graff, 2009; Espasa et al., 2018).

Nelson and Schunn (2009) argue that student understanding of comments given was the single most important factor in whether feedback was implemented. Involving the writer in the process by having them request their own feedback provides a focus for improvement, decreases the expectation that a feedback giver should correct everything, and increases the likelihood the feedback will be implemented (Ducasse & Hill, 2019). In their argument of a more sustainable feedback model rather than teacher feedback, Boud and Malloy (2013) state that students must be engaged and positioned to both create and request their own feedback while involved in aspects of peer assessment. Zhu and Carless (2018) emphasize that “timely in-class dialogue about written comments can enable the exchange of views, negotiation of meaning and enhance learning efficiency [the result of that being] greater engagement and feedback [that] tended to focus on content and argumentation” (p. 890).

Techniques, like scripting the process, are one way to do this and has been found to produce the highest quality peer feedback (Gan & Hattie, 2014; Gielen & De Wever, 2015; Noroozi, et al., 2016). In a longitudinal action research with high school students, DeMartino (2017) found that peer feedback was of higher quality when driven by teacher prompts about quality versus quantity. Requiring that students close the feedback loop by revising their work is related to the dialogic aspect and is a crucial last step according to many researchers (Boud & Molloy, 2013; Ducasse, & Hill, 2019; Wu & Schunn, 2020).
There are several cautions about peer feedback, however, including that teachers should not over-structure it, and that it should be fun. Several studies find that student authorship of feedback versus prescribed revision strategies produced the most substantive revision (Early & Saidy, 2014; Nicol et al., 2014). The law of diminishing returns also applies with multiple studies citing student fatigue with the process after time (Mulder, Pearce, Baik, 2014; Planas Llado et al., 2014; Zou et al., 2018). Other aspects, such as the sensitive nature of assessment in general, point to the care that must be taken to balance peer feedback benefits and limitations—areas where technology and gamification may offer some advantage.

Peer Feedback Study Methods

Peer feedback research spans the past four decades, and although the vast majority is based in undergraduate environments with convenience sampling (Huisman et al., 2019), it is balanced in terms of qualitative and quantitative methods and has focused on two aspects of the process—its effect on writing and students’ perceptions about it. There is a lack of studies, however, that use a control group for comparison (Huisman et al., 2019).

The effect of peer feedback on writing has been largely studied through qualitative coding of revision changes and quantitative scoring of essays using rubrics. Early and Saidy (2014) used qualitative coding of revision suggestions to track the quality and the subsequent revision that occurred. Kayacan and Razi (2017) used writing rubrics for quantitative results of whether their writing improved and open-ended questions for qualitative results of whether the process was beneficial. In one of a few studies to employ a control group, Gielen and DeWever (2015) used a rubric to quantify the peer feedback they received
from three groups. One group used a checklist, another employed a feedback request, and the last one was a control group with no structure. They found mixed results: while structuring did appear to have favorable results for the quality of peer feedback, it was only for the initial stage of the experiment and did not extend through the entire experiment.

Employing qualitative methods such as observation and interview, as well as quantitative survey data, the research on student perceptions is also complex and varied. Many studies focus on the ability of peer feedback experiences to positively influence attitudes toward peer feedback (Singh & Hoon, 2016). Survey data from several studies was used to study the relationship between students’ perceptions and their performance in the peer feedback models (Mulder et al., 2014; Planas et al., 2014; Zou et al., 2018).

The body of research also includes several action research and meta-analyses that offer salient points of consideration for this study. Hovardas et al. (2014) in addition to Chanski & Ellis (2017) represent action research that established how peer feedback is beneficial for students in secondary writing classrooms. Hovardas et al. (2014) used mixed-methods action research to examine the impact of peer feedback on student revisions. Like my proposed study, grades supplied the quantitative data while peer feedback comments provided qualitative data; Both studies pointed to some discrepancies between middle school science peer assessors and experts, but the overall accuracy and quality of the peer assessments given the age of the assessors was encouraging. It is important to note that the population studied was middle school students, which is the youngest population of all the reviewed studies.

Action research provides several relevant features as well. Ballantyne et al.
(2002) conducted a longitudinal study with over 1600 students and staff that demonstrated its iterative nature, how the study evolved between phases in response to student and staff needs. DeMartino (2017) was another action research study conducted in a high school setting and was one of the few longitudinal studies, following students from the 9th grade to 12th. In the end, however, it proved difficult to follow students through different English class placements and the resulting lack of qualitative data negated the overall benefit to such a study, emphasizing the importance of data collection rather than length of time. Meta-analyses are another type of study that reveal important commonalities in the field such the importance of training (van Zundert et al., 2010) and the positive impact of peer feedback on writing improvement (Huisman et al., 2019).

**Impact of Technology to Support Peer Feedback**

Technological innovations have made the peer feedback process easier, but the development of gamification has been argued to be the most important educational development of the 21st century (Van Eck, 2015), so it may hold the greatest potential for integrating complex concepts like peer feedback into high school classroom. Research also indicates, however, that gamification may actually deter the learning process if done ineffectively (Dicheva & Dicheva, 2017), thus there are many points of consideration. First, teachers should not take for granted that students are inherently knowledgeable and comfortable with technology, especially as it pertains to writing. The prediction that 21st century students would be entering our classroom as digital natives has been altered by the reality that while they easily adapt to new technology, they have more experience with social media than with video games and can be surprisingly
awkward and naïve in online interactions (Van Eck, 2015). Case in point, students have found it easier to revise on printed copy rather than on a computer screen (Dave & Russell, 2010), although any writing teacher knows that writing and revision is much more effective when done on computers. Considerations regarding students’ digital competency must be weighed against the outward facility of technology when designing effective peer feedback instruction.

Peergrade, Peerceptiv, PeerStudio, Turnitin.com, and Google Docs are all examples of writing and peer feedback technology that automate the peer feedback process. These platforms provide improvements such as allowing students to review writing asynchronously via computer, enter comments digitally and often anonymously, and some programs that review the peer feedback for quality. Natural Language Processing and Machine Learning have been used to build models for automated peer feedback assessment which can increase its desirable properties, but any perception that the computer is supposed to do the work may also make it too easy for writers to avoid substantive revisions (Nguyen et al., 2017). Automating the process makes it feasible in large classes (Mostert & Snowball, 2012), and the ease that programs like PeerStudio provide in enabling rapid feedback has proved key in fostering revisions (Kulkarni, Bernstein, & Klemmer, 2015). Several studies have cited students’ positive perceptions of peer-feedback in online environments (Kayacan & Razi, 2017; Zhang et al., 2017) and show writing improvement in online peer feedback environments versus traditional ones (Ducasse & Hill, 2019; Gielen & DeWever, 2015; Novakovich, 2016). This may be due in part to students’ proclivity with social media. The research, however, is multifaceted. While the use of a blogging site increased both the quality of peer feedback and writing
performance over traditional methods (Novakovich, 2016), online forums may present obstacles like deterring feedback discussions (Cartney, 2010). So, while the online environment seems suitable for peer feedback because of its convenience, care must be exercised to support the building of trust and communication that are foundations of peer feedback.

**Game-Based Learning and Gamification**

As with all classroom initiatives, another obstacle for success is student engagement and participation, especially when implementing something as complex as peer feedback. Students may not take the task seriously or be fully engaged in peer feedback (McConlogue, 2015), so games may provide a more suitable method for developing such skills especially because they often increase student engagement and motivation (Ryan & Deci, 2000; Sailer & Homner, 2020). Games also offer a way to foster relationships among students (Nash & Brady, 2022), which studies point out is integral when soliciting peer feedback (Saidy & Early, 2016).

**Game-Based Learning**

Definitions of game-based learning characterize it as a type of game play with defined learning outcomes (Shaffer, Halverson, Squire, & Gee, 2005). Games have been incorporated in ELA classrooms in many ways for several decades, but the advent of technology has complicated their use. Games in the 21st century mean digital games. With the Pew Research Center citing that fully 97% of teenagers play digital games (Teens, Video Games and Civics, 2023), it is easy to understand why teachers may want to use them to engage students, but also be skeptical about the amount of instructional time they garner (Lenhart, Kahne, Middaugh, Macgill, Evans, & Vitak, 2008; Nash & Brady, 2022). Video games are
growing in popularity among ELA teachers who favor them for their ability to teach traditional skills such as reading, writing, and literary analysis, but also 21st-century skills such as argument, analytical thinking and collaboration (Lawrence & Sherry, 2021; Nash & Brady, 2022). Gerber, Abrams, Onwuegbuzie, & Benge, C. L. (2014) cited the way in which one high school ELA classroom engaged students in offering peer feedback, although it was related to the creation of a video game rather than writing. English educators are only just exploring the ways that these games can affect student learning, and overall published studies remain nascent (Nash & Brady, 2022).

**Gamification**

Gamification, on the other hand, is defined by Deterding et al. (2011) as the use of game design elements in non-game contexts. This can be done in the classroom to change traditional classrooms into fun and game-like environments to motivate and changing learner behaviors. Out of a concern for instructional value, however, educators should not focus solely on surface-level design features such as leaderboards, experience points, and badges without regard for the contextual interplay of player, game, and narrative (van Eck, 2015). The positive effects of gamification as well as digital game-based learning (DGBL) are based on well-established learning such as situated cognition, play theory, assimilation, and accommodation (van Eck, 2006). There are many elements of game design that can be incorporated into instruction, but in a taxonomy of 21 game design elements, Toda et al. (2019) caution that the increasing number of elements can confuse and drive away designers and educators who wish to gamify their learning activities. Therefore, I will focus on the elements of
storytelling, challenge, and performance in the Amazing Race innovation to motivate students to give peer feedback.

Storytelling is perhaps the most obvious element of gamification as evidenced by its inclusion in almost every discussion of gamification one encounters. It is defined as “The way the story of the game is told (as a script) within the game, via text, voice, or sensorial resources; examples/synonyms: stories told through animated scenes, audio queues or in-game text queues.” (Toda et al., 2019, p. 52). Kapp (2012) labels this as an abstraction of reality taking place in a narrowly defined game space. Narrative context (Reeves & Read, 2010) is another label used to describe the unique quality of games that places the player in a fictionalized situation where they must build their skills.

Challenge is the next game element that can increase motivation for learning. According to Malone and Lepper (1987), challenge involves activities that are neither impossibly hard nor trivially easy so that a participant’s intrinsic motivation is aroused. What is also required is that meaningful goals are established and though attainment is uncertain, clear feedback about one’s progress toward the goals is provided (Malone and Lepper, 1987). The social aspects of competition and cooperation are cited as elements that build the game environment as well (Toda et al., 2019), but in the case of this game, they align with challenge as defined by Malone and Lepper (1987) because they contribute to both the difficulty and reward placed on players to win the game as a team.

The last game element salient for discussion here is that of performance. For players or participants to maintain high levels of engagement in the game, they must receive regular feedback about their progress toward the goal (Reeves & Reed, 2010). That feedback may come in terms of points, badges, or levels
attained. Researchers warn, however, that any reward must represent achievement in the game and that if players feel the challenge is too easy, their motivation will be diminished. This explains why the use of badges without meaning is detrimental to gamification goals (Van Eck, 2015). Malone and Lepper argue that recognition, making performance information visible to others in the form of leaderboards and badges, increases motivation as well (1987).

Building on Deterding et al.’s well-accepted definition and the advice of educational researchers that games must keep the fun intact, I define gamification as the application of game design to non-game applications to make them more engaging and nudge participants to perform certain actions. Gamification in the context of learning can be referred to as gamified learning (Sailer & Homner, 2020), so this study’s innovation will be termed gamified peer feedback (GPF).

Chapter Summary

Peer feedback has been born out of the push for more formative, more sustainable assessment methods for students, and this connection to student improvement means that a comprehensive definition must include closing the feedback loop and having students demonstrate writing growth. Peer feedback has been studied extensively in the past thirty years through a balance of qualitative and quantitative methods, although a lack of control groups has been cited. The beneficial findings for peer feedback are extensive, however, ranging from benefits for the assessor to higher writing performances for assesses despite their ability or their perception of the process. There are a handful of researchers that caution of its limitations, however, including the difficulty in producing substantive revision and improving writing that is already of high quality,
students’ discomfort with assessment in general, and isolated studies that argue benefits for writing cannot be substantiated at all.

One consistency among the studies, however, was the positive effect that training has on the quality of peer feedback and writing. Finally, technology has improved peer feedback methods and results, but gamification may have the potential to make the most impact-motivating students about the process to obtain powerful learning outcomes. There are clear guidelines, however, about the incorporation of specific game elements into instruction for it to be effective. Van Eck (2015) argues that today’s students come with a level of expectation for engagement in our classrooms, but despite the potential game-based learning offers, there are pitfalls like oversimplifying it that teachers must avoid for it to be effective. Gamified peer feedback (GPF) offered in this study has been created with these caveats in mind.
CHAPTER 3

METHOD

This action research investigated how GPF affects the quality of peer feedback and writing improvement of tenth grade ELA students at Southern Charter High School. The research questions were: (1): What impact does GPF have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school? (2) What impact does the type of peer feedback given and received have on their writing improvement? (3) What are the effects of gamified peer feedback on low versus high ability students’ writing improvement?

Research Design

Action research provides a model for classroom teachers to improve their own practice and solve problems within a local context (Mertler, 2014; Stringer, 2008). Unlike traditional quantitative research which requires a control group, action research allows the researcher to engage all participants in an instructional opportunity in their natural classroom environment. Although this can limit generalizability, action research is very relevant to improving environments in which we work (Stringer, 2013). It should also have benefits for marginalized groups in addition to simply improving educational practice (Kemmis, 2006; Stringer, 2008). An action research design was employed in this study because it allowed for the inclusion all students in what may have proven an effective innovation to improve their writing. It also worked toward closing an
opportunity gap that exists for students of color in our school by examining how the innovation affects students of lower writing ability. Action research design “is especially advantageous when a researcher is building a new instrument” (Creswell, 2009, p. 212), in this case, GPF. Described more specifically as a predominantly quantitative multi-method design, this study triangulated qualitative data about the types of peer feedback students produce with quantitative measures of their writing improvement. The iterative nature of action research and the inclusion of both qualitative and quantitative data provided multiple points of insight about the innovation’s effectiveness.

Setting and Participants

This study occurred in a tenth grade Honors English class at Southern Charter High School, a public high school in Charleston, South Carolina. Our school enrollment is 1570 students, 70% White students, 21% Black, 5% Hispanic, 3% two or more races, 1% Asian, .3% American Indian/Alaskan Native, and .2% Hawaiian Native/Pacific Islander. Our school is one of the top-performing high schools within the very competitive district of Charleston County. As the district’s only International Baccalaureate High School, we enroll students from throughout the district who apply to come to our school. While there is a significant number of affluent students in our school, approximately 25% of our students qualify for free/reduced lunch.

Around that same percentage of students report that they do not have technology such as computers and Wi-Fi at home. While many of our students do very well, and most of our students graduate, our students of color, predominantly black, exhibit an opportunity gap in terms of engagement in higher level courses and ultimately college readiness.
The participants for this study consisted of a purposeful sample of all students enrolled in my 10th grade Honors English class, all of whom agreed to participate in the study, a total of 28 students. These students ranged in age from 14 to 16 years of age. Nineteen of the students identified as female, while the remaining nine students identified as male. Demographics of honors classes differ from the demographics of the school. Typically, the number of white students in an honors class is higher than the school average, about 87%, while African American students make up only about 11% of the class total. This class fell within similar demographic ranges, with 24 being Caucasian (86%), 3 being African American (11%) and one student being Hispanic. Four students met the requirements for receiving special education services for disabilities including ADHD and Autism among other disabilities. All students in this class spoke English as their first language. All students in my English class received the innovation for the study because it is part of a writing instruction program that I had been developing for several years. If students did not give their consent to participate in the study, their information and grades would have been removed from the results, but they would still have engaged in the instructional process as part of the class requirements.

**Innovation**

Development of this innovation began several years ago as I looked for ways to improve student writing. I have always loved the TV show The Amazing Race and wanted to incorporate it in my classroom. The Amazing Race is a TV show where teams compete in various challenges as they race around the world. Each time participants get to a new city, they race to find a clue that describes a physical or mental challenge they must complete there. Each episode
ends with how teams are ranked based on their performance in that city. I thought it could work to put students in teams and have them simulate this idea of travelling around the world with the challenges being to write papers and give peer feedback.

After students gave their peer feedback comments, they revised their essays and submitted them for grading. This cycle of watching the video, completing the team-building challenge, receiving writing instruction, offering peer feedback and revising their papers was repeated over the three cycles of essays. Phase One was Revision, Russia which corresponded to the first essay, a 600-word argument essay on the topic of recent Internet technology and how it affects society. Phase Two, Thesis, Thailand, was an 800-word literary analysis essay on Khaled Hosseini’s book named *The Kite Runner*. Phase Three, Organization, Oregon, was a 2,000-word research paper on a relevant social or environmental issue.

Table 3.1
Phase Description of Writing Instruction and Activities

<table>
<thead>
<tr>
<th>Phase/Writing Focus</th>
<th>Essay Type</th>
<th>Team Building Activity</th>
<th>Writing/Revision Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phase One: Revision, Russia</td>
<td>600-word essay about technology</td>
<td>• Scavenger Hunt in library for clue</td>
<td>• Instructional module on Revision and quiz&lt;br&gt;• Give peer feedback comments and discuss&lt;br&gt;• Revise essay after feedback</td>
</tr>
<tr>
<td>Phase/Writing Focus</td>
<td>Essay Type</td>
<td>Team Building Activity</td>
<td>Writing/Revision Instruction</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------------------------------------------</td>
<td>---------------------------------</td>
<td>------------------------------</td>
</tr>
<tr>
<td>Phase Two: Thesis, Thailand</td>
<td>800-word literary analysis essay on <em>The Kite Runner</em></td>
<td>Fun group research about Thailand</td>
<td>Teacher-led thesis instruction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Give peer feedback comments and discuss</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Revise essay after feedback</td>
</tr>
<tr>
<td>Phase Three: Organization, Oregon</td>
<td>2,000-word research paper on social or environmental topic of their choice</td>
<td>Obstacle course group competition</td>
<td>Teacher led instruction about paragraph organization</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Give peer feedback comments and discuss on video</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Revise essay after feedback</td>
</tr>
</tbody>
</table>

Gamification researchers posit that teachers must be selective in the game attributes they choose to incorporate to achieve successful outcomes in their classroom (Toda et. al., 2019). I felt that the game aspects of narrative context, challenge and performance would create an authentic game atmosphere without adding unnecessary distractions for the learners. Van Eck references a K-12 technology integration model, the NTeQ Model (Lowther & Morrison, 1998), to bridge the gap “between designing practical, effective lesson plans for the real-world classroom and the theories of learning and instruction as they exist within games” (p. 180). The elements of this model include that (a) the teacher is technologically competent and assumes the roles of designer, manager, and facilitator; (b) the student actively engages in the learning process, assumes the role of researcher, and becomes technologically competent; (c) the computer is used as a tool, as it is in the workplace, to enhance learning using real-world data.
to solve problems; (d) the lesson is student centered, problem based, and
authentic, and technology is an integral component; and (e) the environment
incorporates multiple resource-rich activities.

With these principles in mind, this 12-week innovation featured me as the
gameshow host guiding students in simulated travel around the world just like
the participants on the actual The Amazing Race show. I created small teams of
3-4 students. I used multi-media technology to create the narrative context of
being in a foreign city on The Amazing Race. On the first day of each writing
phase, students would enter the room as the theme song of the Amazing Race
was played. Then, they would watch a video of the teacher, acting as the
gameshow host, standing in an exotic, fantasy locale like Revision Russia, Thesis
Thailand, or Organization Oregon giving instructions about that phase of the
game. I recorded these videos by standing in front of my promethean board
while YouTube videos of foreign landscapes were playing and recorded myself
with a mobile phone. Figure 3.1 illustrates the use of these teacher-created
videos.

After watching the video, students would participate in a team-building
challenge to earn points. This aspect of the innovation was beneficial in two ways
because it provided the platform for students to build social relationships which
is a pre-cursor to effective peer feedback (Saidy & Early, 2016; Witte, 2013), and it
incorporated the game element of challenge. Team building challenges lasted
approximately 30 minutes of a class period. The first challenge was a scavenger
hunt in the library where teams were racing to be the first ones to find their game
clue. For Thesis, Thailand, students researched then presented things they would
do for fun while visiting Thailand. The last challenge, for the research paper, was
a blind-folded obstacle course where teams competed for time. Figure 3.2 is a picture of students competing in the obstacle course challenge.

Figure 3.1
Teacher-created video for Phase 2: Thesis, Thailand

Note. This is a screenshot of a teacher-created video using a promethean board and YouTube video of Thailand scenery as the background. The video was recorded with an Iphone 12.

Figure 3.2
Students competing in a team building challenge
After the team challenges in each phase, students then received writing instruction specific to the place they were at in the Amazing Race. For example, in Revision, Russia, which was their first essay, the technology essay, they completed an instructional module that introduced them to revision strategies. For the second paper, Thesis, Thailand, the writing instruction consisted of an activity called a thesis throwdown. This activity reinforced that a thesis should be the most revised sentence in the essay and that all paragraphs should support it. For the research paper, Organization, Oregon, instructions were given about paragraphing, transition, and conclusions.

Students were reminded how these writing concepts related to the rubric requirements for that essay and instructed about the number of comments to give, usually 2-3 depending on the number of group members they had. They shared their essays via Google Drive and used the comment feature in Google Docs which tracks the comments. Table 3.1 outlines the writing focus, essay, and activities that corresponded to each phase of the innovation.

The last element of gamification, performance, was included by awarding teams points that led to specific badges at the end of each writing phase. They were given points for winning the team-building challenges, for the number of substantive comments on each essay (which also depended on everyone in the group turning in an essay) and, finally, for grades the students achieved on their essays. The points were kept on a Google Sheet and teams were informed each time the scores were changed during the writing phase. After the papers were graded, however, the team with the most points was awarded the first-place badge for that leg of the race. Subsequent rankings of second through seventh place were awarded to the other teams based on their points and those badges...
were posted on the leaderboard in Canvas. Additional badging such as the teamwork badge was available to reward teams that overcame initial struggles such as one person being late turning in a paper.

At the conclusion of the first two phases, the first-place team was rewarded with gift cards, and remaining groups were entered into a drawing for gift cards. Student groups were changed after the second phase so that students could exchange feedback with multiple peers throughout the semester. Figure 3.3 below shows the badges that student groups earned.

**Data Collection Methods**

The multi-methods data sources for this study included written peer feedback comments, feedback discussion videos, and student essays. In aggregate, these data collection methods offered insight into the effectiveness of gamified peer feedback through triangulation, which is defined as collecting data from multiple sources across time, space, and person (Mertler, 2014). Qualitative data for this study were triangulated in two ways, with multiple data sources and multiple data points by using student essays and feedback discussion videos collected during each of the four phases of the study. Table 3.2 below specifies all data collection methods addressing each research question.

**Student Essays**

Prompts for the three student essays correlate to units of study centered on various texts in the course. Students’ essays were the first method of peer feedback data collection. In each phase, students wrote the required essay in Google Docs, shared them with their Amazing Race group members using the share feature on Google Docs, and then uploaded them to the school’s Learning Management System, Canvas for teacher review.
Figure 3.3
Leaderboard with GPF badges

For each leg of the race, your team is trying to finish all challenges and earn the highest grades on your submitted papers for that leg. If you finish all the challenges, you earn a badge for that location. First place through 8th place will be determined by the average grade of your submitted papers, so revise well! If your team encounters struggles, take heart—you can earn a badge for demonstrating teamwork, perseverance or communication—these badges will be awarded by Mrs. Broome when your team is displaying these attributes! The team with the most badge points will be crowned the overall winner at the end of the semester and will receive gift cards. All the other teams will be put into a random drawing—the more badge points you’ve earned, the more your team’s name goes into the drawing. Another winning group will be drawn to receive gift cards. Of course, grades will be earned along the way as well, so give this your best effort throughout the semester and become a better writer in the process!

These are the badges you're trying to earn:

Note. These badges were created using GIMP (a free, web-based image manipulation program).
Table 3.2
Research Questions and Data Collection Methods Alignment

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data Collection Methods</th>
</tr>
</thead>
</table>
| RQ1: What impact does GPF have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school? | • Student essays  
• Peer feedback comments on student essays  
• Feedback discussion videos |
| RQ2: What impact does the type of peer feedback given and received have on their writing improvement? | • Student essays  
• Peer feedback comments on student essays  
• Feedback discussion videos |
| RQ3: What are the effects of gamified peer feedback on low versus high performing students’ writing improvement? | • Student Essays  
• Peer feedback comments on student essays  
• Feedback discussion videos  
• Writing Rubrics |

Peer Feedback Comments

Students were required to provide at least three comments on their peer’s essay using the comment feature in Google Docs. Using this method, all comments were stored digitally, time stamped, and labeled with the participant giving the feedback which were needed later for coding. Again, this data was both quantitative and qualitative—the number of comments each student makes as well as the type was later analyzed.

Feedback Discussion Videos

Peer feedback data were also collected by recording group dialogue when
students met face-to-face to discuss the written feedback that they provided on the student essays. These feedback discussion videos were recorded by each student group using their cell phones. Two students in the group were asked to record to minimize recording errors. Students submitted these videos by email or Airdrop to the teacher before leaving class. Using this method provided more data than teacher observation would have allowed because multiple groups were meeting at the same time. It also allowed the teacher to attend to student needs in the classroom so that the research was never at the expense of teaching, a danger warned about by Hubbard and Power (2003).

**Data Analysis Methods**

The research questions were addressed using quantitative and qualitative data analysis methods. A priori coding using codes identified in previous studies was conducted on the peer feedback qualitative data which was triangulated from student essays and feedback discussion videos. Descriptive statistics were used to measure the effect that gamification had on the type of peer feedback. Inferential statistics were used to compare coded data with quantitative data obtained from student essays scored with analytic rubrics to determine the effect of peer feedback on writing improvement. Table 3.3 displays how the data aligned with the research questions and data analysis methods that were used throughout the study. (1): What impact does GPF have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school? (2) What impact does the type of peer feedback given and received have on their writing improvement? (3) What are the effects of gamified peer feedback on low versus high ability students’ writing improvement?
Table 3.3

Research Questions, Data Collection, and Data Analysis Alignment

<table>
<thead>
<tr>
<th>Research Questions</th>
<th>Data Collection Methods</th>
<th>Data Analysis Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RQ1: What impact does GPF have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school?</strong></td>
<td>• Student Essays</td>
<td>• A priori coding</td>
</tr>
<tr>
<td></td>
<td>• Peer feedback comments on student essays</td>
<td>• Descriptive statistics</td>
</tr>
<tr>
<td></td>
<td>• Feedback discussion videos</td>
<td></td>
</tr>
<tr>
<td><strong>RQ2: What impact does the type of peer feedback given and received have on their writing improvement?</strong></td>
<td>• Student Essays</td>
<td>• A priori coding</td>
</tr>
<tr>
<td></td>
<td>• Peer feedback comments on student essays</td>
<td>• Analytic rubric scoring</td>
</tr>
<tr>
<td></td>
<td>• Feedback discussion videos</td>
<td>• Independent samples t-test</td>
</tr>
<tr>
<td></td>
<td>• Student essays</td>
<td>• Correlation tests</td>
</tr>
<tr>
<td><strong>RQ3: What are the effects of GPF on low versus high ability students’ writing improvement?</strong></td>
<td>• Peer feedback comments on student essays</td>
<td>• A priori coding</td>
</tr>
<tr>
<td></td>
<td>• Feedback discussion videos</td>
<td>• Analytic rubric scoring</td>
</tr>
<tr>
<td></td>
<td>• Student essays</td>
<td>• Independent samples t-test</td>
</tr>
</tbody>
</table>

A Priori Coding

Data triangulation is defined as collecting data from multiple sources across time, space, and person (Campbell et al., 2020; Denzin, 1978; Mertler, 2014). As referenced in Table 3.4, qualitative data was triangulated in two ways, with multiple data sources and data points using student essays and feedback discussion videos in each study phase. Next, it was coded as either substantive or surface feedback following an a priori method of coding (Creswell, 2008; Guest et al., 2014). These a priori codes—substantive, surface and praise comments
emerged from the literature on peer feedback (Early & Saidy, 2014; McGarrell, & Verbeem, 2007; Zhang et. al., 2017). First, comments given on student essays were sorted through for relevance and coded. Comments pertaining to meaning or structure were labeled substantive. Comments pertaining to grammatical issues were labeled as surface. Praise comments were coded substantive if they reference a meaningful aspect of the writing such as effectiveness of an idea or clear organization within a paragraph. The audio recordings of the peer feedback groups were listened to, transcribed, and coded as substantive or surface as well.

**Writing Rubrics**

Writing rubrics were used to determine participants’ writing improvement on each of the four essays. Rubrics are an accurate method for assessing writing improvement rather than tracking revision changes. Rubrics assess the result of changes made to content, organization, and grammar rather than the sum of individual revisions (Berg, 1999). Although they were teacher created to be auto-populating, the criteria for the rubrics were adapted from national ACT (American College Testing) writing standards. The rubrics vary slightly based on the different essay types given in each of the three phases, but all assess similar requirements with their respective weightings: ideas and analysis (25%), development and support (25%), effective organization (30%), and grammatical structure (20%). The research paper rubric and *The Kite Runner* essay rubric had one additional requirement regarding formatting and revision respectively. As illustrated by Appendices B-D, these Google Sheet rubrics auto-populate the score into the right most column when the teacher grades the essay. All rubrics have a maximum score of 100 with a 5-point deduction if the essay is submitted late.
The first essay was scored two times, once before GPF is administered to determine the students’ baseline writing ability which will be used to answer RQ4, and a second time after GPF. The other three essays were scored after GPF in each phase was given.

Statistical Tests

To address RQ1, descriptive statistics were used to determine what type of peer feedback the students gave because of the innovation. Descriptive statistics were also used to compare the mean writing improvement over the course of the four essays.

RQ2 considered how peer feedback type given and received affected writing improvement. Median splits were used to divide students by who gave or received the most of each type of feedback. Independent samples T-Tests (Adams & Lawrence, 2015) compared each of these groups to determine who showed the most writing improvement.

Next, Pearson’s correlation tests (Adams & Lawrence, 2015) were carried out in each phase with the writing improvement as a dependent variable and the number of each type of comments (received and given) as the independent variable. This type of repeated measures design increases the power of a study because differences between groups is maximized while error variability is minimized (Adams & Lawrence, 2015).

To address RQ3, analytic rubric scores prior to peer feedback in the first phase were to divide the participants into two groups: those with low writing ability and those with high writing ability. Independent samples t-tests were used to determine if there is a statistically significant mean difference in writing improvement between these two groups. All statistical tests throughout the
study used $\alpha = .05$ to determine significance and were conducted using JASP statistical software.

**Procedures and Timeline**

The study began approximately three weeks into an 18-week semester English course in the Spring of 2022. It consisted of three writing assignments spread over eleven weeks of the course and correlated to three major units of study required in the course.

Phase one began with obtaining informed consent forms because all participants are between the ages of 14 and 16. This consent form is included in Appendix A. The next action in Phase One was for students to write the first draft of the first essay which was graded to obtain baseline writing ability about each student. Students then participated in the first GPF team building activity and completed the peer feedback activities. The last step in Phase One was for students to revise their first essay and submit it for final grading by the teacher. All data was collected and stored for analysis. Phases two and three emulated the pattern of phase one, with data collection regarding the first two research questions occurring in each phase. Data collection concluded at the end of Phase 3 and data analysis took place thereafter. See Table 3.4 below for a timeline of implementation.

**Rigor & Trustworthiness**

Several methods were used to pursue rigor and trustworthiness in this study that addressed its qualitative and quantitative aspects and the nature of action research. Those methods included the use of thick, rich descriptions, data triangulation, as well as the validity and reliability of the quantitative measures. Rigor was also increased by the repetition, prolonged engagement, and
Table 3.4
Phase Implementation and Timeline

<table>
<thead>
<tr>
<th>Phase One: Revision, Russia</th>
<th>Activity</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>• Consent forms issued and processed</td>
<td>Week 3-6 of the semester</td>
</tr>
<tr>
<td></td>
<td>• Technology essay first draft completed</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Team building activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Peer feedback activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Technology essay revised</td>
<td></td>
</tr>
<tr>
<td>Phase Two: Thesis, Thailand</td>
<td>• <em>The Kite Runner</em> essay draft written</td>
<td>Week 7-9 of the semester</td>
</tr>
<tr>
<td></td>
<td>• Team building activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Peer feedback activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• <em>The Kite Runner</em> essay revised</td>
<td></td>
</tr>
<tr>
<td>Phase Three: Organization, Oregon</td>
<td>• Research paper draft written</td>
<td>Week 11-15 of the semester</td>
</tr>
<tr>
<td></td>
<td>• Team building activity</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Peer feedback activities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Research paper revised</td>
<td></td>
</tr>
</tbody>
</table>

Persistent observation (Mertler, 2014) of data collection and analysis from phases one through three over the duration of the semester.

**Thick, Rich Descriptions**

The most important measure taken to promote generalizability of knowledge was to provide thick, rich descriptions of the study context and participants. While generalizability is often relegated to quantitative research, some researchers acknowledge that in-depth qualitative research is well-suited for revealing findings that are not unique to a particular participant or setting (Glaser, 2002) … [and that] the rich, highly detailed, and potentially insightful nature of qualitative findings make them especially suitable for extrapolation” (Polit & Beck, 2010, p. 1452). Providing these details allows a reader to determine “the extent to which the overall findings ring true” (Shenton, 2004, p. 69) and
whether study findings may relate to their own. Mertler argues that rigor and credibility is increased when the teacher has conducted previous studies or previous cycles within the same study (2014).

**Data Triangulation**

Triangulation is defined as a qualitative research strategy that uses the convergence of different sources to get richer, fuller data and test validity (Mertler, 2014). Peer feedback data was triangulated from two different sources in the study: the comments on student essays and the feedback discussion videos. Data about writing improvement was also triangulated with data obtained from scoring rubrics from four grading cycles rather than from one grading period.

**Plan for Sharing**

This study findings were shared with several groups who are connected to and may benefit from the findings. These stakeholders include the study participants and their parents, administrators, other ELA teachers at the research location, and our local school board. Student participants within the study were the first group to learn of the findings. Upon completion of the study, a 3- to 5-minute YouTube video, as well as a more detailed written report about study findings, implications, and researcher’s reflections, was shared via email with all above-mentioned groups. Participant names and identifying information was withheld from all published findings, which is important because these students were still attending our school. Moreover, study findings were shared with school administrators including our instructional coach at our school who work closely with teachers to help us improve instruction. Our school board regularly requests information about effective academic programs at our school, so they
were provided summarized information about the study findings. Finally, I shared the study findings with many other teachers in my field via conference presentations such as the National Council Teachers of English and Association for Education Communications & Technology.
CHAPTER 4
ANALYSIS AND FINDINGS

The purpose of this action research was to evaluate the impact of gamified peer feedback on the type of peer feedback that students gave and how that affected the writing improvement of ELA students at Southern Charter High School. The following research questions guided this study:

1. What impact does GPF have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school?
2. What impact does the type of peer feedback given and received have on their writing improvement?
3. What are the effects of gamified peer feedback on low versus high ability students’ writing improvement?

This chapter is divided into two sections: the qualitative data findings and quantitative data analysis. Each section outlines the findings that connect GPF innovation and resulting peer feedback to writing improvement data. All students in the classroom consented to the study, but one student’s data was removed when they did not complete two of the four writing assignments.

Qualitative Coding

Initial Coding

Data analysis began with the initial coding of peer feedback. This data was digitally copied and pasted from each student’s essay from Google Drive into Delve qualitative data analysis software. Each comment was first attribute
coded, which entails labeling each comment with descriptive information (Saldana, 2016) about the student who gave the feedback and which student received the feedback. Simultaneous coding whereby a single excerpt of data is coded with multiple codes (Saldana, 2016) was then applied with the category of writing it referenced and whether that constituted a substantive comment or a surface comment. The substantive and surface coding also represented provisional coding which utilizes a pre-determined list of researcher-generated codes (Miles and Huberman, 1984).

Clear delineation between substantive and surface was critical in the coding process. Previous studies were used to establish the categories and they defined substantive comments as high-level suggestions that affect meaning, i.e. comments about thesis, argument, evidence or organization (Faigley & Witte, 1981; Wu & Schunn, 2020; Zang et al., 2017). In contrast, surface comments were defined as relating to low-level issues such as form, control of language and conventions—essentially error-detection at the sentence-level (Early & Saidy, 2014; Faigley & Witte, 1981; McGarrel & Verbeem, 2007).

Simultaneous coding of each item with the categories such as adding content, thesis, or organization made the delineation between substantive and surface comments more discernible. Consider Cameron’s comment that "your real-life examples are amazing, but I think you should back them up a little more for the readers to better understand the consequence a bit better" This comment was simultaneously coded as adding content. It pertains to the effectiveness of the writing to convey meaning which is a substantive writing concern (Wu & Schunn, 2020). Another student’s comment was about organization: "maybe save more of the detail or the body paragraphs so you don't repeat things and have
your intro be 3-4 sentences." Organization is another substantive writing concern (Wu & Schunn, 2020).

By contrast, simultaneous coding of sentence level concerns like formatting and punctuation (Faigley & Witte, 1981; Wu & Schunn, 2020) issues lent reliability to the coding of surface comments. These included comments such as "you’re supposed to indent the first line of a paragraph" which was coded as formatting as well as surface. Another clearly surface comment was "there are 2 periods here."

Each comment had to be carefully considered. For example, this comment: "I wouldn't advise you to use the word "nowadays" in the transition. It seems a bit conversational." Although this comment seems to be a sentence-level concern, usually indicates a surface-level comment, it is actually a high-level concern. Their discussion of how the wording implies the incorrect tone for writing is a valuable learning objective for writers at this level. Because high school students use informal language so often in texting and social media, how to write with a more formal tone is an important feedback comment for them to give, necessitating a substantive code.

Adding citations to an essay was another comment that was atypical in terms of how it was coded. Comments related to formatting are usually considered surface, but in the context of instruction regarding plagiarism, it constituted a substantive consideration. In contrast, adding content is typically considered substantive. One student's comment about adding a title was coded surface because it is not a concern for instruction at the high school level. Figure 4.1 provides a screenshot of these simultaneous layers of coding in Delve.
Figure 4.1

Feedback Discussion Video Transcript and Corresponding Codes

[Transcript and Codes]

May 19th 3:12 4:13 Research paper

- so far, I liked how you wrote a scenario out for your first paragraph to draw people in and you said racism is often blown off as a type of discrimination. I think you could use a different word for often

- Research Paper, praise wording, sub

I said I liked how your thesis was straightforward and how you included the definition for racism instead of just assuming that the reader knows what you're talking about

- Research Paper, no improvement suggested, praise

make sure you're adding quotes from your slides and then citing them afterwards and then to go along with that, I highlighted one of them and it had the quotation marks but there was no citation after it and there were several of them so just make sure you cite them afterward

- plagiarism, sub

- paper

you used really good evidence and you have very good vocabulary, and in your second to last paragraph it's a little bit wordy but its very interesting and adding new information to keep the flow going

- add content, Research Paper, praise, reduce content, sub

I said I liked how the organization of your paper's paragraphs was going good so far and then I said one of the quotes that you used was a really strong piece of evidence and I think it relates to your paper really good

- Research Paper, no improvement suggested, praise

I said whenever you're quoting something like explain what you're getting at to get your point across and um I think its your third paragraph it's a little long so if you can, find a way to break it up

- add content, Research Paper, organization, sub

- paper

in your thesis statement explain how people are hurting them and how they help our ecosystem so that the reader knows, and you may want to add information to one of your paragraphs...

- add content, Research Paper, sub

I said I would say "significantly impacts" to have a stronger effect, and your last couple of paragraphs were really short and find a way to lengthen them

- add content, Research Paper, praise wording, sub

- I liked that you used ethos bio it really drew me in, and then the fourth paragraph thought you could possibly shorten it and for your sentence "aid of sea turtles", instead of using at in the next sentence you could use a different...
As Valerie Janesick (2011) observes, in addition to systematic analysis, “the qualitative researcher should expect to uncover some information through informed hunches, intuition, and serendipitous occurrences that, in turn, will lead to a richer and more powerful explanation of the setting, context, and participants in any given study” (p. 148, as cited in Saldana 2016, p. 64).

Researcher autonomy is emphasized in qualitative research where provisional codes are used, where one may too easily find what they are looking for (Dey, 1993). To counter this, qualitative researchers should revise, delete and expand codes to incorporate their experience and resulting hypotheses or hunches (Saldana, 2016).

Ultimately, I adhered to previous definitions for substantive and surface in much of the coding. As the coding progressed, however, it became clear that these categories are relative to the writing level of the student. What may be a low-level surface comment about a transition word for an undergraduate student would be considered high-level substantive for a tenth grader because it represents an area of writing that is significant for revision at their level. There were some sentence-level comments that, after consideration, were labeled substantive for this reason. For example, Sarah offered, “instead of saying this disease, I think you should say ‘coral bleaching’ because it partners with your intro.” Comments like these were coded as wording, many of which seemed like surface-level initially, but because they pertain to the cohesiveness of an essay (a significant aspect of writing), they qualify as substantive. Camille’s comment “um, I think its your third paragraph it’s like a little long so if you can, find a way to break it up” is another example of what seems like a reference to form
which is usually a surface concern, but is substantive because it deals with the organization of ideas within an essay. Careful attention was taken in both the initial and subsequent rounds of coding to place comments in the appropriate categories given the developmental level of high school writers.

Another important aspect of the analysis was to look beyond the lack of sophistication sometimes present in student dialogue to discover the quality of feedback they might be struggling to express. In this exchange below, Leigh Ann comments on two important aspects for writing improvement in Jed’s essay. She articulates the first one: the transition between ideas. The second suggestion is an equally important aspect of writing for students at this level to master, but she does not articulate cohesiveness, a substantive concern, specifically.

Leigh Ann: “I asked Gavin where the quote started and ended because he had a double quote and I think you should make it more clear and put a transition sentence at the end of your paragraph. Or you can do whatever the heck you want.

Jed: on the sentence it says “the American Dream”…it generalized what the American Dream was; you should make it more specific to say that’s Bernie Sanders’ American Dream.

Leigh Ann: I think you should elaborate on why you disagree with Bernie.
Both of Leigh Ann’s comments were coded as substantive.

A modification made to the provisional coding of substantive and surface was the addition of comments related to praise. Research validates that these comments can have significant effect on writing improvement (Patchan et al., 2016), so I wanted to track them, but they did not always coincide with a suggestion for any improvement. If they did not, they were also coded “no improvement suggested.” Out of the total 631 student comments recorded in the study, 169 of them were praise, and of those, 71 were praise only, with no suggestion for improvement. Examples from these two different types of praise comments are included below:

Elena’s comment on Brandon’s research paper May 5, 2022
I really like how you look into the history here! This would be a very good place to mention historical allusion (praise with suggestion)

Lexi’s comment, feedback discussion video May 4, 2022.
I agree I think that you used good questions about why now if how if they weren’t having to pay back then, what makes it any different now and the strong literary device of logos
and you had a good explanation of that

(praise only)

Dialogic comments were another category that arose from the coding. These comments revealed the conversational and socially interactive nature of successful peer feedback emphasized in many studies (Boud, 2015; Espasa et al., 2018; Zhu & Carless, 2018). Comments like these, although isolated (a total of five comments), point to a deep level of understanding that can emerge between students during this process. They represent the highest levels of engagement that can emerge in peer feedback interactions because they show the intellectual exchange that writing can evoke. An example of this is given below:

Kaya’s comment, I think this is kinda implying that only feedback discussion Christians have these beliefs, but in reality, video May 4, 2022. it’s also several other religions.

Subsequent coding

Subsequent phases of coding involved a careful review of each item to determine that every comment was assigned to both a giver and a receiver of the comment since this was going to be statistically analyzed. The most common change that was made in this round of coding was the clarification between substantive and surface codes, although less than 10% of the original codes were revised for this reason. Codes under “wording” garnered the most modification because there were cases when a single word affected meaning, making it substantive, or an entire phrase could be a less significant, surface comment. For example, the following comment references the “flow” of the essay and reducing content, which were often indicative of substantive change. Upon analysis, however, the resulting revision is relatively minor, and therefore best described
as a surface comment: “Try taking out “in the article,” to make it flow a little easier. Or add it to the end of the sentence.”—Ivan on Micheal’s research paper. And although many suggestions for wording revision were indicated at the sentence level, some indicate substantive revision that is valuable for a student writer to address, such as in: “maybe re-word this sentence. I found it a little hard to understand (mainly towards the end).”—Ivan on Rusty’s research paper, May 19th, 2022.

Painstaking analysis was applied to every code as evidenced by Leah’s comment on Gary’s research paper. At a glance, it was very simple, asking: “because?,” Upon reviewing the selected text tied to the comment in Google Drive, however, it was clear that her comment was in reference to his statement, “I disagree with Bernie that college should be totally free”. Her open-ended comment was a very effective way to encourage a substantive revision for clarification of what he meant. Other codes that received modification in subsequent phases included the “citing” category that had eleven codes in it and referred to writers citing their sources. Upon review, most of them addressed either praise, adding content, or plagiarism, and were listed exclusively in those categories rather than a separate one. Another category with eleven codes, tone, however, seemed significant to report as an independent code because it is important to writing instruction at this level. “Improve vocabulary” was another small, but significant category to report because it was another area addressed frequently in instruction. The low frequency of comments with this code indicates that more instruction is needed in the future because it is usually an area of weakness for high school writers. The category of “wording” was
renamed “phrase wording” so that it would not be confused with improved vocabulary which refers to the use of a single word.

In all, students produced 467 substantive peer feedback comments compared to 108 surface comments during the semester. The most frequent type of comments made substantive in nature including adding content, clarification, organization, phrase wording, transition, thesis, etc. Two hundred-thirty-four praise comments addressed substantive concerns as well. Examples of these type of comments include comments such as: “I like the examples and evidence you used to support your claim because I think it makes your essay stronger and more understandable.” —Jose on Matthew’s technology essay. Another substantive praise comment said, “I really like how you look into the history here! This would be a very good place to mention historical allusion.” —Perrin on Kerry’s research paper May 4, 2022. Clearly, surface comments accounted for the fewest number of comments with the highest frequency being comments about punctuation (n = 35) and formatting (n = 34). Table 4.1 below provides a chart of the final codes that list all substantive and surface comments that the students offered.

Table 4.1
Peer Feedback Codes by Category, Number and Type

<table>
<thead>
<tr>
<th>Peer Feedback Code</th>
<th>Code Frequency</th>
<th>Peer Feedback Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Add content</td>
<td>220</td>
<td>Substantive</td>
</tr>
<tr>
<td>Clarify</td>
<td>85</td>
<td>Substantive</td>
</tr>
<tr>
<td>Phrase wording</td>
<td>62</td>
<td>Substantive</td>
</tr>
<tr>
<td>Transition</td>
<td>59</td>
<td>Substantive</td>
</tr>
<tr>
<td>Thesis</td>
<td>58</td>
<td>Substantive</td>
</tr>
<tr>
<td>Reduce Content</td>
<td>49</td>
<td>Substantive</td>
</tr>
<tr>
<td>Cohesiveness</td>
<td>41</td>
<td>Substantive</td>
</tr>
<tr>
<td>Provide Examples</td>
<td>25</td>
<td>Substantive</td>
</tr>
<tr>
<td>Peer Feedback Code</td>
<td>Code Frequency</td>
<td>Peer Feedback Type</td>
</tr>
<tr>
<td>-----------------------</td>
<td>----------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>Introduction</td>
<td>19</td>
<td>Substantive</td>
</tr>
<tr>
<td>Flow</td>
<td>18</td>
<td>Substantive</td>
</tr>
<tr>
<td>Plagiarism</td>
<td>18</td>
<td>Substantive</td>
</tr>
<tr>
<td>Quotes</td>
<td>7</td>
<td>Substantive</td>
</tr>
<tr>
<td>Conclusion</td>
<td>7</td>
<td>Substantive</td>
</tr>
<tr>
<td>Tone</td>
<td>11</td>
<td>Substantive</td>
</tr>
<tr>
<td>Passive Voice</td>
<td>1</td>
<td>Substantive</td>
</tr>
<tr>
<td>Punctuation</td>
<td>28</td>
<td>Surface</td>
</tr>
<tr>
<td>Formatting</td>
<td>28</td>
<td>Surface</td>
</tr>
<tr>
<td>Citing</td>
<td>7</td>
<td>Surface</td>
</tr>
<tr>
<td>Comma</td>
<td>3</td>
<td>Surface</td>
</tr>
<tr>
<td>Sentence Structure</td>
<td>7</td>
<td>Surface</td>
</tr>
<tr>
<td>Run-on Sentence</td>
<td>6</td>
<td>Surface</td>
</tr>
<tr>
<td>Verb Tense</td>
<td>2</td>
<td>Surface</td>
</tr>
<tr>
<td>Spelling</td>
<td>1</td>
<td>Surface</td>
</tr>
</tbody>
</table>

Hand Counting

While Delve coding was sufficient to determine the number of comments each student gave, the most accurate way to determine the number and type of comments that each student received involved hand counting. Each of the 28 students had a stack of four post-it notes with the number and type of comments—substantive (sub), surface (sur) and NIP (No Improvement/Praise)—each student received listed by essay. Figure 4.2 shows the process of hand-counting the codes from each student’s essay from Delve.

Quantitative Data Analysis

Quantitative data in the study consisted of the number of each type of peer feedback comments obtained from the peer feedback coding as well as essay scores from the rubrics. Although the four rubrics differed slightly in language specific to the different writing tasks, they all addressed the same essential writing categories—introduction and thesis, development of ideas, organization,
and language use. In addition to being scoring instruments, however, rubrics are used by students to understand the specifics of a writing assignment. For that reason, it is helpful to them if the rubric explicitly describes items such as a thesis for an argument essay which is different than a thesis for a literary analysis. For example, the first rubric used for the technology essay cited that “the argument’s thesis reflects nuance and precision in thought and purpose. The argument establishes and employs an insightful context for analysis of the issue and its perspectives. The analysis examines implications, complexities, and tensions, and/or underlying values and assumptions.” The rubric for the second essay, *The Kite Runner* literary analysis differed somewhat explaining that the “introduction establishes the topic of the essay and gives appropriate background; the paragraph should "hook" the reader by making the essay seem interesting and sophisticated analysis of literary device is revealed.

The second category on each rubric referenced the development and support for ideas. On the first and last rubric, this was called development and On *The Kite Runner* essay rubric this was described as analysis of text while the research rubric detailed it as incorporating research. The third and fourth
categories pertained to organization and language use respectively. *The Kite Runner* essay had an additional category which scored revision between the first and second draft since this was a focus of the instruction for that phase. The research rubric added a category for formatting that addressed research citations and MLA requirements. These rubrics are included in Appendices B-E.

**Descriptive Statistics**

The first statistical tests conducted in JASP were descriptive in nature to examine the effect of the game innovation on the quality of peer feedback. Quality is determined by the number of comments that are substantive in nature versus the lesser quality surface and praise comments. One student’s data were removed from the sample when they did not complete two of the essays. The remaining students (n = 27) provided more substantive comments (M = 17.30, SD = 8.46) than surface comments (M = 4.00, SD = 3.35) while Praise/No improvement comments were the fewest (M = 2.67, SD = 3.81). These results are presented in Table 4.2 below.

<table>
<thead>
<tr>
<th>Type of Comment</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantive</td>
<td>17.30</td>
<td>8.46</td>
</tr>
<tr>
<td>Surface</td>
<td>4.00</td>
<td>3.35</td>
</tr>
<tr>
<td>Praise/No Improvement</td>
<td>2.67</td>
<td>3.81</td>
</tr>
</tbody>
</table>

The results presented in Table 4.2 indicate higher quality of peer feedback when students’ essays were revised with the gamified peer feedback system. The mean grade for the class on each essay is presented in Table 4.3 below. Each essay cycle evidenced consistently higher writing improvement with the highest gains reported (+9.63) between the pre-revised technology essay
(M = 68.04, SD = 11.38) and the revised technology essay (M = 77.18, SD = 9.99).

Students averaged 7.52 points of increase between the revised technology essay and *The Kite Runner* essay. The lowest increase was reported between *The Kite Runner* essay and the research paper with a 4.24 point gain. This information is presented by Table 4.3 below and illustrated by Figure 4.3.

Table 4.3

Grade Means for Each Essay Cycle

<table>
<thead>
<tr>
<th>Essay</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-revised Technology Essay</td>
<td>67.78</td>
<td>11.52</td>
</tr>
<tr>
<td>Revised Technology Essay</td>
<td>77.41</td>
<td>10.12</td>
</tr>
<tr>
<td><em>The Kite Runner</em> Essay</td>
<td>84.93</td>
<td>8.60</td>
</tr>
<tr>
<td>Research Paper</td>
<td>89.17</td>
<td>11.06</td>
</tr>
</tbody>
</table>

Figure 4.3

Mean Writing Improvement for Each Essay Cycle
The next calculation involved determining a writing improvement score for each student to represent the amount of writing growth that each student demonstrated. The writing improvement score was determined by averaging the last two essays (The Kite Runner essay and research paper) minus the average of the first two grades (pre-revised technology essay and revised technology essay). Four students did not submit the final research paper; their improvement scores were based on the growth from the average of the first two essays to The Kite Runner essay. All rubric grades as well and the writing improvement scores were recorded on a Google Sheet document where additional data calculations were made.

To answer RQ2, a median split of the data sample was again utilized for the number of given and received comments in each feedback types. These six categories were as follows: substantive given, substantive received, surface given, surface received, praise/no improvement given and praise/no improvement received. The two groups in each category were documented in the Google Sheet for later testing.

The Google Sheet was also used to determine writers of high ability versus low ability needed to answer RQ3. A median split in the scores ($Mdn=71$) on the pre-revised technology essay produced these two groups. A section of this Google Sheet is given in Figure 4.4 below.

**Inferential Statistics**

Independent samples t-tests and correlation tests were administered in JASP to examine the effect of comments given or received on writing improvement as well as the relationship between the frequency of each type of comment given or received and writing improvement (RQ2). Normality tests
Figure 4.4

Google Sheet Calculations of Quantitative Data

<table>
<thead>
<tr>
<th>Ability grouping pre-revised tech</th>
<th>revised tech</th>
<th>Kite Runner</th>
<th>Research paper</th>
<th>grade impro</th>
<th>4 p Sub given</th>
<th>Sur give</th>
<th>NIP given</th>
<th>Sub Received</th>
<th>Sur Received</th>
<th>NIP Received</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54</td>
<td>83</td>
<td>87</td>
<td>97</td>
<td>23.5</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>18</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>67</td>
<td>71</td>
<td>58</td>
<td>77</td>
<td>-1.5</td>
<td>16</td>
<td>5</td>
<td>0</td>
<td>19</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>54</td>
<td>75</td>
<td>78</td>
<td>87</td>
<td>18</td>
<td>31</td>
<td>0</td>
<td>1</td>
<td>22</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>54</td>
<td>63</td>
<td>78</td>
<td>94</td>
<td>27.5</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>67</td>
<td>79</td>
<td>80</td>
<td>87</td>
<td>10.5</td>
<td>24</td>
<td>2</td>
<td>4</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>54</td>
<td>67</td>
<td>90</td>
<td>29.5</td>
<td>33</td>
<td>4</td>
<td>2</td>
<td>17</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>63</td>
<td>80</td>
<td>74</td>
<td>20.5</td>
<td>24</td>
<td>9</td>
<td>1</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>1</td>
<td>67</td>
<td>79</td>
<td>82</td>
<td>97</td>
<td>16.5</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>63</td>
<td>83</td>
<td>87</td>
<td>81</td>
<td>11</td>
<td>3</td>
<td>15</td>
<td>0</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>1</td>
<td>58</td>
<td>71</td>
<td>80</td>
<td>94</td>
<td>22.5</td>
<td>25</td>
<td>4</td>
<td>5</td>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>1</td>
<td>58</td>
<td>92</td>
<td>80</td>
<td>94</td>
<td>12</td>
<td>8</td>
<td>8</td>
<td>1</td>
<td>18</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>46</td>
<td>63</td>
<td>82</td>
<td>51</td>
<td>12</td>
<td>14</td>
<td>4</td>
<td>0</td>
<td>15</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>63</td>
<td>67</td>
<td>75</td>
<td>10</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>67</td>
<td>63</td>
<td>92</td>
<td>87</td>
<td>24.5</td>
<td>17</td>
<td>0</td>
<td>16</td>
<td>22</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>83</td>
<td>92</td>
<td>90</td>
<td>97</td>
<td>6</td>
<td>10</td>
<td>5</td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>67</td>
<td>82</td>
<td>11</td>
<td>11</td>
<td>4</td>
<td>2</td>
<td>13</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>79</td>
<td>78</td>
<td>84</td>
<td>6</td>
<td>14</td>
<td>5</td>
<td>2</td>
<td>16</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>79</td>
<td>83</td>
<td>87</td>
<td>97</td>
<td>11</td>
<td>27</td>
<td>1</td>
<td>1</td>
<td>22</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>96</td>
<td>94</td>
<td>97</td>
<td>10</td>
<td>8</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>71</td>
<td>93</td>
<td>100</td>
<td>25.5</td>
<td>19</td>
<td>4</td>
<td>0</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>83</td>
<td>83</td>
<td>82</td>
<td>97</td>
<td>6.5</td>
<td>17</td>
<td>3</td>
<td>0</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>83</td>
<td>92</td>
<td>97</td>
<td>100</td>
<td>11</td>
<td>32</td>
<td>9</td>
<td>0</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>79</td>
<td>90</td>
<td>15</td>
<td>25</td>
<td>2</td>
<td>0</td>
<td>10</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>75</td>
<td>75</td>
<td>85</td>
<td>94</td>
<td>14.5</td>
<td>21</td>
<td>5</td>
<td>4</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>88</td>
<td>92</td>
<td>89</td>
<td>87</td>
<td>-2</td>
<td>23</td>
<td>0</td>
<td>1</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>83</td>
<td>100</td>
<td>94</td>
<td>20</td>
<td>15</td>
<td>2</td>
<td>11</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>83</td>
<td>79</td>
<td>97</td>
<td>84</td>
<td>9.5</td>
<td>18</td>
<td>5</td>
<td>9</td>
<td>19</td>
<td>0</td>
</tr>
</tbody>
</table>
were conducted on each of the data sets to determine which type of statistical testing was appropriate. The results of these normality tests are given first followed by the results of the correlation tests and independent samples t-tests.

**Normality Tests**

Shapiro-Wilks test were conducted on the data groups in RQ2 to assess distribution normality. The first normality test was conducted on the data grouped by total comments given and received to prepare for correlation testing. Both samples were found to be normally distributed and illustrated by Table 4.4 below.

Table 4.4

<table>
<thead>
<tr>
<th>Peer Feedback</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Given</td>
<td>.96</td>
<td>.40</td>
</tr>
<tr>
<td>Total Received</td>
<td>.98</td>
<td>.93</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution.

Normality tests were then applied to the data grouped by peer feedback type received and given. Table 4.5 outlines the results of these tests which found that only three of the six data sets to be normally distributed: surface given, praise/no improvement given, and praise/no improvement received (p > .05). The last Normality tests were applied to the data samples that had been produced by the median splits for the types of feedback received and given. To prepare for independent sample t-testing, the data were tested with Shapiro-Wilks and Levene’s test for equality of variances. The data sets for total comments given and received were found to be normally distributed as
Table 4.5

Shapiro-Wilk Normality Tests—Peer Feedback Type Given and Received

<table>
<thead>
<tr>
<th>Peer Feedback Category</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Substantive Received</td>
<td>.96</td>
<td>.37</td>
</tr>
<tr>
<td>Substantive Given</td>
<td>.95</td>
<td>.23</td>
</tr>
<tr>
<td>Surface Received</td>
<td>.98</td>
<td>.92</td>
</tr>
<tr>
<td>Surface Given</td>
<td>.88</td>
<td>.01</td>
</tr>
<tr>
<td>Praise/No Improvement Given</td>
<td>.73</td>
<td>&lt;.00</td>
</tr>
<tr>
<td>Praise/No Improvement Received</td>
<td>.82</td>
<td>&lt;.00</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution evidenced Shapiro-Wilk’s test (p > .05); and there was homogeneity of variances, as assessed by Levene’s test for equality of variances (p > .05). The assumptions for each test are provided in the tables 4.6 and 4.7 below.

Table 4.6

Shapiro-Wilk Normality Tests—Comments Given vs. Received Correlation

<table>
<thead>
<tr>
<th>Peer Feedback Category</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fewest Comments Received</td>
<td>.97</td>
<td>.87</td>
</tr>
<tr>
<td>Most Comments Given</td>
<td>.97</td>
<td>.86</td>
</tr>
<tr>
<td>Most Comments Received</td>
<td>.96</td>
<td>.62</td>
</tr>
<tr>
<td>Fewest Comments Given</td>
<td>.92</td>
<td>.02</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution.

Table 4.7

Levene’s Equality of Variances—Comments Given vs. Received Correlation

<table>
<thead>
<tr>
<th>Peer Feedback Category</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Comments Received</td>
<td>.549</td>
<td>1</td>
<td>25</td>
<td>.466</td>
</tr>
<tr>
<td>Total Comments Given</td>
<td>.006</td>
<td>1</td>
<td>25</td>
<td>.940</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution.

The next normality tests were applied to the median split groups by peer feedback type given and received. The data sets for total comments given and received were found to be normally distributed as evidenced Shapiro-Wilk’s test.
(p > .05); and there was homogeneity of variances, as assessed by Levene’s test for equality of variances (p > .05). The assumptions for each test are provided in the tables 4.8 and 4.9 below.

Table 4.8

Shapiro-Wilk Normality Tests—Peer Feedback Type and Grade Improvement

<table>
<thead>
<tr>
<th>Peer Feedback Type</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Substantive Given</td>
<td>.98</td>
<td>.939</td>
</tr>
<tr>
<td>Least Substantive Received</td>
<td>.97</td>
<td>.868</td>
</tr>
<tr>
<td>Most Praise Received</td>
<td>.96</td>
<td>.75</td>
</tr>
<tr>
<td>Least Praise Received</td>
<td>.96</td>
<td>.61</td>
</tr>
<tr>
<td>Most Surface Given</td>
<td>.96</td>
<td>.57</td>
</tr>
<tr>
<td>Least Surface Given</td>
<td>.95</td>
<td>.69</td>
</tr>
<tr>
<td>Least Substantive Given</td>
<td>.95</td>
<td>.64</td>
</tr>
<tr>
<td>Least Praise Given</td>
<td>.95</td>
<td>.55</td>
</tr>
<tr>
<td>Most Praise Given</td>
<td>.94</td>
<td>.53</td>
</tr>
<tr>
<td>Most Substantive Received</td>
<td>.94</td>
<td>.41</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution.

Table 4.9

Levene’s Equality of Variances—Peer Feedback Type and Grade Improvement

<table>
<thead>
<tr>
<th>Peer Feedback Type</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Received</td>
<td>.33</td>
<td>1</td>
<td>25</td>
<td>.57</td>
</tr>
<tr>
<td>Substantive Received</td>
<td>.43</td>
<td>1</td>
<td>25</td>
<td>.52</td>
</tr>
<tr>
<td>Substantive Given</td>
<td>.36</td>
<td>1</td>
<td>25</td>
<td>.56</td>
</tr>
<tr>
<td>Surface Given</td>
<td>.11</td>
<td>1</td>
<td>25</td>
<td>.750</td>
</tr>
<tr>
<td>Praise/No Improvement Given</td>
<td>.01</td>
<td>1</td>
<td>25</td>
<td>.93</td>
</tr>
<tr>
<td>Praise/No Improvement Received</td>
<td>1.53</td>
<td>1</td>
<td>25</td>
<td>.23</td>
</tr>
</tbody>
</table>

Note. Significant results (p < .05) indicate a non-normal distribution.

The last normality tests were conducted for the grade improvement scores on the students of high versus low ability. These data were normally distributed, as evidenced by Shapiro-Wilk’s test (p > .05), and there was homogeneity of variances, as assessed by Levene’s test for equality of variances (p = .061). See tables 4.10 and 4.11 below.
Table 4.10
Shapiro-Wilk Normality Tests—Grade Improvement by Ability Group

<table>
<thead>
<tr>
<th>Ability group</th>
<th>W</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Ability</td>
<td>.95</td>
<td>.65</td>
</tr>
<tr>
<td>Low Ability</td>
<td>.95</td>
<td>.61</td>
</tr>
</tbody>
</table>

Table 4.11
Levene’s Equality of Variances—Grade Improvement by Ability Group

<table>
<thead>
<tr>
<th>Grouping Variable</th>
<th>F</th>
<th>df1</th>
<th>df2</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability Level</td>
<td>1.35</td>
<td>1</td>
<td>25</td>
<td>.26</td>
</tr>
</tbody>
</table>

Correlation tests

The first correlation tests were used to evaluate the relationship between total comments given versus the total comments received and writing improvement. Given the normality of this data, Pearson correlation tests were conducted. Grade improvement and total comments received demonstrated small correlation ($r = .26, p = .20$). Grade improvement and total comments given resulted in slightly lower correlation ($r = .23, p = .24$). These results are displayed in Table 4.12 below.

Correlation tests were then administered to assess the linear association between the types of comments received or given, and students’ writing improvement. For the normality-distributed data sets, Pearson tests were used to test correlation. For the three non-normal data sets, non-parametric Spearman tests were conducted. Results revealed a small correlation between writing improvement and surface comments received ($r = .21, p = .30$). Writing improvement and substantive comments received showed a similar correlation...
Table 4.12

Correlation of Total Comments Given vs. Received and Writing Improvement

<table>
<thead>
<tr>
<th>Variable</th>
<th>$r$</th>
<th>Test</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Received</td>
<td>.26</td>
<td>Pearson</td>
<td>.20</td>
</tr>
<tr>
<td>Total Given</td>
<td>.23</td>
<td>Pearson</td>
<td>.24</td>
</tr>
</tbody>
</table>

($r = .20, p = .32$). Spearman test showed a small positive correlation between grade improvement and praise/no improvement given comments ($r = .15, p = .46$) while grade improvement and substantive comments given produced a similar correlation using the Pearson test ($r = .14, p = .49$). The smallest positive correlation was demonstrated between grade improvement and praise/no improvement received comments ($r = .02, p = .94$). Grade improvement and surface comments given was the only category to show a negative correlation ($r = -.09, p = .66$). These values are given in table 4.13 below.

Table 4.13

Correlation Between Frequency of Comment Types and Writing Improvement

<table>
<thead>
<tr>
<th>Type of Comment</th>
<th>Test</th>
<th>$r$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface Received</td>
<td>Pearson</td>
<td>.21</td>
<td>.30</td>
</tr>
<tr>
<td>Substantive Received</td>
<td>Pearson</td>
<td>.20</td>
<td>.32</td>
</tr>
<tr>
<td>Praise/No Improvement Given</td>
<td>Spearman</td>
<td>.15</td>
<td>.46</td>
</tr>
<tr>
<td>Substantive Given</td>
<td>Pearson</td>
<td>.14</td>
<td>.49</td>
</tr>
<tr>
<td>Praise/No Improvement Received</td>
<td>Spearman</td>
<td>.02</td>
<td>.94</td>
</tr>
<tr>
<td>Surface Given</td>
<td>Spearman</td>
<td>-.09</td>
<td>.66</td>
</tr>
</tbody>
</table>

**Independent Samples T-Tests**

Independent samples t-tests were also used to examine the relationships between grade improvement types of feedback given and received. For each test, students
were divided by median split to obtain the highest and lowest givers and receivers of each type of feedback. Given that all data was normally distributed, the first test was conducted on the samples separated by total comments given versus total comments received. As shown in Table 4.15 below, the group who gave the most comments demonstrated highest grade improvement ($M = 15.73, SD = 8.09$). The group who received the most comments demonstrated next highest grade improvement ($M = 15.37, SD = 8.37$). The group who gave the fewest comments reported a mean grade improvement of ($M = 12.58, SD = 8.16$) while the group who received the fewest comments had the lowest grade improvement ($M = 12.50, SD = 7.87$). These results are given by Table 4.14 below.

Table 4.14
Grade Improvement by Highest Givers or Receivers

<table>
<thead>
<tr>
<th>Comment Type</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Comments Given</td>
<td>15.73</td>
<td>8.09</td>
</tr>
<tr>
<td>Most Comments Received</td>
<td>15.37</td>
<td>8.37</td>
</tr>
<tr>
<td>Fewest Comments Given</td>
<td>12.58</td>
<td>8.16</td>
</tr>
<tr>
<td>Fewest Comments Received</td>
<td>12.50</td>
<td>7.87</td>
</tr>
</tbody>
</table>

The next tests compared the writing improvement between groups in terms of the type of peer feedback the given or received. The group that gave the most substantive feedback presented higher grade improvement ($M = 15.46, SD = 8.54$) versus the group that gave the least substantive feedback ($M = 12.62, SD = 7.73$). The group that received the most substantive feedback demonstrated the second highest grade improvement of all the groups ($M = 16.13, SD = 8.19$) while the group that gave the least substantive feedback presented one of the lowest mean improvements ($M = 12.62, SD = 7.73$).
The sample was then tested based on surface feedback given and received. The group that gave the most surface feedback presented higher grade improvement ($M = 14.60$, $SD = 8.54$) while those that gave the least surface feedback scored lower in writing improvement ($M = 13.36$, $SD = 7.84$). The group that received the most surface presented show mean improvement of ($M = 14.60$, $SD = 8.54$) versus the group that received least surface ($M = 13.36$, $SD = 8.18$).

The next tests compared students on the variable of praise/no improvement given or received. The group that gave the most praise presented the highest mean average of all the groups ($M = 16.50$, $SD = 7.40$) with the group that gave the least praise presenting one of the lowest mean scores ($M = 12.17$, $SD = 8.41$). The group that received the most praise earned a mean score improvement of ($M = 12.96$, $SD = 7.21$). The group that received the least praise earned a mean score improvement of ($M = 15.00$, $SD = 8.94$). These results are presented in Table 4.15 and Figure 4.5.

Table 4.15

Grade Improvement by Feedback Type Given and Received

<table>
<thead>
<tr>
<th>Comment Type</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most Praise Given</td>
<td>16.50</td>
<td>7.40</td>
</tr>
<tr>
<td>Most Substantive Received</td>
<td>16.13</td>
<td>8.19</td>
</tr>
<tr>
<td>Most Substantive Given</td>
<td>15.46</td>
<td>8.54</td>
</tr>
<tr>
<td>Least Praise Received</td>
<td>15.00</td>
<td>8.94</td>
</tr>
<tr>
<td>Most Surface Given</td>
<td>14.60</td>
<td>8.54</td>
</tr>
<tr>
<td>Most Surface Received</td>
<td>14.59</td>
<td>8.32</td>
</tr>
<tr>
<td>Least Surface Received</td>
<td>13.36</td>
<td>8.18</td>
</tr>
<tr>
<td>Least Surface Given</td>
<td>13.36</td>
<td>7.84</td>
</tr>
<tr>
<td>Most Praise Received</td>
<td>12.96</td>
<td>7.21</td>
</tr>
<tr>
<td>Least Substantive Given</td>
<td>12.62</td>
<td>7.73</td>
</tr>
<tr>
<td>Least Praise Given</td>
<td>12.17</td>
<td>8.41</td>
</tr>
<tr>
<td>Least Substantive Received</td>
<td>11.54</td>
<td>7.62</td>
</tr>
</tbody>
</table>
The last statistical tests were used to determine the differences outlined in RQ3 regarding the effects of the innovation on writing improvement in students of low ability versus those of high ability. These groups were identified using a median split of grades on the first essay, the pre-revised technology essay. One student’s data were deleted from the sample because they did not complete the first and final essays which resulted in 13 students per group. An independent-samples t-test was conducted to determine if there were differences in grade improvement between the low ability group and the high ability group. The results showed that students of low ability presented higher mean grade improvement ($M = 16.89, SD = 8.44$) compared to students of high ability ($M = 11.07, SD = 6.84$). These results are presented in Table. 4.16 and Figure 4.6.

Chapter Summary

This chapter delivered the data analysis methods and presented the qualitative and quantitative findings from the peer feedback and writing
Table 4.16

Grade Improvement by Ability

<table>
<thead>
<tr>
<th>Ability Level</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Ability</td>
<td>16.89</td>
<td>8.44</td>
</tr>
<tr>
<td>High Ability</td>
<td>11.07</td>
<td>6.84</td>
</tr>
</tbody>
</table>

Figure 4.6

Mean Grade Improvement Comparison Between High and Low Ability Groups

Improvement data collected during this study. Qualitative peer feedback data from student essays and feedback discussion videos were coded and quantified. Descriptive statistics were then used to address RQ1. Quantitative data from peer feedback coding and writing rubrics were then analyzed using inferential tests to address RQ2 and RQ3. Total comments given and received as well as peer feedback type given and received were tested for correlation to grade improvement. Independent samples t-tests were also used to compare total comments given and received, peer feedback type given and received and finally writing ability to grade improvement.
CHAPTER 5
DISCUSSION, IMPLICATIONS, AND LIMITATIONS

This chapter details the findings of how gamified peer feedback affected the writing improvement of high school English students. These findings are situated by research in the field regarding peer feedback. This study purposed to investigate how a gamified peer feedback model in a digital writing environment affects the quality of peer feedback and subsequent writing improvement for tenth grade ELA students at Southern Charter High School. The study focused on the following research questions: (1) What impact does gamified peer feedback have on the quality of peer feedback given by tenth grade ELA students at Southern Charter High school? (2) What impact does the type of peer feedback given and received have on their writing improvement? (3) What are the effects of gamified peer feedback on low versus high ability students’ writing improvement? Several key findings emerged: how peer feedback benefited students overall, how different types of feedback offered different value, and how the benefits differed for students based on their writing ability. These findings are examined in the (a) discussion, (b) implications, (c) limitations and (d) conclusions below.

Discussion

A primary goal of the study was to determine the effect of gamification on whether it led students to produce substantive peer feedback rather than surface feedback. Qualitative and quantitative data from written comments, group
discussions, and rubric scores were triangulated, and then meticulously coded and quantified in their respective categories. Next, writing scores were obtained from the rubrics and used to determine how student writing may have improved. Various descriptive and inferential statistical tests were then performed on these groups differentiated by the type of peer feedback they gave or received, and how it affected their writing improvement. The findings of this study suggest that gamification positively influences the amount of substantive feedback versus surface feedback that students give, but that both surface feedback as well as praise feedback may play an unexpected role in the writing improvement process.

Research Question 1: What impact does gamified peer feedback have on the quality of peer feedback given by tenth grade students at Southern Charter High School?

Peer feedback has been proven to help students improve their writing, but it must be carefully structured to be effective (McConlogue, 2015; Nguyen, & Khau, 2022 Simmons, 2019). Framing that instruction into game challenges helped me to guide certain aspects of the peer feedback without making it mundane for students. Kayacan and Razi (2017) point to the involvement of learners internalizing their own growth; in terms of writing, understanding writing as process-based rather than product-based is essential to writing improvement. Making the game an ongoing, semester-long interaction encouraged students to engage with the writing process as researchers advocate (Fitzgerald, 1987; Santangelo et al., 2007). All students, whether they were skilled or struggling, were engaged in giving feedback, which research shows can also be a powerful tool for improving one’s writing skills (Strijbos et al., 2010).
Gamifying the process of peer feedback also allowed several key features of effective peer feedback to occur: situating learning in a context of competition and purpose, encouraging collaboration, and finally, making it fun. Students seemed very engaged in the process, whether it was nearly running down the hall to find a clue, asking if they could give extra comments, or talking excitedly when they met with their groups. The group discussions allowed for the dialogic benefits of peer feedback to occur, something that I did not expect, but that several studies referenced (Espasa et al., 2016; Yim et al., 2014). The video evidence of these discussions also provided triangulation of data which helped delineate between some substantive and surface comments that were sometimes difficult to discern.

Using student teams to give peer feedback is grounded in Vygotsky’s socio-cultural learning theory that learning is not an individual, secluded activity, but one mediated by social interaction (Vygotsky & Cole, 1978). Collaborative learning involving use of small peer groups also finds support in composition theories which emphasize the importance of peer interaction in writing development. Students may even benefit more from peer feedback than teacher feedback (Karegianes, Pascarella, Pflaum, 1980) because of an enhanced sense of audience, raised awareness of own problems, and fostered ownership of their own texts. Having students accountable to each other for game performance also seemed to solve one traditional problem that occurred in my classroom, which was students not turning in their essays on time. I was pleasantly surprised at how few times a student did not have their essay, and how the problem decreased over time.
Overall, gamification benefitted the quality of feedback that students gave because research has also shown that novice students often defer to surface writing concerns if they are not taught to do otherwise (Faigley & Witte, 1981, Hanjani, 2019). In all, students produced 467 substantive peer feedback comments compared to 108 surface comments which is an average of 17.30 substantive comments to 4 surface comments per student.

A finding of this study, however, is that although previous research has used the substantive and surface delineations, the boundaries of those prescribed labels must consider the developmental level of the writer. In the body of research, substantive comments refer to high-level writing suggestions that affect meaning and that surface comments relate to low-level writing concerns like grammar errors (Early & Saidy, 2014; Faigley & Witte, 1981). The implication here is that surface comments are quick fixes which do not encourage the writer to question their writing. Substantive comments such as asking the writer to add more detail or re-organize the flow of ideas engage the writer in a more thoughtful evaluation of their writing overall, are what leads a writer to improve their writing. Because the onus for improvement is on the writer, however, their ability level must be considered. For example, a comment about the wording of a sentence may be a surface concern for a graduate student but be a substantive exercise for a writer in high school. If substantive comments encourage substantive revisions and surface comments result in surface revisions, the goal of feedback would be to encourage the most sophisticated evaluation a writer can make, which would vary depending on the ability level of the writer. The areas for desired writing growth in each context are what define substantive
feedback while those that are less desirable because they do not encourage meaningful improvement are surface feedback.

The most significant finding in the coding process, however, pertained to the quality and range of substantive comments that students gave. The most frequent comment was about adding content (n=220), often regarded as one of the most quality suggested revisions to offer. The next most frequent comments regarded clarification, organization, transition, and thesis. These are significant areas of writing to address in a high school course as well, reflected in the English Language Arts standards for our state, Common Core writing standards and National ACT writing Standards. Having students apply these grade-level writing concepts to their own writing is a powerful instructional experience. This phenomenon was bolstered by the finding that students’ comments originated from revision instruction provided in the course such as adding content, organization, and transition. In approximately half of the comments, students used the exact language referenced by the teacher like better transition, or flow of ideas, with the other half using informal wording to relate substantive concerns.

There was only one code out of the 457 that did not harken back to writing instruction offered in one of the phases. That code, which referred to passive voice, came from another. Students were not expected to seek outside help with their writing, but this student sought extra help from another teacher that they trusted. That teacher was not an ELA teacher, but rather a foreign language teacher, whose comments about passive voice align more with instruction students typically receive later such as in twelfth grade. With comments being so closely tied to areas of instruction in the course, this one code demonstrates how areas of writing can remain unfamiliar to students unless explicitly taught in a
course. It also demonstrates how adult writers may be far removed from where a student is in terms of ability, not working in student’s zone of proximal development when they are commenting on areas like passive voice, an area of writing that no students commented on during the study.

Some feedback that students gave did not fit into the expected categories of substantive or surface, however. These 72 comments, an average of 2.67 comments per student, were identified in the analysis section as praise/no improvement because they did not offer any suggestions for improvement. These comments often did, however, validate important writing aims like cohesiveness or meaning. Considering previous studies about the value of praise (Patchan et al, 2016), however, they still speak to the quality of peer feedback that students produced. This was a surprising aspect of the study because these comments were not encouraged by the instruction or incentivized in any way. Low ability students often offered these as their only comments, especially in the beginning phases, indicating they could not find any areas for improvement in essays written by a high ability student, although praise comments were also given by high-ability students, often in addition to the required number of comments that were required. These scenarios reveal that praise may somehow be an intuitive part of the feedback process where students may be trying to establish rapport. Research has shown that this type of congenial relationship is vital for peer feedback to be effective (Boud, 2015; Saidy & Early, 2016).

Dialogic comments were another surprising category that arose from the coding, that considered neither substantive nor surface in nature. An example of a dialogic comment is one like, “I agree whenever I am hanging out with friends its a much better time if you interact with each other instead of looking at
phones.” —Ellie’s comment on Kaya’s technology essay. These comments show the transfer of meaning between writer and reader that is social in nature (Boud 2015, Espasa et. al., 2018). Dialogic feedback, like praise comments, were neither instructed nor incentivized, so their emergence indicates that they, too, are perhaps innate to peer feedback interaction. Writers getting feedback about how their words impacted the reader intellectually is clearly valuable, given the understanding of what writing aims to do which is to communicate meaning. The scarcity of these types of comments in the study, however, may indicate that too much restriction was placed on students making “substantive” comments. Questions such as “where do you agree or disagree with the writer?” may provide more dimension to future peer feedback sessions rather than focusing solely on concepts like form and organization.

**Research Question 2: What impact does the type of peer feedback given and received have on their writing improvement?**

Mean writing scores increased consistently, an average of 7 points between each of the three phases of the study. This writing improvement can be reasonably attributed to the only writing instruction that occurred in the course—gamified peer feedback and the largely substantive peer feedback it produced. The most significant gains were reported between the pre-revised technology essay ($M = 67.78, SD = 11.52$) and the revised technology essay ($M = 77.41, SD = 10.12$). This speaks to the effectiveness of gamification because although the act of peer feedback was new to them, they engaged with it effectively. Research indicates that peer feedback is often more impactful in the beginning versus later in the process when students may become fatigued (Mulder et. al., 2014; Planas et. al, 2014). Indeed, this law of diminishing returns
was evident in this study when gains in improvement occurred between the first two writing assignments and decreased each time over the subsequent essays.

It is significant to note here, then, that the essays, although similar in some respects such as requirements for organization and language conventions, were scaffolded to increase in difficulty over the course of the semester, and so the consistency of benefits across the study phases is significant. For example, the first two essays, the pre-revised technology essay and the revised technology essay, prompted students to write from their own experiences with technology and supplement it with some texts studied in class, with one of the most important texts being a Netflix documentary they were able to watch. The third essay, a literary analysis of The Kite Runner, was more difficult first, in terms of being a literary analysis, and secondly having a longer length requirement (600 words). The last essay, the research paper, was the longest with a length requirement of 1,000 words. Students had to write that essay based on their own research and then meet MLA formatting requirements as well. In the context of increasing difficulty, writing score gains across these the later phases seem even more noteworthy.

Many studies have found that giving peer feedback results in writing improvement for reasons such as fostering a growth mindset and increased familiarity with the assessment standards (Boud & Malloy, 2013; Feltham & Sharen, 2015; Moore & Teacher, 2013). Others have looked at the value of giving peer feedback versus receiving it (Chanski & Ellis, 2017; Huisman et al., 2018; van Popta et al., 2017), but only a few have looked at the effect that the type of feedback had on writing improvement of American high school students. Correlation tests in this study explored this question. Firstly, the positive
correlations in five of the six peer feedback types suggest that students who engage in peer feedback, regardless of the type, show better writing improvement (Berg, 1991; O’Donovan et al., 2004). Furthermore, they revealed that both comments given ($r = .26, p = .20$) and comments received ($r = .23, p = .24$) showed a small positive correlation to grade improvement. Similar results were found when correlations were broken down over peer feedback types and whether they were given or received. These tests showed the highest correlations to be between surface comments received ($r = .21, p = .302$) and substantive comments received ($r = .20, p = .322$) supporting the results found in the total comment correlation tests: receiving peer feedback leads to writing improvement. The results also show, however, that receiving praise had the smallest correlation of all the categories ($r = .02, p = .94$), results that are in keeping with research saying praise may diminish revision quality (Cheng et al., 2015; Patchun et al., 2016).

Surface given comments showed the only negative correlation to grade improvement. The key concept here is that these are comments that are given, which speak to the effort and or ability of the student giving them. This finding substantiates research in the field that students who defer to making these types of low-quality comments do not improve their writing (Saidy & Early, 2016) because surface comments reflect less effort and do not indicate interaction with the assessment standards. The results were not statistically significant, but this could also be due to the small sample size.

To explore this question further, independent samples t-tests were also performed between the highest givers and highest receivers when all peer feedback types were combined. Students who gave the most comments had the
highest mean improvement scores ($M = 15.73$, $SD = 8.09$). Students who received the most comments performed only slightly lower ($M = 15.37$, $SD = 8.54$), but both groups outperformed students who gave the fewest comments ($M = 12.58$, $SD = 8.16$) and those who received the fewest comments ($M = 12.50$, $SD = 7.87$). These results further the research in the field in a few ways. First, it corroborates previous findings that peer feedback has clear benefits for students who provide it (Chanski & Ellis, 2017; Huisman et al., 2018; van Popta et al., 2017). The assertion that students may become more familiar with the assessment criteria is supported by the givers in the study being defined by the number of instances where they gave feedback, the number of times they interacted with the criteria for good writing, regardless of the quality or type of the feedback they gave. It is noteworthy that there were clear similarities in mean writing improvement with students who gave the most comments ($M = 15.73$, $SD = 8.09$) or received the most comments ($M = 15.37$, $SD = 8.37$) versus the students who gave the fewest comments ($M = 12.58$, $SD = 8.16$) or received the fewest comments ($M = 12.50$, $SD = 7.87$). This strengthens findings in the study about the value of engaging in peer feedback, especially for those giving it. It is also noteworthy and not surprising given the research that students who gave the fewest comments, students are interacting the least with assessment standards and therefore showed the least amount of writing improvement. Whether this lack of giving comments is tied to writing ability, however, needs further research.

The impact of ability is also unclear as it relates to the benefit that giving feedback has. There seems to be a slightly higher benefit to giving rather than receiving, although it is unknown whether these are the higher-performing students who often give more effort to their work.
The next round of independent samples t-testing occurred with the mean writing improvement in students differentiated by the type of feedback they gave or received. Median splits were used to separate students based on who were the highest givers and receivers of each of the three types of feedback—substantive, surface and praise and then tested for mean grade improvement. This analysis yielded results that varied slightly from the correlations tests. Surprisingly, the highest mean grade improvement was presented by the group who gave the most praise ($M = 16.50$, $SD = 7.40$). Students who received the most praise ($M = 12.96$, $SD = 7.21$) did not improve as much as those giving praise, but they did outperform students who gave the least praise ($M = 12.17$, $SD = 8.41$). This is a significant finding because it is the type of feedback that improved writing the most across the two groups, even though other studies have found that praise did not always improve outcomes.

The independent samples t-test also revealed that receiving substantive feedback, however, is almost as impactful for improving writing. Students who received the most substantive comments scored almost as high as the group that gave the most praise ($M = 16.13$, $SD = 8.19$). This finding corroborates earlier findings in the correlation tests about how substantive feedback leads to improved writing. It also validates the innovation of gamified peer feedback, the aim of which is to increase substantive feedback.

What independent samples t-tests revealed about the role of surface feedback is noteworthy, however, because many researchers have minimized its ability to affect writing improvement (Bridwell, 1980; Faigley & Witte, 1981, Saidy & Early, 2016). Research in the field characterizes surface changes as minimal and therefore unrelated to significant writing improvement, but this
study finds that students who gave the most surface feedback ($M = 14.60$, $SD = 8.54$) performed just slightly lower than the two substantive groups, and very close to the students who received the most surface feedback ($M = 14.59$, $SD = 8.32$) indicating that although like previous studies suggest, surface feedback is not as valuable as substantive feedback, it does not diminish writing improvement in any way, and still leads to some writing improvement. This indicates that, for the high school classroom, students should be encouraged to give substantive revision, but not discouraged when they give surface feedback.

Lastly, the lowest mean writing improvement ($M = 11.54$, $SD = 7.62$) came from the group that received the least substantive feedback. In other words, students who did not receive adequate feedback are not improving their writing.

It should be considered that higher-scoring essays are more difficult to improve, but the median writing scores for the class average 82 which indicates these more advanced writers in the group still had some room to improve. Combined with the powerful influence that receiving substantive feedback ($M=16.13$, $SD = 8.19$) had on improving other students’ writing, this emphasizes the role that substantive peer feedback can play in writing improvement.

**Research Question 3: What are the effects of gamified peer feedback on the writing improvement of low versus high ability students?**

The last research question is one of the most important questions to answer in the study for several reasons. An advantage of action research is the ability to benefit marginalized groups by improving environments in which we work (Kemmis, 2006; Stringer, 2008). ELA teachers are often perplexed about how to engage low ability students and how to improve their learning outcomes. The question of whether peer feedback can help high ability students improve
their writing which is already of a high quality is also important, however (Gielen et al., 2010).

In a study that examined ability in relationship to peer feedback and writing improvement, Patchun and Schunn (2016) found that high ability reviewers gave higher quality peer feedback and had higher text quality, but the differences in text quality did not vary as much as expected given the differences in ability. Low quality papers did not receive more comments, suggesting that reviewers had a threshold in the amount of feedback they offered, despite what was needed. They found a similar threshold in the number of revisions students were willing to make: regardless of ability: students across the study implemented only two-thirds of the feedback they received. They found that both high and low ability students were equally likely to implement revisions, but that low ability students relied more on feedback from their low ability counterparts than from high ability reviewers, emphasizing the role that zone of proximal development may play.

My study differed in that writing improvement was measured by rubric score rather than the number of revisions made, and there was a difference in performance between the groups. Using a median split of the sample based on the pre-revised draft of the technology essay, independent samples t-tests showed that the low ability group had the highest mean grade improvement ($M = 16.89$, $SD = 8.44$) over the high ability group ($M = 11.07$, $SD = 6.84$).

My study aligns with the Patchan & Schunn (2016) study in several ways. In both studies, an initial writing sample was used to assess students’ writing abilities for the sake of grouping. Both studies also found that both high and low ability groups benefitted from peer feedback, albeit in different ways. Their
study indicates that high ability students are better at diagnosing writing problems, and that may help them to write better in the initial stage, but that low ability students use peer feedback to their advantage to catch up. These findings are promising for teachers who need reassurance that low ability students are not left behind in this process. Peer feedback is still of value to high ability students, however, who averaged over 11 points of increase, a letter grade of improvement, because of their effort.

**Implications**

**Personal Implications**

Continuing to gamify peer feedback with the Amazing Race in my classroom will always require preparation, atypical lesson-planning, and effort, but the benefits of it have only been reinforced to me because of this study. Seeing students form quick relationships to win a game challenge creates a congenial classroom environment, a requisite for peer feedback cited in many studies (Marsh, 2018). Making the process competitive was also important because they seem incentivized to help one another improve their writing as part of the game. The most rewarding aspect for me as an ELA teacher, however, has been seeing the sophistication of their comments about a peer’s writing. I would never have guessed they would be such competent writing tutors.

Prior to using the Amazing Race and peer feedback, I relied on the more traditional method of having students write and turn it in to me for a grade. This traditional method asks students to write for an audience of one—the teacher—which led to several negative outcomes. One, was that my students often drafted an essay quickly and relied on me as their sole source of feedback. When I did not give them enough feedback with their grade, they would become upset and
want all the reasons why their writing did not achieve a particular grade.
Involving students in a peer feedback process before a final revision is made is more aligned to the recursive writing process that I want them to follow. With peer feedback, they start to see writing as having a larger purpose than just to receive a grade. They begin to understand how it may communicate meaning to other people such as their peers. They realize that writing can be a process that can be improved between several stages. Most importantly, they see a connection between the improvements they make and their final grade. As a teacher, it has been much more satisfying to me to open an essay, see revisions that students have made, and realize that they are working as hard as I am to improve their writing.

As I write this final evaluation for this study, however, I realize that much of what I have reported appears to be one-sided, mostly the benefits of gamification and using peer feedback. I know that part of that bias comes from all the semesters of hard work that I put into making this game successful. As I explained earlier, this was an instructional activity that I had been developing for years before it became the subject of my action research. It would not have been easy to admit defeat, had my study found no evidence of improvement in my students’ writing.

The years of piloting were necessary, however, in terms of having all the moving parts work well enough to collect the data for this study. Gamification is not easy in the classroom because it requires much more work beyond traditional lesson planning, and it requires the teacher to take on a very non-traditional role of not having full control. Although the game aspect went smoothly, one change I made was in removing the final essay for the class due to end-of-course testing.
requirements they needed to focus on. One area for improvement moving forward is to incorporate this writing test in the culmination of the game.

One of my core values as an educator, however, has always been that learning should be fun and the classroom should be a place for experimentation. Given the results of my study overall, I am more compelled than ever to continue using the Amazing Race to involve students of all ability levels in giving peer feedback on their writing. Because of the positive results of my research thus far, I am even using the game this semester with students who are below grade level in their writing. This should be one of the most convincing reasons for other educators to try gamification, use peer feedback, or do both because as much effort as it takes to do this, the results for me far outweigh the effort and risk to do it.

Implications for the ELA classroom

Gamifying the process may also be unfamiliar territory for many teachers who are not typically trained to use games, much less create them. I can say from experience that I spent a great deal of time calculating points, creating leaderboards, making silly videos, none of which seems like teaching. Spending this time to gamify writing instruction can seem like a risk, even to someone who has researched the value of peer feedback and gamification. Harnessing the power of students to give feedback, however, frees the teacher from the responsibility and allows the teacher more time to facilitate other aspects of the lesson such as gamifying, but whether gamification is the only way to inspire students to give feedback or not remains the subject of future studies.

Allowing students to learn from one another and construct their own knowledge is perhaps one of the strongest aspects of using peer feedback, but
this requires the teacher to step back from direct involvement as the writing
couch, which may be difficult for some teachers. I have learned, however, that
students can take on this role and their feedback to other student writers can be
remarkable. Opening the peer feedback opportunity gives students a larger
audience for their writing. It is also an audience of their peers, whom they
typically prefer over adults, and who are closer to their intellectual ability which
cecourages learning in their zone of proximal development.

Researchers from across the field of education are warning that students
are becoming disengaged with traditional methods of instruction. ELA
classrooms must become places where hierarchies and limiting structures must
be rearranged in favor of more social and collaborative environments (Marsh,
2018). As we innovate and use techniques like gamification, however, Van Eck
warns “if we get it wrong, we may not get a second chance” (2015, p. 24) which
means teachers must be competent to employ the technology that students find
engaging in ways that teach what we know to be important. Whether it is having
our students write evaluations about video games or collaborating on an affinity
space for writing like FanFiction.net, or using gamified peer feedback, teachers
must be open to innovation and mindful of research to remain effective in our
classrooms.

**Implications for Future Research**

As optimistic as these results are, however, there are many facets of the
peer feedback process and gamification that warrant more study, especially
about high school ELA classrooms. Most of the previous studies regarding peer
feedback on writing performance were done in college classrooms and EFL
classrooms, so continued studies about how peer feedback can be used in high
school classrooms is needed. Information about how peer feedback can be used in high school ELA classrooms is especially relevant given the burdens that these teachers face in giving students feedback about their writing. Findings must be replicated to be generalizable to the writing problem evidenced by national assessments in 2011 and 2017 (U.S. Department of Education).

The correlation between comments given and grade improvement was established in this study, but how does writing ability factor in? With larger sample sizes, regression tests could be used to determine the influence of multiple variables like giving feedback and high ability have on writing improvement. This study used mostly quantitative data to measure writing improvement, but a qualitative study at the high school level of actual revisions made would be valuable to the field of research as well. Rather than determining writing improvement, analyzing how peer feedback comments relate to overall writing performance, such as meeting national proficiency standards, needs to be determined.

**Limitations**

While action research offers the teacher-researcher an opportunity to explore the effects of an innovation in their sphere of influence, it limits the generalization of findings because of several factors. The positionality of the teacher-researcher and the lack of a peer rater in the coding and grading was a limitation of the study design. The scale and size of a fifteen-week innovation with so much time and effort is difficult to be unbiased about. Peer rating would increase the reliability of the quantitative data. Further, I felt strongly that all my students should have the opportunity to learn from the innovation and to participate in the gamified aspect of it, but of course this eliminated the
possibility of having a control group. Comparing the gamification of peer feedback to another classroom that used peer feedback in a more traditional way would offer insight into whether gamification of the peer feedback process is necessary.

Another limitation of this study was the demographics of the sample. Honors students do not represent the students of lowest ability in our ELA classrooms. The question of how to help struggling writers is a very relevant concern for most teachers, and one that peer feedback shows promise in being able to address. This study’s findings that the writers of lowest ability in the sample showed the most improvement certainly implies that the benefits of peer feedback could extend to writers of even lower ability, but this study cannot confirm that.

Sample size was another limitation in the study. Limiting the sample to one class allowed for more thick, rich description of the peer feedback comments, which was very valuable to the interpretation of findings. The disadvantage to a smaller sample was not having normal sample distributions in the data for the statistical tests. Although the correlations and differences among groups were intriguing and aligned to previous studies in some ways, the results were not statistically significant. Larger sample sizes with more diverse participants would provide more insight into whether gamification played a crucial role in engaging students with substantive feedback or whether other methods could produce the same results. A study that utilized a control group, or qualitative data obtained through an interview or survey about the students’ feelings about the game, would provide more insight into the role that gamification played.
Closing Thoughts

Writing is an effortful task, and writers crave to know that someone hears them. As teachers, it is incumbent on us to support that process, but this is an overwhelming and unrealistic task when we are the only ones reading our students’ writing. Gamification is a powerful way to engage students, but the more important take-away from this study is the role of peer feedback. Involving students in the writing process through peer feedback is proving to be a powerful tool that writing teachers can use to elicit several benefits. Involving students’ peers supports the dialogic process that we know writing is—a social and communicative process that should not be limited to student and teacher. The more people we can involve in the dialogue, the more meaningful the writing process becomes. We often talk in ELA about giving students authentic writing opportunities, which means, having them write for real audiences so they begin to understand what writing is—communication over space and time. Although the benefits of giving and receiving peer feedback are becoming well-documented, giving students a larger audience for their writing is an added advantage aspect of peer feedback at the high school level.

As we advance in the digital age, social media is becoming an all-consuming entity for our students who are searching for ways to express themselves. As writing teachers we want to teach them the power that formal writing has in making their voices more effective. GPF is a tool that we can use to combat the writing deficiencies we are seeing in students and engage them in learning to write well. In doing so, we are sharing the power that we know comes with it.
REFERENCES


Chanski, S., & Ellis, L. (2017). Which helps writers more, receiving peer feedback or giving it? The English Journal, 106(6), 54-60.


http://dspace.hebron.edu:80/xmlui/handle/123456789/69


Jones, S. (2013) Gamification vs. Game-Based Learning - Theories, Methods, and Controversies, [Conference Presentation] Symposium for Teaching and
Learning with Technology Conference. Metro State University of Denver. Denver, Colorado, United States.


APPENDIX A

CONSENT FORM AND IRB APPROVAL

UNIVERSITY OF SOUTH CAROLINA

CONSENT TO BE A RESEARCH SUBJECT

The Effects of Gamified Peer Feedback on Student Writing in High School English Language Arts

KEY INFORMATION ABOUT THIS RESEARCH STUDY:
You are invited to volunteer for a research study conducted by Kerise Broome, an English teacher at James Island Charter High School. I am also doctoral candidate in the Department of Education, at the University of South Carolina. The purpose of this study is to evaluate the impact of gamified peer feedback in a high school English Language Arts classroom. You are being asked to participate in this study because you are a tenth grade Honors English student in my course.

The following is a short summary of this study to help you decide whether to be a part of this study.

PROCEDURES AND DURATION
This study will take occur from week 3 to week 18 of our semester English course. As a student, you will be participating in four essay cycles every four weeks in which you will write an essay, give peer feedback to other students, and revise your essay as part of the requirements for class. I will collect peer feedback data and grades during each of the four essay cycles for all students who consent to the study and analyze this data at the end of the semester. If you agree to participate in this study, you are allowing your information (peer feedback comments and your writing grades) be analyzed to determine if the activities we did helped my students to improve their writing. If you allow your information to be used, there is nothing additional that you must do, and all identifying information including your name will be removed from the final published results so that all student information is anonymous.
RISKS/DISCOMFORTS:
While there is minimal risk to all students that district software with grades and personal information could be compromised, there is no additional risks to students in this study considering that all information such as student work and grades will be kept on district-issued software and a password-protected computer.

BENEFITS:
Taking part in this study by allowing your information to be used is not likely to benefit you personally. However, this research may help me and other teachers understand how games can enhance classroom activities like peer feedback and whether peer feedback can help students write better.

CONFIDENTIALITY OF RECORDS:
Information obtained about you during this research may be published, but you will not be identified. Information that is obtained concerning this research that can be identified with you will remain confidential to the extent possible within State and Federal law. The investigators associated with this study, the sponsor, and the Institutional Review Board will have access to identifying information. All records in South Carolina are subject to subpoena by a court of law. Study information will be securely stored in locked files and on password-protected computers.

VOLUNTARY PARTICIPATION:
Participation in this research study is voluntary. You are free not to participate, or to stop participating at any time, for any reason without negative consequences. Your grade in this course will be not be affected by whether or not you give consent for your information to be used. If you wish to withdraw from the study at any time, please email Mrs. Broome.

I have been given a chance to ask questions about this research study. These questions have been answered to my satisfaction. If I have any more questions about my participation in this study, I can contact Kerise Broome at kerise_broome.c@charleston.k12.sc.us.

Concerns about your rights as a research subject are to be directed to, Lisa Johnson, Assistant Director, Office of Research Compliance, University of South Carolina, 1600 Hampton Street, Suite 414D, Columbia, SC 29208, phone: (803) 777-6670 or email: LisaJ@mailbox.sc.edu.

I agree to participate in this study. I have been given a copy of this form for my own records.

If you wish to participate, you should sign below.
Kerise Broome
3458 Walter Drive
Johns Island, SC 29455

Re: Pro00117731

Dear Mrs. Kerise Broome:

This is to certify that research study entitled The Effects of Gamified Peer Feedback on Student Writing in High School English Language Arts was reviewed on 12/10/2021 by the Office of Research Compliance, which is an administrative office that supports the University of South Carolina Institutional Review Board (USC IRB). The Office of Research Compliance, on behalf of the Institutional Review Board, has determined that the referenced research study is not subject to the Protection of Human Subject Regulations in accordance with the Code of Federal Regulations 45 CFR 46 et. seq.

No further oversight by the USC IRB is required. However, the investigator should inform the Office of Research Compliance prior to making any substantive changes in the research methods, as this may alter the status of the project and require another review.

If you have questions, contact Lisa M. Johnson at lisaj@mailbox.sc.edu or (803) 777-6670.

Sincerely,

Lisa M. Johnson
ORC Assistant Director and IRB Manager
APPENDIX B

Analytic Scoring Rubric for Phase One Technology Essay

<table>
<thead>
<tr>
<th>Component</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>Final Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Thesis and overall analysis</strong></td>
<td>The writer's intentions are difficult to discern. Attempts at analysis are unclear or irrelevant.</td>
<td>The argument's thesis reflects some clarity in thought and purpose. Analysis is simplistic or somewhat unclear.</td>
<td>The argument's thesis reflects clarity in thought and purpose; analysis recognizes implications and tensions.</td>
<td>The argument's thesis reflects precision in thought and purpose. The analysis addresses implications and tensions.</td>
<td>The argument's thesis reflects nuance and precision in thought and purpose. The argument establishes and employs an insightful context for analysis of the issue and its perspectives. The analysis examines implications, complexities and tensions, and/or underlying values and assumptions.</td>
<td>Development of ideas and support for claims are weak, confused, or disjointed.</td>
<td>Development of ideas and support for claims are mostly relevant but are overly general or simplistic.</td>
<td>Lines of clear reasoning and illustration adequately convey the significance of the argument. Qualifications and complications extend ideas and analysis.</td>
<td>A mostly integrated line of purposeful reasoning and illustration capably conveys the significance of the argument. Qualifications and complications enrich ideas and analysis.</td>
</tr>
<tr>
<td><strong>Development and Support</strong></td>
<td>Ideas lack development, and claims lack support.</td>
<td>Development of ideas is inconsistent and often unclear. Transitions between and within paragraphs are misleading or poorly formed.</td>
<td>Development of ideas is consistent and ideas are logically grouped. Transitions between and within paragraphs clearly clarify the relationships among ideas.</td>
<td>Ideas are logically grouped and sequenced. Transitions between and within paragraphs clarify the effectiveness of the argument. Transitions between and within paragraphs consistently clarify the relationships among ideas.</td>
<td>Development of ideas and support for claims deepen insight and broaden context. An integrated line of skillful reasoning and illustration effectively conveys the significance of the argument. Skillful organizational strategy. The response is unified by a controlling idea or purpose, and a logical sequencing of ideas contributes to the effectiveness of the argument. Transitions between and within paragraphs consistently clarify the relationships among ideas.</td>
<td>Organization</td>
<td>Grouping of ideas is inconsistent and often unclear. Transitions between and within paragraphs are misleading or poorly formed.</td>
<td>The response largely coheres, with most ideas logically grouped. Transitions between and within paragraphs clearly clarify the relationships among ideas.</td>
<td>Ideas are logically grouped and sequenced. Transitions between and within paragraphs clarify the effectiveness of the argument. Transitions between and within paragraphs consistently clarify the relationships among ideas.</td>
</tr>
<tr>
<td><strong>Language Use</strong></td>
<td>Sentence structures are often unclear. Errors in grammar, usage, and mechanics are pervasive and often impede understanding.</td>
<td>The use of language is consistent and often unclear. Word choice is rudimentary and frequently imprecise. Sentence structures are sometimes unclear. Distracting errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding.</td>
<td>Word choice is general and occasionally imprecise. Sentence structures are usually clear but show little variety. Distracting errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding.</td>
<td>The use of language conveys the argument with clarity. Word choice is adequate and sometimes precise. Sentence structures are clear and demonstrate some variety. Errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding.</td>
<td>Word choice is precise. Sentence structures are clear and varied often. While minor errors in grammar, usage, and mechanics may be present, they do not impede understanding.</td>
<td>Language Use</td>
<td>The use of language is consistent and often unclear. Word choice is rudimentary and frequently imprecise. Sentence structures are sometimes unclear. Distracting errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding.</td>
<td>The use of language conveys the argument with clarity. Word choice is adequate and sometimes precise. Sentence structures are clear and demonstrate some variety. Errors in grammar, usage, and mechanics may be present, but they generally do not impede understanding.</td>
<td>Word choice is precise. Sentence structures are clear and varied often. While minor errors in grammar, usage, and mechanics may be present, they do not impede understanding.</td>
</tr>
</tbody>
</table>

118
APPENDIX C

Analytic Scoring Rubric for The Kite Runner Essay

<table>
<thead>
<tr>
<th>Analysis of text (25 points)</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction paragraph (15 points)</td>
<td>Introduction paragraph is almost meets the standard</td>
<td>Introduction establishes the topic of the paper and gives appropriate background; the paragraph should 'hook' the reader by making paper seem interesting and sophisticated analysis of literary device is revealed</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of text (35 points)</td>
<td>paper meets the required length and content is analytical versus plot summary; well-selected examples from the book are given; understanding of theme and its development is emerging but may not be clear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>organization (25 points)</td>
<td>ideas are logically organized for the most part but some ideas may be out of order or some paragraphs may be underdeveloped</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>grammar and sentence structure (10 points)</td>
<td>several errors in grammar or sentence structure are made</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>revision (15 points)</td>
<td>although some revision is demonstrated, errors reveal that more revision was needed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Final Grade

points detracted if paper was late
APPENDIX D

Analytic Scoring Rubric for Research Paper

<table>
<thead>
<tr>
<th>Research Paper Rubric</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td><strong>Content Development</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
</tbody>
</table>

| **Incorporating Research** |
| 1 | The writer fails to include information that connects to thesis and clearly supports ideas. |
| 2 | Development of ideas and support for thesis are weak, confused, or disjointed. |
| 3 | The writer includes information that connects to thesis and clearly supports ideas. Development of ideas and support for thesis are mostly relevant but are overly general or simplistic. |
| 4 | Content is mostly well-supported and informative and well-connected to a strong thesis. Research (0-2 citations) is properly cited in text and includes direct quotes, indirect quotes, paraphrasing, and primary sources. Development of ideas and support for thesis are deeper. Mostly distinguishing between common knowledge and information that should be cited. |
| 5 | All content is extremely well-supported and insightful. Information connects to an insightful thesis, and the writer expertly integrates their own ideas and research. Research (0-2 citations) is properly cited in text and includes direct quotes, indirect quotes, paraphrasing. Writer expertly distinguishes between common knowledge and information that should be cited. |

| **Organization** |
| 1 | Does not exhibit an organizational structure, minimal grouping of ideas or transition. |
| 2 | Exhibits a basic organizational structure, most ideas logically grouped. Transitions between and within paragraphs are missing or poorly formed. |
| 3 | Shows clear organization, reflects emergent controlling idea or purpose. Ideas are logically grouped and sequenced. Transitions between and within paragraphs clarify relationships among ideas. |
| 4 | Shows productive organization, mostly unified by a controlling idea or purpose, and a logical sequencing of ideas contributes to the effectiveness of the argument. Transitions between and within paragraphs clarify relationships among ideas. |

| **Language Use** |
| 1 | The use of language is inconsistent and often unclear. Word choice is imprecise and often difficult to comprehend. Sentence structures are often unclear. Stylistic and register choices are difficult to identify. Errors in grammar, usage, and mechanics are pervasive and often impede understanding. |
| 2 | The use of language is basic and only somewhat clear. Word choice is general and occasionally imprecise. Sentence structures are usually clear but show little variety. Stylistic and register choices, including voice and tone, are not always appropriate for the rhetorical purpose. Distracting errors in grammar, usage, and mechanics are present, but they generally do not impede understanding. |
| 3 | The use of language conveys the argument with clarity. Word choice is adequate and sometimes precise. Sentence structures are clear and demonstrate some variety. Stylistic and register choices, including voice and tone, are appropriate for the rhetorical purpose. While errors in grammar, usage, and mechanics are present, they rarely impede understanding. |

| **Formatting** |
| 1 | Information is not cited properly resulting in plagiarism. Parenthetical citations do not match the Works Cited. Works Cited not accurately formatted (MLA, APA) or there may be fewer than the required sources. |
| 2 | Formatting throughout paper is attempted, but not correct MLA. Formatting throughout paper is mostly correct, but some small errors were made. |

| 3 | Formatted and cited properly in MLA format, and all formatting is correct throughout paper. |

**final grade**