Practice and Efficacy of Peer Writing Feedback in a Large First-Year Engineering Course

Michael Roman Ekoniak III

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Engineering Education

Marie C. Paretti
Holly Matusovich
Elizabeth D. McNair
Katrina M. Powell

4/20/2018
Blacksburg, Virginia

Keywords: Engineering Education, Writing in the Disciplines, Peer Feedback,

Formative Assessment, Writing Process

Practice and Efficacy of Peer Writing Feedback in a Large First-Year Engineering Course

Michael Roman Ekoniak III

ABSTRACT

Engineering educators and industry professionals recognize the need for graduates to be effective communicators, yet the effective teaching of communication remains a persistent contemporary issue, with studies continuing to indicate that engineering graduates are insufficiently prepared for workplace communication. Despite compelling arguments that that writing should be treated as a situated activity, writing instruction is often separated from content instruction within engineering curricula. Even when they are integrated, it is often in a way does not optimally support improvement of students' writing skills. Writing studies scholarship identifies best practices that include treating writing as a process, with pedagogy that includes writing and revising drafts based on feedback and revision. However, writing assignments in engineering courses often not process-oriented.

Challenges in addressing this problem include epistemology (i.e. instructors believe that learning to write and learning to engineer are separate skills), self-efficacy (i.e. instructors not feeling qualified to effectively provide feedback or writing instruction), and resources (i.e. inclusion of feedback and revision is unfeasible within key constraints of many engineering courses – limited instructor time and large student-faculty ratios).

A potential solution is to use peer feedback, where students provide each other feedback on drafts, which can support a process approach while addressing these challenges. Research outside engineering has demonstrated that peer feedback can be as or more effective than instructor feedback; to bring that work into engineering, this study examines peer feedback in the context of a first-year engineering course through a quasi-experimental intervention in which feedback and revision were incorporated into an existing assignment using several variations of peer feedback. Interventions included two types of feedback training as well as feedback from single peers and multiple peers.

Results support peer feedback in this context: it was statistically indistinguishable from instructor feedback when students were given sufficient instruction. Feedback from multiple peers, in fact was more effective than instructor feedback in improving writing quality, and in-class instruction was more effective than a handout only in helping students provide effective feedback. However, some general feedback recommendations – notably that readerly feedback should be encouraged directive feedback discouraged – were not supported. While writing studies literature encourages feedback that takes the position of the reader, readerly peer feedback reduced revision quality in this study. Directive feedback, on the other hand, caused improvements in writing quality, supporting previous hypotheses that the tightly-constrained genres in which engineers write justify more use of directive feedback.

Practice and Efficacy of Peer Writing Feedback in a Large First-Year Engineering Course

Michael Roman Ekoniak III

GENERAL AUDIENCE ABSTRACT

Engineering graduates must be effective communicators, yet studies continue to indicate that engineering graduates are insufficiently prepared for workplace communication. This indicates that the effective teaching of communication remains an important area of research. While research shows that writing should integrated into engineering content courses to best facilitate learning, it is common for writing to be seen as a separate topic of study. Even when they are integrated, best practices of writing pedagogy developed in writing studies research are not used: writing assignments in engineering courses often include no feedback on writing and no revision, which are important for students' writing skill development. An additional challenge is lack of resources: limited instructor time and large class sizes in many engineering courses make providing opportunities for feedback and revision a challenge.

This study examines one potential remedy to these challenges: using peer feedback, where students provide each other feedback on drafts. Results support using peer feedback: it was as effective as instructor feedback when students were provided with sufficient instruction. Additionally, feedback from multiple peers was *more* effective than instructor feedback. Overall, this research supports the use of peer feedback in large engineering courses.

Dedication

To my parents, Loretta Ekoniak and Michael Ekoniak Jr., for always encouraging me to follow my curiosity; you made me a lifelong-learner before I even understood what that means.

#foreverastudent

Acknowledgements

To my advisor, Dr. Marie Paretti, I cannot even begin to fully express my gratitude. Through your example, I learned what kind of scholar I could hope to be. Every time I tried to give up, you both made me feel it was OK to do so, yet your confidence in my ability to persist drove me to finish. Your advice and encouragement made me believe in myself, but your empathy and understanding were the critical factors I'd not be here without.

To my committee members, Holly Matusovich, Lisa McNair, and Katy Powell. Holly, I don't think I really understood what research really was before becoming your student; you helped me appreciate the salience of the "Ph" in Ph.D. Lisa, you were the first person I ever learned about the field of engineering education from and made me interested in being a part of it. Katy, you introduced me to ways of thinking about the world that have forever changed my perspective.

To Jean Mohammadi-Aragh and Molly Scanlon for collaborating on the original design and implementation of this study and Loretta Ekoniak for assisting with analysis.

To the amazing people I've formed relationships with throughout my tenure at Virginia Tech. I didn't go to grad school with any intent to pursue a doctorate, and without your friendships I'd probably not have begun the journey and certainly not have finished it. One of the great tragedies of grad school is meeting such amazing people and then dispersing throughout the world.

Molly Scanlon, I'd never even have started this process without your encouragement; while our lives have since diverged, I cannot overstate my gratitude for you being a part of it. To the other members of my OG rhet/comp family, Matt Sharp and Megan O'Neill, you influenced every aspect of who I am today; whether our intellectual conversations or the nights we barely remember was the greater influence I will never know but I'd not take either back for anything.

To my fellow engineering education nerds. Ben Lutz, I think I can safely say that without each other neither of us would be the people, scholars, and educators we are and strive to be. Jean Mohammadi-Aragh, Phil Brown, Chris Venters, Kevin Sevilla, Andrea Goncher, and Rachel Kajfez, we were the end of an era in EngE (there were no graduates yet when we started!). Y'all make me look forward to June every year.

Mandy Wright, you showed me that we have as much to learn from our emotional side as our analytical side. Rachel Corell, I'm glad I was sorted into your house, and that you made me appreciate it; our many awesome trips together kept me sane and got me through. Liz Misiewicz, I'd never have passed my prelim if I hadn't met you – at a time I struggled to write a sentence, our conversations prepared me to convince my committee I was ready to do this.

Table of Contents

List of Figures

List of Tables

Glossary of Terms

Contemporary Issue Report: The writing assignment that was the focus of this study was the Contemporary Issue Report (CIR). The terms *Contemporary Issue Report*, *CIR*, and *report* are used interchangeably in this document.

Draft: The initial CIR as submitted by students prior to revising.

Revision: Can refer either to the complete revised draft submitted by students after the intervention (synonymous with *final*), or to an individual change made to the draft when revising.

Nature of Revisions: The types of revisions an author makes when revising their draft and their classification based on the codes described in Section 3.6.

Nature of Feedback: The types of feedback a reviewer gives on an author's draft and its classification based on the analytic framework coding frame described in Sections 2.5 and 3.5.2.

Feedback Conditions: The two variables manipulated for the study intervention: *peer review instruction* and *peer review type.* Crossing the conditions results in *groups*.

Peer Review Instruction: The method by which students were taught to give peer feedback, either *handout-only* or *in-class plus handout*.

Handout-only: Peer review instruction consisting of a sheet of instructions (handout) on how to provide quality peer feedback that was distributed via the course management system.

In-class Plus Handout: Peer review instruction consisting of an in-class workshop in addition to the handout, which was received by all peer reviewers.

Peer Review Type: Refers to whether a student provided feedback to and received feedback from one peer (single-peer) or two peers (multiple-peer).

Group: Crossing the feedback conditions results in four *groups* of students who received a specific combination of feedback conditions (e.g. handout-only, multiple-peer). An additional baseline group received feedback from their instructor.

Feedback Item: An individual unit of feedback, as defined in section 3.5.1. During analysis, each feedback item was classified by its *feedback features,* as defined in the coding frame for feedback (see section 3.5.2). The term feedback item is used in this document when it is necessary to refer to units of feedback that are either *comments* or *edits.*

Comment: A comment is a feedback item consisting of a reviewer's written remarks on or about an author's paper. In the analysis phase of this study, comments were classified along five dimensions of the coding frame for feedback items: mode, tone, focus, localization, and type.

Edit: An edit is a feedback item consisting of a reviewer's direct changes to or proofreading marks on an author's paper. In most literature, edits fall under the broader definitions of feedback or comments, but they are distinguished from comments in the analysis conducted for this study because they do not vary within the dimension of the coding frame (i.e. it makes no sense to distinguish "positive edits" and "negative edits")

Chapter 1: Introduction

1.1. Statement of the Problem

Effective written communication skills are essential for engineers, as is widely recognized in the field of engineering education and by US and international engineering professional and program accreditation organizations (e.g. ABET, 2010; British Computer Society, 2012; IEEE, 2004). Since implementation of the EC2000 accreditation criteria, communication has been also been explicitly required by ABET for engineering programs. Specifically, criterion 3g states that programs must demonstrate that their graduates develop "an ability to communicate effectively" (ABET, 2010), and even though controversial proposed revisions to ABET criteria set to take effect in 2019 have been accused of "watering down" the other professional skills, communication remains a separate and even somewhat expanded outcome (Flaherty, 2015). In addition to its importance as a professional skill, learning to communicate as an engineer is tightly bound with the formation of one's engineering identity, a process known as development of a discursive identity (Allie et al., 2009).

Although nearly two decades have passed since adoption of EC 2000, studies continue to indicate that employers view engineering graduates as insufficiently prepared for communication in the workplace (e.g. Donnell, Aller, Alley, & Kedrowicz, 2011). At the same time, many engineering graduates continue to separate communication skills from 'real engineering' in their conceptualization of workplace tasks, even while being aware of the centrality of those skills to their work (Anderson, Courter, McGlamery, Nathans-Kelly, & Nicometo, 2010; Trevelyan, 2010).

One reason for this state of affairs is that the teaching of writing is often separated from the teaching of content knowledge within engineering curricula (Leydens & Schneider, 2009, p. 260). Many scholars (Artemeva, Logie, & St-Martin, 1999; Bazerman, 2015; Paretti, McNair, & Leydens, 2014) emphasize that writing pedagogy should be based on an understanding of writing as a *situated activity* – one in which the knowing of something is inseparable from the doing and therefore benefits from being learned in contexts similar to those in which it will be practiced (Brown, Collins, & Duguid, 1989). Thus, separation of "learning to write" from "learning to engineer" in engineering curricula is problematic for both skill and identity development.

Even when communication and content are integrated in the same courses, this integration is often done in a way that does not support improvement of students' writing skills as effectively as it could. Best practices of writing pedagogy treat writing as a process rather than a product; this approach includes writing and revising drafts based on formative assessment (feedback) rather than receiving only summative evaluative feedback of a final product – often referred to as the "one-and-done" approach (Chandrasegaran, 2008; Murray, 1972). Incorporating revision in the writing process is recognized as an invaluable component of improving written work, but writing assignments in engineering courses frequently neither acknowledge nor incorporate this process (Swarts & Odell, 2001; Zhu, 2004). Although there are a variety of reasons for this gap, part of the challenge rests with engineering instructors, who may not feel qualified to effectively provide feedback or who may perceive that draft feedback is unfeasible within key constraints of many engineering courses: limited instructor time and large student-faculty ratios (Matusovich, Paretti, Motto, & Cross, 2012).

One potential way to address these instructor concerns is to use guided peer feedback, where other students in the class provide feedback on students' writing rather than the instructor.

Peer feedback can be done outside of class as a homework assignment, eliminating the burden on class time. At the same time, online tools and general guidance can alleviate concerns about instructor competence, and many universities have resources available through writing centers serving students across disciplines to assist instructors with integration of writing assignments in their classes. Equally important, recent research in domains other than engineering has also shown that peer feedback can be as or more effective than instructor feedback when student writing is reviewed by multiple peers rather than a single peer, potentially assuaging concerns about the utility of peer feedback (K. Cho & MacArthur, 2010; Nelson & Schunn, 2009; Patchan, Schunn, & Clark, 2011). Moreover, more fully integrating writing and rewriting as a collaborative professional practice could help graduating engineers more readily recognize writing as engineering work.

1.2. Research Questions and Approach

While multi-peer feedback has been shown to be useful in other domains, little scholarship has explored its value in engineering. Because the use of peer feedback could enable the closer integration of learning to engineer and learning to write as an engineer, the application of this approach within engineering courses would benefit from systematic study. Thus, in order to make informed recommendations for using peer review to improve the teaching of written communication, the overarching goal of this study is

> **to better understand how peer review of writing can be used to effectively support development of engineering students' writing.**

This goal can be viewed as an optimization problem: feedback and revision are important components of writing pedagogy, but the resources required to provide feedback increase with student to instructor ratio, which makes instructor feedback impractical in large courses. Peer review can address this constraint but also consumes student resources; multiple-peer review, in which students provide feedback to and receive feedback from multiple peers, further increases student resources. At the same time, peer review can also increase resources required of the instructor: higher if in-class time must be used for training, albeit as a fixed cost to instructors rather than increasing with instructor-student ratio, but lower if training can be provided via a handout or similar external resource. To address the overarching goal of this study, I use a quasi-experimental design (Creswell, 2009), manipulating inputs to each process in order to purposefully generate data on peer review within a range of resource costs. Inputs manipulated included peer feedback instruction (handout-only, in-class plus handout) and number of peer reviews (single peer, multi-peer), as shown in Figure 1.

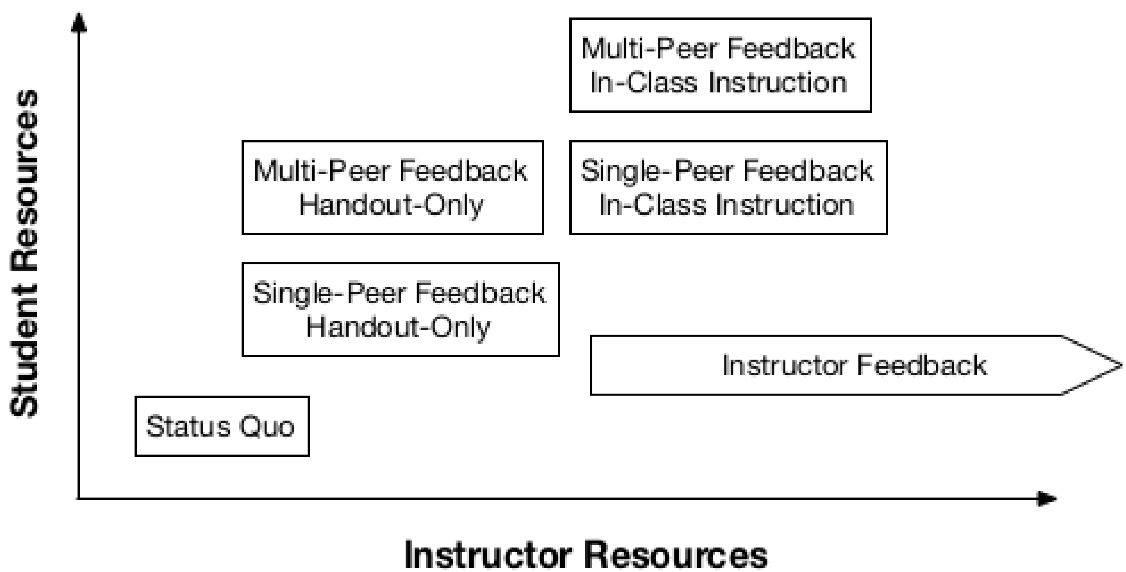


Figure 1: Resource (time, effort, etc.) Costs of Various Kinds of Feedback

Within this framework, I answer the following research questions:

RQ1: How do different types of feedback instruction (in-class + handout, handout-only) affect the practice of peer review (i.e. the nature of reviewer feedback) for different peer review types (single peer, multi-peer)?

RQ2: What factors influence the efficacy of peer review (i.e. the nature of authorial revisions)?

To approach these questions, the study is situated in a large first-year engineering (FYE) course for four reasons. First, FYE courses are often large and resource-constrained. Second, these courses are students' entry point into the engineering curriculum and thus a useful site to begin understanding how students develop between beginning their education and graduation; although this study did not address this development the results provide a starting point for future work. Third, FYE programs often explicitly include communication in student learning outcomes. Finally, by integrating communication into students' first exposure to an engineering course, the centrality of writing to engineering practice is made explicit as early as possible. Understanding peer review in this context, especially with the challenges associated with large classes (Topping & Ehly, 1998) and the well-documented challenges students face in moving from secondary to university academic discourse (Bielenberg, 2011; Lea, 1998), is especially relevant to the field of engineering education.

To answer the research questions, I analyzed a corpus of 119 research reports written by first-year engineering students. This corpus consists of students' first drafts, feedback on those drafts, and final reports. Feedback was received from either a single peer or from multiple peers based on the quasi-experimental approach described above; an additional 47 reports that were revised based on instructor feedback were also analyzed for comparison. I characterized the feedback and subsequent revisions using qualitative content analysis, which is a method to transform qualitative data such as text into categorical quantitative data appropriate for analysis with standard statistical techniques (Schreier, 2012). In addition to the categorization of feedback and revisions, changes from draft and final reports were scored using an instrument based on the rubric students and instructors used for the original assignment; thus, when determining whether and how the quality of reports changed from draft to final submission, analysis was aligned with the same criteria students used to make those revisions.

1.3. Intellectual Merit

This dissertation contributes to the bodies of knowledge in engineering education and writing studies in several important ways. I conducted a quasi-experimental study that varied both peer review instruction (training via either in-class workshop or only and handout) and peer review type (single peer or multiple peer) in a large first-year engineering design course and analyzed the effect on the types of peer feedback students made and the resulting improvement from initial draft to final submission. Related work of this type has been conducted in other disciplines (e.g. K. Cho & MacArthur, 2010), and within engineering professional communication (e.g. Artemeva & Logie, 2003), and upper-level content courses (e.g. Jensen & Fischer, 2005), but to date only my Master's project (Ekoniak, 2014) has characterized peer feedback in a first-year engineering setting.

To answer the first research question, I varied both peer review instruction (training via either in-class workshop or only and handout) and peer review type (single peer or multiple peer) in a large first-year engineering design course and analyzed the effect on the types of peer feedback students made. To establish a baseline for comparison, some students received feedback from their instructor who had not received any special training on providing feedback on writing. There were five major findings related to the first research question:

1. Multiple-peer feedback was at least as effective as instructor feedback in terms of improvement from draft to final version across multiple dimensions (technical quality, writing quality, and mechanics).
2. For writing quality, multiple-peer feedback was more effective than instructor feedback.
3. In-class feedback instruction improved the practice of peer review over handout-only.
4. Most of the improvements in the practice of peer review resulting from the in-class instruction were only statistically significant when a student reviewed multiple reports.
5. Multiple peer review was the only factor predicting corrections based on assignment rules.

To answer the second research question, I analyzed how the types of feedback students received predicted they types of revisions students made and the resulting quality improvement. I also considered indirect effects of peer review instruction and type on revisions and quality improvement. There were four major findings related to the second research question:

1. Directive feedback, in which the writer tells the reader what to do, was more helpful for quality improvement than anticipated.
2. There were no quality metrics for which readerly comments (those where the reviewer takes the perspective of a reader rather than an evaluator) were a positive predictor.
3. Localization (i.e. the degree to which a comment is directly linked to a specific section of text to improve) is important.
4. Macrostructure revisions (e.g. major changes in content or organization), while important for quality improvement, were not predictable.

These nine major findings and their relationship to current literature are summarized in Table 36.

These results extend previous research (K. Cho & MacArthur, 2010; K. Cho, Schunn, & Wilson, 2006; Patchan, 2011) by demonstrating that peer feedback is as or more effective than instructor feedback to the context of engineering. Thus, several problematic status-quo aspects of engineering communication curricula – the typical separation of communication and content learning (Leydens & Schneider, 2009, p. 260), adoption of a "transmission model" of communication (Paretti et al., 2014, p. 614; Leydens, 2008) and challenges for improving these practices due to resource constraints such as limited instructor time and large student-faculty ratios (Matusovich et al., 2012) – can be addressed by using peer feedback to enable the adoption of process-oriented writing assignments in engineering courses.

These results also underscore the importance of using feedback from multiple peers rather than just a single peer and training peers to provide quality feedback. Previous studies (e.g. K. Cho & MacArthur, 2010; K. Cho & Schunn, 2007; K. Cho, Schunn, & Wilson, 2006) have shown that multiple-peer feedback is more effective than single-peer both for both providing and implementing feedback; the results of this study showed that feedback from two peers did result in the provision of better feedback; further investigation would be needed to determine whether an increased number of peers could further improve revisions based on that feedback, as seen in studies done in other contexts. The benefits of training peers to provide feedback has been empirically tested previously (Ozogul & Sullivan, 2007; Sluijsmans, Brand-Gruwel, & Merriënboer, 2002; Weaver, 1995), but those studies did not examine differences in efficacy for alternative types of training. This study showed that training using an in-class workshop is preferable to using only a handout with respect to peers following the training recommendations, though even the handout-only group showed quality improvements. However, the results also indicate that some general recommendations for writing feedback – notably an emphasis on readerly feedback and discouraging directive feedback – that are applicable to feedback from experts may not be applicable for peers. While writing studies literature encourages feedback that takes the position of the reader (Straub & Lunsford, 1996a), and previous studies have supported the applicability of this recommendation to peer feedback (e.g. K. Cho & MacArthur, 2010), this study found that readerly peer feedback actually reduced revision quality among first-year engineering students. These findings represent important contributions to writing studies as they help nuance generalized findings across university writing to specific disciplinary contexts. Following the work of Patton and Smith Taylor (Patton & Smith Taylor, 2013), this study demonstrates that effective feedback may not be generic across disciplines or academic years, and in turn deepens our understanding of situated writing pedagogies.

1.4. Summary of Remaining Chapters

In this chapter, I provided an overview of the problem that this dissertation addresses: written communication is an essential engineering skill and has been the subject of increasing focus within the curriculum, but reports suggest that major gaps still exist. In part, these gaps may arise from the common use of a "one and done" approach to writing assignments, typically due to resource constraints, that does not match best practices from writing studies. My research suggests that peer review introduced with either handouts or dedicated in-class instruction could be used to address these constraints while improving writing pedagogy by incorporating feedback and revision in these assignments. I also provide an overview of the approach, research questions and methods I used to better understand how peer feedback can be used in first-year engineering.

In Chapter 2, I review prior work relevant to this research. First, I discuss the importance and history of engineering communication. Next, I describe key best practices in writing pedagogy, summarize research that indicates that practices in engineering curricula conflict with best practices, and discuss barriers that cause and maintain this conflict. Finally, I argue, using theory and empirical evidence, that peer review can address these barriers and potentially improve writing pedagogy in engineering.

In Chapter 3, I present my research methods. I begin by discussing my research questions and quasi-experimental research design in depth. Next, I provide contextual details of the implementation of the quasi-experimental design and data collection. In the second half of the chapter, I describe how I conducted the analysis, including an overview of how qualitative content analysis (QCA) is used to transform qualitative into quantitative data and a detailed description of how QCA was operationalized in this study. Next, I discuss the quantitative analysis, including choice of statistical techniques. Finally, I conclude with a discussion of the trustworthiness and limitations of this research design.

In Chapters 4 and 5, I present and discuss the results of the resulting analysis. In Chapter 4, I begin by comparing the draft-revision quality improvement demonstrated by students who participated in peer review with those who received traditional instructor feedback in order to validate the assumption that the equivalence between peer review and instructor feedback reported in previous studies is transferrable to the first-year engineering context. Next, I analyze the effects of peer feedback instruction and multiple vs single peer review on the practice of peer review in order to answer the first research question. Finally, I answer the second research question by analyzing which factors influence the efficacy of peer review, i.e. the causes of quality improvement. In Chapter 5, I first discuss these results in the context of extant literature and describe how this study confirms, extends, or contradicts current knowledge. Next, I discuss the implications of this study for engineering curricula and instructors. Finally, I identify the limitations of this study and suggest future directions for research.

Chapter 2: Literature Review

Written communication skills are essential for engineering graduates. The literature provides strong support for this claim, but also identifies many ways that common practices in engineering education insufficiently address this need. The goal of this study is to provide evidence-based recommendations for improving writing pedagogy in engineering education classrooms via peer review.

To lay the foundation for this work, I begin this chapter by discussing the importance of engineering communication and the continued mismatch between engineering graduates' communication skills and workplace expectations. Next, I identify salient practices from writing pedagogy that conflict with common practices in engineering curricula. In the third section, I consider epistemic and practical barriers to adoption of these best practices. Finally, I provide evidence suggesting that the use of peer review can address some of these issues.

2.1. Communication as a Core Engineering Skill

Engineering educators have acknowledged the importance of communication as an engineering professional skill from the earliest days of formalized engineering education (Kynell-Hunt, 2000). Despite this acknowledgement, there has been a lack of consensus on the role, location, and practice of writing instruction in engineering curricula. Studies continue to demonstrate the centrality of communication skills in engineering work, but engineering culture often divides skills into a dichotomy of *hard* and *soft*, with "soft" skills such as communication hard seen as separate from – and lesser than – the "hard" engineering skills (Leydens & Schneider, 2009, p. 260,261; see also Anderson et al., 2010; Trevelyan, 2010). Even the more recent shift in language to "technical" and "professional" skills maintains both the dichotomy and the perceived hierarchy. These binary views combined with constantly increasing technical content in engineering courses make it difficult to incorporate communication skills into the core engineering coursework, despite compelling arguments that communication skills are best learned in context (Paretti et al., 2014). This tension drives the need to continue to explore effective approaches to communication instruction throughout the engineering curriculum.

2.1.1. Foundations of "Teaching Writing" in Engineering

To situate this need, it is helpful to first understand the history of communication instruction in engineering and how it has evolved over time. For most of the 19th century, American engineers learned informally through apprenticeship or vocational training (Seely, 1999). With the development of the land-grant system, which was specifically enacted to promote practical and economically valuable knowledge (National Research Council, 1969), engineering began to move into the domain of higher education and thus professionalization. Under these conditions, the first literacy concerns in engineering were motivated as much by status concerns – the desire to be seen as equals by the other professions – as they were by their applicability to engineering work (Kynell, 1999, p. 144). By the time of the inaugural meeting of the Society for the Promotion of Engineering Education (SPEE, now ASEE) in 1893, motivations had shifted and Mansfield Merriman (1893, p. 259) argued the importance of literacy skills in doing engineering work, although he dismissed their formal inclusion in the curriculum. A decade later at the same meeting, however, a University of Nebraska professor presented a plan for "a combined cultural and technical engineering" curriculum, arguing that engineers need a broad liberal education as well as a deep technical one in order to be true professionals rather than "doers of other men's thought" (Chatburn, 1907).

Modern conceptions of engineering communication trace their roots to a seminal conversation at the 1911 SPEE meeting (Kynell, 1995), where Samuel Earle, an English professor at Tufts College, presented his “radical experiment” in teaching engineers technical writing (Earle, 1911). Until this point, when students studied writing it consisted mostly of grammar, mechanics, and spelling drills, and was decontextualized from disciplinary work (Kynell, 1995, p. 88). Earle described his integrated approach at Tufts, arguing that writing in context would facilitate learning and abstract thought, improve assessment of conceptual knowledge, and allow engineers to engage with audiences outside their profession. Earle’s ideas generated extensive discussion at the meeting – mostly positive – eventually leading to his appointment as the first chair of the SPEE Committee on English. Thus, the need to integrate communication into engineering courses persists more than 100 years after Earle first argued for it (Kynell, 1995, pp. 90–91; Raymond, 1911; Paretti et al., 2014).

While Earle was quite influential in solidifying support for writing instruction within the engineering education community, his work was in some ways undermined by its own success. Throughout the 20th century, an increasing number of subjects falling outside of the traditional liberal arts were added to the domain of higher education. Boundaries between disciplines became increasingly sharp as they tried to legitimate themselves by claiming ownership of regions of knowledge. During the rise of the early engineering English programs, many schools of engineering had English departments. Even though technical writing was in high demand, teachers were often derided from both sides: the English graduate programs that they came from perceived their status as beneath literary scholars, and engineering students criticized them for being insufficiently masculine. English graduates came to see interest in technical writing or composition as “professional suicide” (Connors, 1982, p. 337). By 1939, a survey showed that there were only five remaining English departments in schools of engineering, down from a peak of over two dozen (Connors, 1982, p. 338).

This lack of professional status combined with increased demand for technical writing outside the boundaries of engineering led to the development of technical communication as its own distinct discipline, separate from engineering or English. One effect of this change was that the integrated approach promoted by Earle was supplanted by the standalone technical writing course, which often does not provide the integrated knowledge and skills engineering graduates need (Ford, 2004; Paretti et al., 2014; Pinelli, Barclay, Keene, Kennedy, & Hecht, 1995; Wolfe, 2009).

2.1.2. A Contemporary Issue

This need for more integrated, targeted communication instruction is evident in the consistent concerns about engineers' ability to communicate and calls to action spanning the past 125 years. While the conversation about engineering communication curricula began in the 19th century and spawned an entire discipline, concerns about the communication skills of engineering graduates remain a central issue today.

In one of the first surveys of engineering employers in 1917, "the great majority of respondents felt that recently hired engineers did not have adequate English skills to perform their work" (Kynell, 1995, p. 91), concerns which persisted throughout the 20th century (Dudley, 1971). In the 1990s, surveys (Evans, Beakley, Crouch, & Yamaguchi, 1993; Pinelli et al., 1995) indicated that recent graduates and employers valued technical writing and communication skills second only to problem posing and solving, yet felt these were the skills most lacking in graduates. A review of industry perspectives summarized the state of the issue in 2001:

> the academic world has begun to understand that engineers learn much of their communication expertise through immersion in the workplace, but the workplace expects

> the academic world to fully prepare its engineering graduates in advance of professional practice (Donnell et al., 2011, p. 9).

Today, research continues to indicate that engineering program graduates and their employers still see engineering programs as insufficiently preparing students for professional communication.

The most recent large-scale effort to address this gap came with the advent of EC 2000, the accreditation criteria developed by ABET during the 1990s to help move engineering curricula from a focus on counting courses to a focus on outcomes (Williams, 2001, 2002). In particular, EC 2000 led to a major shift in the way engineering programs integrate communication because it required programs to demonstrate that their graduates possess "an ability to communicate effectively" (ABET, 2010; Williams, 2001, 2002). More recently, even as controversial proposed accreditation changes have been accused of "watering down" the so-called professional skills, the communication criterion has been updated to "[c]ommunicate effectively with a range of audiences through various media," arguably broadening and strengthening it (Flaherty, 2015). Still, engineering educators struggle with constraints that make it challenging to effectively teach these skills (Matusovich et al., 2012).

2.2. Learning and Communication

While the history of writing in engineering reveals the past and continued need to develop engineers' communication skills and to improve writing pedagogy in engineering, writing studies theories provide the frameworks that guide how to do so. Like engineering education, writing education and its theoretical foundations have developed significantly over time (see Nystrand, Greene, & Wiemelt, 1993). Two developments are particularly salient for this study: the pedagogical move from a product to a process-oriented approach to teaching writing, and the larger shift from viewing writing as an individual act of encoding meaning to a fundamentally social practice of constructing meaning. These developments are particularly important because they conflict with widely-held beliefs and practices within engineering, which are discussed in section 2.3. writing as an engineer also has implications for professional identity development (Lea & Street, 2006; Winsor, 1996) and, as a consequence, can introduce identity conflict when dominant ideologies of engineering conflict with students' existing identities (Sandler, Siiverberg, & Hall, 1996). In this section, I discuss writing pedagogy developments and the epistemological foundation of constructivism on which they are built as well as the implications for identity development.

2.2.1. Writing Pedagogy Should be Process-Oriented

The first key shift in writing pedagogy that informs this dissertation is the shift from product-centered to process-centered approaches that occurred in the late 1970s and early 80s. It may seem trivial to position writing as a process since in a sense every activity is a process, but the shift here is important because it represents a shift in pedagogies from attention on what students *make* to what students *do.* The product-centered pedagogies emphasize correctness, form, and style and frame instruction around the final text. From this perspective, Villanueva (2003, p. 18) describes teaching as having students

> look at texts, analyze and discuss what happens in those texts, and then produce something of their own that followed the patterns they found in those texts. Ideas were to be provided by the text, the form provided by the text, with evaluation based on how well the student paper emulated the ideal text. The process was rather like having students watch and discuss a videotape of a prima ballerina and having the students attempt the same dance, with the students then being evaluated based on how well they approximated

the ballerina's performance without knowing how the ballerina came to master those steps. No attention was given to the process of arriving at the product.

The process approach (*Figure 2*), on the other hand, does not emphasize emulation of idealized forms, but rather supports the student through the stages of prewriting, drafting, and revising (Murray, 1972).

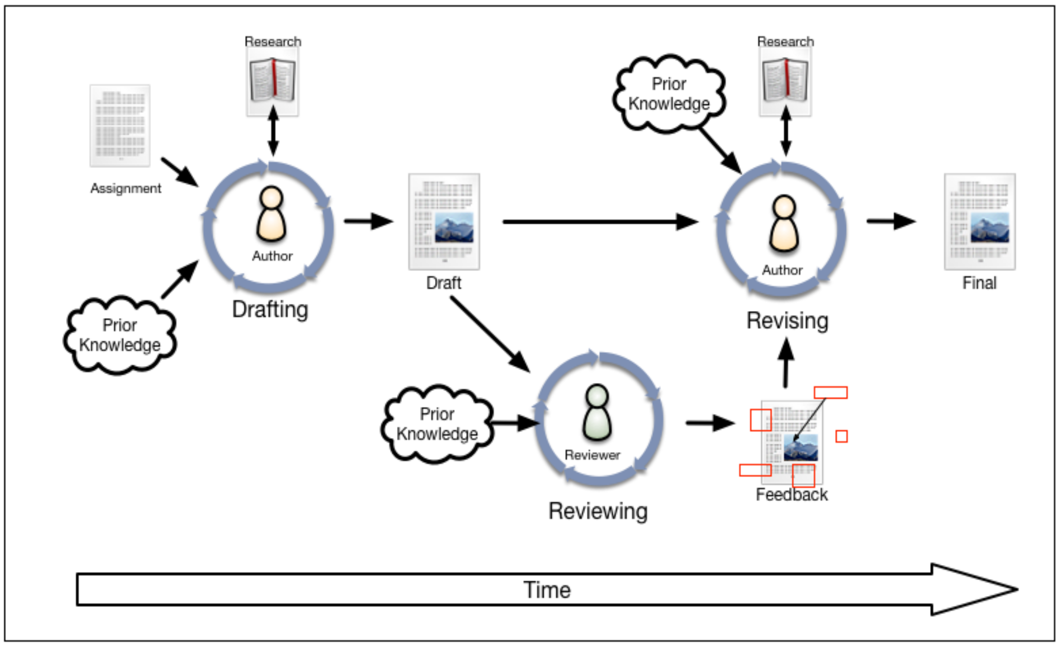


Figure 2: Writing as Process with Feedback and Revision

A key difference between the two approaches is in the role of assessment. In the product approach, assessment is summative: "students write essays and teachers describe their strengths and weaknesses, grading them accordingly [… assuming] that evaluating the product of composing is equivalent to intervening in the process" (Knoblauch & Brannon, 1981, p. 1). Process pedagogy, on the other hand, emphasizes formative assessment: teachers should give students feedback on drafts during writing, and students should have the opportunity to revise their writing based on that feedback (Garrison, 1974; Murray, 1972). Crucially, the feedback given for the purposes of formative assessment must also be appropriate to the maturity of the draft: sentence-level feedback is not only unnecessary in early stages but is ultimately confusing to the writer (Sommers, 1982). Instead, in early stages of drafting and revising, teachers provide guidance related to issues such as focus, organization, and argument, with stylistic and mechanical concerns addressed in later stages of editing and proofreading.

It should be noted that since process theory rose to prominence, writing studies scholars have also developed a set of theories collectively known as post-process that challenge some assumptions of process theory. Kalan (2015) synthesized the post-process literature and identified seven key aspects:

1. Writing cannot be reduced to a single codified process to be taught.
2. Essayist literacy and the rhetoric of assertion should be challenged in order to broaden genre possibilities.
3. Writing should liberate students' agencies.
4. Writing is not an individual activity taught through a simple classroom pedagogy.
5. Teachers need to move beyond the classroom as the only rhetorical situation and their role as the possessor of the techne of writing.
6. Written texts should be regarded as products of a complicated web of cultural practices, social interactions, power differentials, and discursive conventions.
7. Teaching writing is basically teaching rhetorical sensitivity and hermeneutic guessing through a large number of literate activities.

Despite these criticisms, post-process theories still generally view writing as a process. The most salient criticism to this study is the first: that "writing cannot be reduced to a single codified process to be taught." However, the intervention used in this research does not do this; rather, the intervention was situated within the context of the course and the findings, summarized in

Section 1.3 and detailed in Chapter 4, highlight important disciplinary nuances that in themselves contribute to this post-process understanding.

2.2.2. Writing is a Situated Activity

The second major development in writing pedagogy influencing this study moves beyond the cognitive to the social domain. The primary theoretical reason for the shift from product to process approaches described in section 2.2.1 lies in the cognitive domain in that it treats writing as an internal cognitive process (Flower & Hayes, 1981), with the role of the teacher focused on providing input to that process through responses (Keh, 1990). Many composition scholars viewed this approach as too limiting (Berlin, Faigley) and, building on Vygotsky's social-historical theories, brought about the "social turn" in composition (McComiskey, 2015). The social turn has numerous consequences, for the purposes of this study the most important being that writing is – and therefore writing pedagogy should be – inherently *situated*.

The socio-historical movement in psychology had broad influence and the social turn in composition was contemporaneous with many related movements across disciplines that Greeno (2005, p. 79) collectively calls situativity, or the situative perspective. This perspective is more than simply a change of reference frame: the situated perspective entails an epistemic shift, rejecting the premise that knowledge is fundamentally contained within individuals and instead adopting a social epistemology in which knowledge is "distributed among people and their environment, including the objects, artifacts, tools, books, and communities they are a part of" (J. G. Greeno, Collins, & Resnick, 1996, p. 20). This social epistemology is what differentiates the situative perspective from those that are simply socially-informed, such as Linda Flowers' updated social cognitive theory of writing, which is still considered cognitivist (McComiskey, 2000, p. 49). A key distinction is that when knowledge is considered to be situated in its environment, knowing and doing are inextricably linked such that attempts to separate knowledge from its context and teach it independently are at best inefficient and perhaps counterproductive (Brown et al., 1989). Under this situative perspective, learning is often viewed as the result of legitimate peripheral participation leading to meaningful participation (Lave & Wenger, 1991) within a community of practice (Wenger, 1998).

Many theories used in writing studies are informed by the situative perspective, including discourse communities, rhetorical genre theory and activity theory (Paretti et al., 2014). Shared among these theories is a rejection of the transmission model of communication; instead they adopt a functional model that views texts in terms of their communicative purpose (Nystrand et al., 1993). Taken as a whole, the major implication of the situative perspective for this study is therefore that learning to write as an engineer cannot happen solely as an isolated activity devoid of disciplinary context. Instead, communication assignments should be situated within engineering courses and teach the use of authentic genres in order to help students enter the discourse community of engineering and thus develop a discursive identity (Allie et al., 2009; Artemeva et al., 1999).

2.2.3. Writing and Identity Development

The situative perspective also helps refocus the role of communication in engineering work. Many of the calls to improve writing education in engineering curricula (see section 2.1) implicitly conceptualize writing solely as a *professional skill* – an individual cognitive skill that simply requires improved learning and replication of particular textual features. In contrast, the situative perspective suggests an alternative – viewing writing as a means through which students develop a particular *academic literacy* – and, correspondingly, an important means through which students develop a professional identity (Lea & Street, 2006; Winsor, 1996). From this perspective, learning to communicate as an engineer is not a separate process from learning to be an engineer: to learn engineering is “a process of coming to participate in the discourse of the engineering community and taking on the identity of being a member of this community” (Allie et al., 2009, p. 2). This interwoven relationship between fluency in the discourse of engineering and coming to identify and be identified as an engineer – the development of a *discursive identity* – reinforce the notion that “core” engineering courses should integrate communication tasks (Allie et al., 2009; Bielenberg, 2011).

It is also critical to recognize that we hold multiple identities and that in developing a professional identity we do not discard a previously-held identity and acquire a new one, but rather must integrate these multiple dimensions of identity into a new sense of self (Jones & McEwen, 2000). This aspect of identity development is particularly salient when students experience the dominant ideologies of engineering as conflicting with their existing identities. This conflict contributes to what has been called the “chilly climate” (Sandler et al., 1996) implicated in the persistent underrepresentation of women and minorities in engineering (Marra, Rodgers, Shen, & Bogue, 2012). Attempts to organize the curriculum to better support discursive identity development should ensure that “students are given the opportunity to take on discursive identities that embody these concerns” (Allie et al., 2009, p. 4).

2.2.4. Constructivism and Process-Oriented Pedagogy

The shift from product to process-oriented writing pedagogy, as well as the social turn in writing, are both grounded in constructivism, an epistemology that sees knowledge as created through the interaction of ideas, or prior knowledge, and experiences. In constructivist theory, experiences become prior knowledge, but prior knowledge also influences how we perceive experiences. There are several branches of constructivism, all of which are grounded in the work of Swiss psychologist Jean Piaget, who developed the concepts of schemata, perturbation, assimilation, and accommodation to explain how people learn (Hofer & Pintrich, 1997, p. 101).

Schemata are our mental models of the world we live in; that is, schemata “hold" our prior knowledge. It is important to note again that schemata are models. For example, we have probably seen many thousands of chairs throughout life, but in our knowledge of chairs, we don’t remember every detail of every variation of chair that we’ve seen. Instead we each have a schema for *chair* that is a generalized mental model of what that word means (Piaget, 1971).

New experiences that don’t quite match our schema are called *perturbations;* when a perturbation occurs, we either assimilate or accommodate the new experience. *Assimilation* is the process by which we interpret new experiences in a way that conforms to our extant schemata, thereby leaving the schemata unchanged. Individuals see a new object that fits their schema for chair and they interpret it as such. *Accommodation*, on the other hand, is the process by which we form a new schema to resolve a state of *disequilibrium,* which occurs when the mind cannot assimilate the new experience into existing schemata and therefore creates a new schema to resolve the disequilibrium. This entire process is called *equilibration* and is how constructivism explains learning (Rowell, 1989).

These concepts are the foundation of formative assessment and help explain the criticality of treating writing as a process rather than a product. In order to learn – to create new or modify existing schema – the presence of perturbations is a necessary but insufficient condition. First, the perturbation must cause a disequilibrium, and second, the learner must resolve it through equilibration. Summative feedback does not encourage equilibration because the student can easily ignore, dismiss, or misunderstand the response, leaving their schema unchanged or worse, and neither teacher nor student are aware of it. One way that the writing process approach supports learning, then, is by ensuring that feedback actually causes a state of disequilibrium and that the student resolves that state during revision.

While Piaget's theories aim to describe the internal processes of learners, they do not directly address teaching or the relationship between teacher and learner or the social aspects of learning that are also critical to studying formative feedback in writing. For that, I turn to Lev Vygotsky, who built on Piaget's work and examined its implications for teaching, developing *social* constructivist theory. Two key components of social constructivism are the Zone of Proximal Development (shown in Figure 3) and Situated Learning.

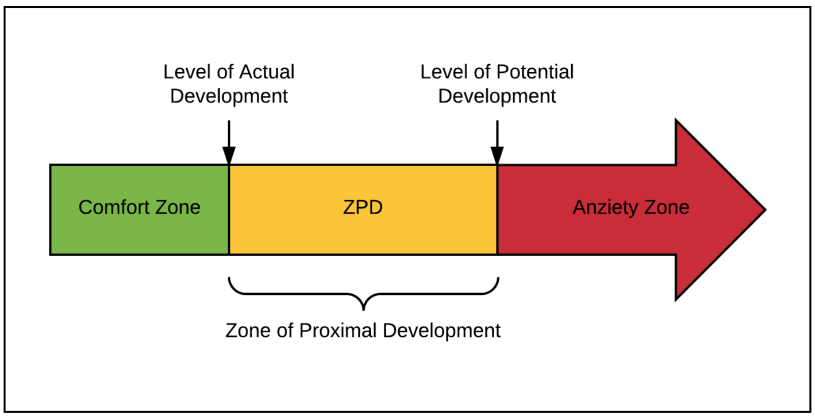


Figure 3: Zone of Proximal Development

In a domain like communication that consists of concepts and skills with increasing complexity, a learner has a subset of the domain in which they can act independently (the comfort zone in Figure 3), the upper bound of which is their level of actual development. At the far end of the domain lie knowledge and skills that are beyond the learner's capacity and which can produce anxiety, the lower bound of which is their level of potential development. Between these bounds is their Zone of Proximal Development (ZPD), the zone in which the learner can accomplish tasks with appropriate guidance (Bransford, Brown, & Cocking, 2000, p. 81).

Teaching, in the social constructivist sense, should aim to identify the learner's ZPD and work within it, designing activities that scaffold the learner's development. As the learner's independent competence increases, the ZPD moves further along the continuum (Brown et al., 1989, p. 456).

In part, this concept supports the idea of providing different types of feedback at different stages of writing. The use of formative assessment in process-oriented writing pedagogy is one such type of scaffolding. Feedback helps the writer identify and understand ways that their text does and does not accomplish their communicative goals and revision allows them to practice improvement. In order for this process to facilitate communicative skill development, it is crucial that feedback be appropriate to their Zone of Proximal Development. As discussed later in the next section, the ZPD is an important factor when understanding different types of feedback.

2.2.5. Constructivism and Peer Review

There are several relevant types of feedback from the perspective of constructivist learning theory, and the ZPD specifically. In her study of peer tutoring Chi (1996) differentiated between various forms of feedback including corrective, prompting, and scaffolding. Corrective feedback is when the tutor simply tells the tutee what to do or what is incorrect. When a peer tutor prompts—by asking a question to elicit elaboration—the result is self-construction of knowledge by the tutee when they answer. Scaffolding, on the other hand, implies a dialogical relationship between tutor and tutee resulting in co-construction of knowledge between the tutee and tutor.

These differences have implications from a constructivist perspective. Prompting and scaffolding correspond to Piagetian and Vygotskian theories respectively. With prompting feedback, the tutor essentially provides the perturbation in the Piagetian model by questioning the learner, which in turn prompts the (internal) learning through self-construction. When scaffolding, on the other hand, the tutor purposefully works within the learner's zone of proximal development to increase learning. Perhaps the most important distinction is that scaffolding requires the tutor to be aware of the "starting point of the tutee" (Topping & Ehly, 1998, p. 17) and purposefully guide the tutee by working above their current level (but within the ZPD). This means that scaffolding is usually inappropriate for "true peer" assessment – true peers are at the same level of understanding and aren't in a position to scaffold. Instructors, on the other hand, sometimes have "expert blindness" that makes them unable to recognize the starting point of the (novice) student enough to work within their ZPD (Bransford et al., 2000). As discussed in sections 2.4 and 3.2, design of writing assignments using peer review should be informed by these nuances.

2.3. Writing in the Engineering Curriculum

While theoretically-informed writing pedagogy is integral to practice in composition and technical writing courses housed in English and writing departments, it has had far less influence in engineering classrooms despite continued emphasis on communications in engineering. The outcomes-based ABET accreditation criteria introduced in 2000 formally included communication as one of the 11 areas that engineering programs must demonstrate for accreditation; however, outcomes-based assessment does not mandate how these criteria are operationalized in the curriculum, leaving programs wide latitude on implementation. Several implementation strategies that have been used include standalone technical communication classes, writing-intensive disciplinary classes, and multi-year integrative curricula (Paretti et al., 2014).

It is common for engineering faculty to value communication, yet see teaching it as outside their domain: it is *important* but *separate* from content instruction (Matusovich et al., 2012). This view reflects engineering faculty's epistemic beliefs, lack of pedagogical knowledge, and self-efficacy concerns within the domain of communication. These issues, combined with the pragmatic barriers of time and resource constraints, result in engineering programs operationalizing written communication in ways that are often at odds with the best practices of contemporary writing pedagogy described in the previous section – specifically, best practices indicating that writing instruction should be both situated and process-oriented.

Perhaps the most fundamental reason for the separation of communication from content in engineering curricula is epistemic: engineering faculty view them as distinct types of knowledge. Claude Shannon, an electrical engineer at Bell Labs who is often regarded as the founder of information theory, wrote in "A Mathematical Theory of Communication" that:

> The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently the messages have meaning […]. These semantic aspects of communication are irrelevant to the engineering problem. (Shannon, 1948)

Engineering faculty often view writing similarly, as a "simple transcription of experience into prose" (Broadhead, 1999, p. 24), with the work of communication being to more perfectly encode and decode knowledge into and from texts and reduce noise (Nystrand et al., 1993, p. 278); in other words, engineers often adopt what is called the "transmission model" of communication (Paretti et al., 2014, p. 614; Leydens, 2008).

Engineering faculty see themselves as subject matter experts and their role as teachers of *engineering knowledge* and conclude that some other, more qualified, subject matter expert will teach *communication knowledge*. A study by Zhu (Zhu, 2004) comparing engineering and business faculty views on the importance of writing found that engineering faculty consider written communication a critical professional skill; however, they also see writing competency as distinct from technical knowledge and do not emphasize its integration in disciplinary courses. In contrast, business faculty did not make this distinction, believing that writing skills *are* business skills. One professor of Business Information Systems, which has significant content overlap with computer science, stated that the emphasis on writing ‘‘is what I think differentiates us from the College of Engineering and Computer Sciences” (Zhu, 2004, p. 36). In other words, the separation of content and communication in engineering curricula is partially the result of engineering faculty having an epistemology of communication that supports that separation.

Similarly, in a study investigating faculty beliefs about teaching communication (Matusovich et al., 2012) over half of the engineering faculty interviewed indicated that “someone else” had responsibility for teaching communication. The “someone else” that faculty identify was split between elsewhere in engineering – usually a capstone design class – and outside of engineering – usually a technical communication class.

Perhaps this should not be surprising since a defining characteristic of engineering work is discretizing a problem space, defining boundaries around subdomains, and dividing labor among specialists in various fields. In fact, the construction of a “technical” vs. “non-technical” dualism – and the implied “real engineering” vs. “not-really engineering” dichotomy – is prevalent in engineering education and practice (Schneider & Munakata-Marr, 2013, p. 156), and has been theorized to broadly affect the profession in ways ranging from creating a chilly climate that encourages a lack of diversity (Cech & Waidzunas, 2011) to limiting engineers’ ability to think critically and examine the societal implications of engineering work (Cech, 2013; Leydens, 2013).

2.4. Peer Review and Response

One promising approach that has the potential to facilitate integration of process-oriented writing pedagogies into disciplinary engineering courses is the use of peer review. Engineering faculty value written communication, but the aforementioned barriers – lack of knowledge of effective writing pedagogy, time and resource constraints, and low self-efficacy beliefs – mean that writing assignments are left out of disciplinary courses altogether or implemented in ways that suboptimally support student learning. While faculty training can address lack of knowledge, peer review is a promising candidate for addressing time and resource constraints while preserving both the situated nature of learning to write and the need for formative feedback in the writing process.

Feedback on writing is a widely practiced part of writing pedagogy whose importance was stressed during the rise of the process movement; it was a common practice long before the empirical studies described above attempted to systematically understand it in the 1990s. Peer feedback was advocated by scholars as early as the 1970s (Elbow, 1973) and its value was taken as a "given" by composition teachers by the 1990s (Weaver, 1995, p. 32). Despite peer review's broad support by composition theorists and ubiquity among practitioners, for many years there was a lack of evidence-based research demonstrating its efficacy (DiPardo & Freedman, 1988; Newkirk, 1984; Weaver, 1995).

Addressing this lack of evidence, researchers began to empirically investigate peer review. Early work starting in the late 1980s focused on student and teacher perceptions of the value of peer review, and efficacy as measured by draft-to-draft revisions and grade changes. In the mid 2000s, scholars extended the detailed study of instructor feedback to include peer feedback. This work includes characterizing the types of feedback peer reviewers give (e.g. K. Cho, Schunn, & Charney, 2006); comparing peer, writing instructor, and content instructor comments (e.g. K. Cho, Schunn, & Wilson, 2006; Paulus, 1999); and investigating the effect of review on the reviewer's own writing (e.g. Jensen & Fischer, 2005; Patchan & Schunn, 2015). The results of these studies have generally supported the use of peer feedback and even identified ways it can be superior to instructor feedback. Although this work is promising, a recent meta-analysis of peer assessment studies (Topping, 2010) called for further empirical investigation of peer feedback using experimental and quasi-experimental study designs with large sample sizes in order to produce generalizable results (p. 343).

2.4.1. Student Perceptions of Peer Review

In early empirical peer review research, scholars sought to understand students' perceptions of peer review. For example, a large national study of successful secondary school writing teachers and their students found that students valued teacher-written comments on drafts the highest of any type of feedback, and teacher response of any kind – conference, comments on final papers, and numerical grades – over peer response of any type; especially notable is the direct conflict with teachers, who valued peer response higher (Freedman, Greenleaf, & Sperling, 1987). Another study, on the other hand, found that graduate students in three psychology courses rated peer review as highly beneficial (Haaga, 1993). One possible explanation for this difference is students' conception of the purpose of feedback: if they see the purpose of feedback as to improve their grade, then feedback from a peer who is not assigning a grade may not be valued (Weaver, 1995). In light of these results, it is particularly interesting that when Cho, Schunn and Charney (2006) asked students to rate the helpfulness of feedback comments without knowing whether the comment source was a peer or instructor, peer and instructor comments were perceived as equally helpful. Moreover, regardless of perception, peer feedback has consistently been shown to increase draft-to-draft quality (Topping & Ehly, 1998).

To further explore the difference between student and instructor perceptions of peer review, researchers conducted a follow-up study with a large and diverse sample – 708 students from 16 courses (12 undergraduate, 4 graduate) at 4 universities over three years (K. Cho, Schunn, & Wilson, 2006). Each paper was evaluated by at least four peers and one expert. They compared the numerical ratings (grades) that peers and instructors assigned to papers to determine reliability and validity from both the instructor and student perspectives. Their findings showed that an asymmetry of information may explain the difference between student and instructor perspectives. Because instructors have access to the entire set of peer ratings, they are aware when the mean student ratings are close to the instructor's ratings (i.e. student ratings have high validity from the instructor view). Students, on the other hand, only see the ratings that they personally receive, and this limited information leads to low perceived reliability and validity from the students' view. This result helps to explain why students rate peer review more poorly than would be expected given its efficacy (2006, p. 898).

2.4.2. Scaffolding Peer Feedback

The mismatch between student perceptions of peer review and the reality demonstrated by empirical studies foregrounds the necessity for scaffolding in the peer review process. Although peer feedback is generally effective in improving writing performance, several scaffolds, including training and blind review, can strengthen the effect. Students who receive feedback from peers who have been trained in effective review techniques show more improvement than those who receive feedback from untrained peers (Ozogul & Sullivan, 2007;

Sluijsmans et al., 2002; Weaver, 1995). The use of modelling, question and response forms, and computer-mediated feedback also have a positive effect (Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Gielen, Tops, Dochy, Onghena, & Smeets, 2010).

One consideration for training is that correction is often dismissed as ineffective in both the peer review and composition literature and it is often suggested that reviewers should be trained to avoid it. This type of feedback tends to cause the tutee to blindly accept the correction, rather than building deep understanding (Chi, 1996, p. 4). On the other hand, the classroom context and presence of other types of feedback may mitigate against blind acceptance, so avoiding it entirely may be too broad a recommendation (Patton & Smith Taylor, 2013). These factors suggest that when teaching students to give peer feedback in engineering classes, instructors should focus on training students to:

1. give primarily prompting-type feedback, and
2. use corrective feedback judiciously and never exclusively.

Additionally, instructors should also discuss the *why* of assignment constraints. For example, corrective peer comments such as “need to cite source,” “use the headings from the template,” or “include index terms” would be more likely to promote learning rather than blind correction when the learner understands the reasons those conventions exist.

Finally, the use of blind peer review can increase performance, which is especially salient because of the prevalence of in-person peer response groups in writing classes. Blind review improves the honesty of feedback for two reasons: it mitigates students’ discomfort with critiquing their peers (Haaga, 1993; Loretto, DeMartino, & Godley, 2016), and it also encourages them to act as an authentic, curious reader who can admit not understanding concepts to a “competitor” in the class (Weaver, 1995).

These components of scaffolding are key considerations in the overall design of the intervention described in Section 3.2. Training is also the primary experimental variable for the first research question, where I seek to understand whether training can effectively be implemented without the use of classroom time.

2.5. Analytic Framework: Features of Feedback Comments

Because peer review has been studied in a number of different domains outside engineering, I ground this study in existing frameworks to categorize different dimensions of feedback. A number of researchers have empirically examined types and features of feedback at the comment level. One of the more influential taxonomies, and the one that forms the primary frame for this work, was developed by Straub and Lunsford (1996b) by analyzing the comments that twelve respected writing scholars made on a set of sample essays. In the mid 2000s, researchers also began to examine peer feedback at the comment level using this and other taxonomies to (K. Cho, Schunn, & Charney, 2006; Nelson & Schunn, 2009). In this section, I discuss the various features of feedback that researchers have identified and their effects, using Straub and Lundsford (1996b) and the related work by Smith Taylor and Patton (2006) as primary frameworks and noting variations where applicable.

2.5.1. Mode: Reviewer Control with Directive vs Facilitative Response

The first dimension of Straub and Lunsford's taxonomy, mode (presented in Table 1), places comments on a continuum reflecting the level of control the commenter takes over the author's text. When using directive modes, the commenter takes the role of editor, whereas using facilitative modes puts the commenter in the role of coach or reader. Directive comments are usually considered product-oriented as opposed to facilitative, which are process-oriented

(Patton & Smith Taylor, 2013). Writing studies literature stresses that facilitative comments are more desirable than directive comments because directive comments appropriate student texts and lead to primarily surface-level revisions that are simple implementations of the directive in the comment (Brannon & Knoblauch, 1982; Knoblauch & Brannon, 1981; Sommers, 1982; Straub & Lunsford, 1996a).

Other researchers have called into question some of the assumptions made in this earlier work. O'Neill and Fife (1999) argue that contextual factors influence how students read and interpret comments. For example, while Straub and Lunsford made little distinction between direct edits and directive comments, students whose previous teachers' feedback was primarily edits experienced even directive comments as coming from a caring and thorough reader and therefore more dialogical (p. 44). They also found that students read comments through their perception of the teacher's ethos: when students were asked about the tone of teacher comments, their responses indicated that tone of the classroom environment impacted their perception of the tone of written comments. These findings suggest that a simple reading of the text of a comment is insufficient to determine how the student perceives the level of control exerted, and that researchers must use caution when interpreting written comments out of context.

Similarly, Patton and Smith Taylor (2013) call into question the broad presumption that directive feedback is inappropriate and to be avoided. First, they argue that comments that are directive on the surface can still *implicitly* invite choice, because directive comments such as "explain why this is important," can "position the student to make important rhetorical judgments" when revising (n.p.). Second, they argue that disciplinary factors in the sciences and engineering mean that writing

> typically functions in more tightly constrained genre systems than does writing in the humanities. Not only are some directive comments necessary to foreground the factual and situational constraints, but also directive comments may elicit more rhetorical thinking than humanists assume (n.p.).

Because of these factors, Patton and Smith Taylor suggest that a mix of facilitative and directive comments may be most appropriate for engineering writing and argue that the open-ended nature of well-constructed facilitative comments can effectively “ripple through an entire project,” and that therefore the absolute number of facilitative comments may underestimate their effect.

Table 1: Comparison of Several Categorizations of Mode

Straub and Lunsford			Smith Taylor and Patton	
Category Name	Definition	Examples	Definition	Category Name
Corrections	The teacher makes a change in the text	"When I have tried my best, ~~It should have made me feel~~ as If I have achieved something." [adds above the crossouts: "When I have tried my best, I should feel as if . . .. "]	Edit	Edit
Negative Evaluations	Presented as an objective criticism about the writing	Poor sentence structure. This paragraph needs detail and clarity.		
Qualified Negative Evaluations	Qualified or draws attention to its subjective nature	I need to know more about specific techniques. This is a bit too technical for me.		
Praise	Positive judgment about the writing; presented objectively or acknowledges subjectivity of the responder	This is an interesting way to start the paper. I think you've got a good start here. Yes!	Authoritative comments are evaluative and usually directive. They may be positive or negative, but most often give the student advice about how to improve an aspect of the report. They may simply offer an evaluation, such as 'good job.'	Authoritative
Imperatives	Requests a change in the text or some action by the student, usually by means of a command	Put the conclusion in a separate paragraph Be careful with your transitions.		
Advice	Recommends or suggests a change in the text, offering advice that leaves a measure of choice to the student. Often conditional mood.	You might add some detail here.		
Indirect Requests	Uses a question that begins with "Can you" or "Could you" to prompt the student to engage in some activity of revision.	Can you give an example? Could you develop this argument?		
Problem-Posing Questions	Uses a question to identify a problem in the text or some issue that needs to be considered. Problem-posing comments do not directly call on the student to make changes.	Can't this be said of fall in any northern city? Is this paragraph really necessary? Do you see this material about your membership in the gang as your key point?	Coaching comments prompt the student authors to think further about a topic. These comments are often stated as questions.	Coaching
Heuristic Questions	The teacher asks the student to add to or think further about the content of the writing.	What beach? How old were you? In what ways has your illness changed your outlook and behavior?		
Reflective Statements	This category is a catch-all for descriptive, interpretive, explanatory, reader-response, and hortatorical statements-and all other statements that are not evaluative, directive, or advisory. They usually present the teacher's reflections on the writing, either as an instructor or as an individual reader.	[I'd consider adding an example.] It would solidify your point. In academic writing, the trick Is to express your opinion with authority-but to make sure your argument is more than just your opinion This is the kind of personal involvement you can use throughout the essay. Here is your second argument I actually stopped to think if that might be why Auburn has the name it does. I hadn't thought of people hibernating. It took me a while to figure how this sentence follows from the previous one.	Readerly comments express the professor's experience as a reader of the report. They may indicate confusion or appreciation, but they speak directly about the process of reading and understanding.	Readerly

2.5.2. Tone: Comment Affect

In Straub and Lunsford's model, the mode categories are also divided to reflect a comment's positive and negative language, but only for evaluative comments. For example, their evaluative modes include praise, negative evaluations and positive evaluations but advice and reflective comments do not make this distinction. I argue, however, that this aspect reflects a different construct than reviewer control, and in fact while those categories are present in their analysis of writing experts' comments (e.g., p. 183), their detailed definitions of mode exclude them (pp. 191-193). Smith Taylor and Patton (2006) also do not include this construct in their definition of mode (see Table 1); they do however separately categorize whether a comment is "positive" or not. As a result, I treat this affective issue as a separate dimension, tone.

Nelson and Schunn (2009) hypothesized from a cognitive psychology perspective that while other features of comments affect *understanding* of feedback, the tone, or affective language, affects *agreement* with feedback and likelihood of students making revisions. Two studies conducted using this model had differing results: the first found that comment affect had no effect (Nelson & Schunn, 2009) on implementation, while the second (Patchan, Schunn, & Correnti, 2016) found a small positive impact on implementation for praise (i.e. compliments) and a moderate negative effect from mitigating praise (i.e. compliments used to offset a criticism).

Traditionally, teachers' response styles are overwhelmingly negative and critical, primarily concerned with correcting students' mistakes (Christensen, 1962). Writing scholars, in contrast, stress the importance of positive comments and praise as positive reinforcement (Connors & Lunsford, 1993; Daiker, 1989; Gee, 1972), and have also found that positive comments build students' confidence as writers and their belief they can improve (Dragga, 1988).

2.5.3. Focus: Comment Level

Straub and Lunsford also classify comments according to focus, which "identifies *what* a comment refers to in the writing," whereas mode identifies "*how* the comment is framed" (Straub & Lunsford, 1996a, p. 158). Several classifications of focus have been used in the literature (Table 2). While there are some differences between these classifications, they generally view focus as high-level (global, content, high prose, substance), low-level (local, form, low prose) or outside the writing (extra-textual, other). Novice writers have difficulty identifying and addressing higher order issues with their writing (Hayes et al., 1987), so expert writing teachers emphasize these higher order issues in their feedback comments (Straub & Lunsford, 1996a). When revising, comments about higher order issues lead to revisions that increase paper quality, although when these comments deal with writing issues like organization they are more likely to be ignored than when they deal with substance issues such as inaccurate content (Patchan et al., 2016).

Table 2: Comparison of Several Categorizations of Focus

Author	Categories	Subcategories
(Straub & Lunsford, 1996a)	Global	Ideas, Development, Global Structure
	Local	Local Structure, Wording, Correctness
	Extra-Textual	N/A
(Smith Taylor, 2007; 2003)	Content	Sufficiency, Validity, Coherence, Organization, Appropriateness
	Form	Mechanics, Style, Word Choice, Design
	Other (2003)	Meta, Comprehension
(Treglia, 2009)	N/A	Ideas, Grammar/Editing, Form, Generic, Process/Academic, Formatting
(Patchan et al., 2016)	Low Prose	N/A
	High Prose	N/A
	Substance	N/A

2.5.4. Localization

A final key dimension of comments found in the literature, albeit not one identified by Straub and Lundsford, is localization. Localization describes *where* a comment refers to in the draft; when a comment is localized, it "explicitly refers to the location of the issue" (Patchan et al., 2016, p. 1100). It is different from focus, although comments with low-level focus are more likely to be localized. Localization is a particularly salient comment feature when comparing *end comments* and *marginalia;* marginalia are localized by default (unless the commenter indicates otherwise, a comment refers to something in its vicinity), whereas end comments are non-localized by default. Guidelines for giving feedback often recommend end comments because of the higher-level focus, yet guidelines also recommend localization, which is lower in end comments (Nilson, 2003).

An extreme example of localization is proofreading marks; their position at the exact location of the change or suggestion leaves no ambiguity as to where the reviewer is referring to with their feedback. A margin comment such as “could use more references” could be localized due to its proximity to a paragraph or an arrow pointing at a particular area; in contrast, the same comment would be non-localized if it were an end comment. By comparison, an end comment noting that “the middle of the paper needs work” would be more localized, and “could you make third section clearer?” would be even more so.

Evidence on the effect of localization is mixed: localized comments are more likely to be understood (Nelson & Schunn, 2009) and more likely to be implemented (Nelson & Schunn, 2009; Patchan et al., 2016) than non-localized comments, but also slightly reduce revision quality (Patchan et al., 2016), as shown in Figure 4.

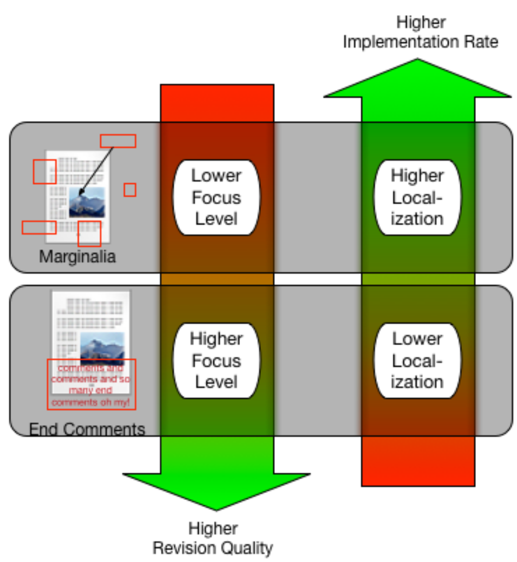


Figure 4: Interaction of Revision Quality and Implementation Rate with Margin and End Comments

2.6. Summary

In this chapter, I argued that while written communication is an important skill for engineers and is recognized by engineering faculty as such, curricular and pedagogical practices often suboptimally support student development of this skill. Key practices of effective writing pedagogy include contextualization and process support: writing assignments should be embedded within disciplinary courses and students should be given feedback on their writing and asked to revise it. Problematically, when engineering faculty embed writing assignments they often see their role as limited to simply requiring students to write in their classes – including assignments with written deliverables such as reports or documentation. Feedback is often limited and summative – a grade and/or comments on the written deliverable.

This state of practice conflicts with the consensus that formative assessment is a key to developing writing skills, that feedback should be given during the process of writing and inform subsequent revision. The most significant barrier to improving practice in engineering courses is resources: large student-faculty ratios make providing feedback on drafts difficult or impossible. To address this barrier, I propose that peer review could be used instead of instructor review. Prior research has shown that peer feedback can be as or more effective than expert feedback, but this has not been empirically tested in engineering, nor have researchers explored how to train students to give effective peer feedback. My study addresses these gaps, drawing on research in related domains to guide data collection and analysis.

Chapter 3: Methods

In the review of the literature, I argued that are there several factors constraining the improvement of writing pedagogy in engineering courses. Specifically, assignments do not approach the task of writing as an iterative, recursive process but rather employ a product-centric *one and done* approach. While this mismatch can be partially attributed to lack of knowledge and training on the part of engineering instructors, there are often structural barriers to providing the feedback and revision opportunities that are key to developing writing skills: large class sizes and time constraints prevent instructors from providing adequate feedback during the writing process. I hypothesize that these barriers can be addressed through the use of peer feedback in place of instructor feedback. Some research has shown that peer feedback can be as effective as or more effective than expert feedback (K. Cho & MacArthur, 2010; K. Cho, Schunn, & Charney, 2006; K. Cho, Schunn, & Wilson, 2006), but this has not been empirically studied in the context of engineering. I addressed this gap in this dissertation by answering research questions in support of the following overarching goal:

> **to better understand how peer review of writing can be used to effectively support development of engineering students' writing.**

To achieve this, I conducted a quasi-experimental content analytical study in which I analyzed the practice and efficacy of peer review of writing in a large first-year engineering course. This analysis used existing data from an intervention I developed with colleagues and deployed at a large mid-Atlantic research university in 2013. In the remainder of this chapter, I first discuss my research approach, including the quasi-experimental design and associated research sub-questions. Next, I provide a brief overview of the intervention from which the data to be analyzed resulted. Finally, I detail the methods and procedures I will use for analysis.

3.1. Research Approach

To address the overarching goal, I decomposed the process model of feedback and revision (Figure 5) and examined its two constituent processes—reviewing and revising—through separate research questions. I used a quasi-experimental design, manipulating inputs to each process in order to purposefully generate data on how these processes are affected by practices with a range of resource costs. The research questions are:

RQ1: How do different types of feedback instruction (in-class + handout, handout-only) affect the practice of peer review (i.e. the nature of reviewer feedback) for different peer review types (single peer, multi-peer)?

RQ2: What factors influence the efficacy of peer review (i.e. the nature of authorial revisions)?

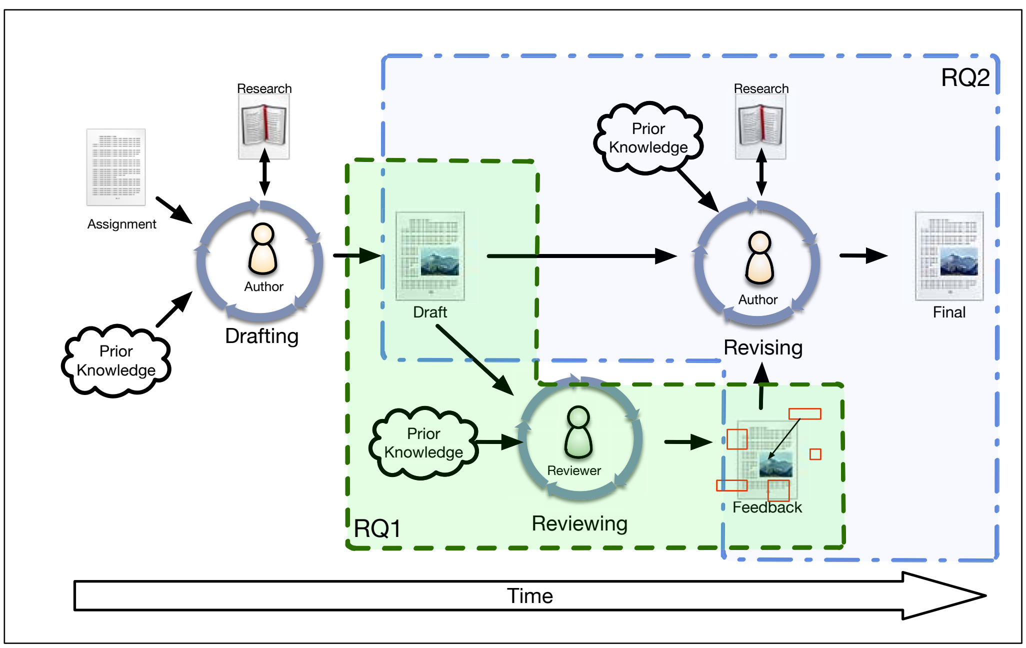


Figure 5: Research Question Location within Process Model

As shown in Figure 5, the research questions consider temporally-distinct aspects of the writing process model. The first research question considers how predictors that are fixed at the time of draft submission affect the types of feedback reviewers produce. The second question considers how predictors that are fixed at the start of the revision process affect revisions and quality improvement (i.e. the dependent variables for RQ1 are independent variables for RQ2).

The quasi-experimental design allowed me to manipulate inputs to the processes and analyze how those changes influence outcomes by applying treatment(s) to some groups while withholding them from others and comparing the outcomes (Creswell, 2009, p. 12). At the time of the study, first-year engineering students at the study site were enrolled in large multi-section courses, providing an ideal opportunity to apply variations of a treatment to entire sections and compare outcomes between groups. The large class size and high student-faculty ratio that constrain the direct application of pedagogical practices from composition courses afford a practical advantage here: the ability to test multiple variations of an intervention in parallel in order to empirically analyze their effects.

3.1.1. RQ1: The Reviewing Process

The first question concerns the process of providing peer feedback. As I discuss in the previous chapter, prior research suggests that some types of feedback are more effective in eliciting revisions that enhance paper quality than others, though some studies produce conflicting results. It makes sense, then, to train students in how to give feedback effectively, but recommendations on how to do so are largely absent in the extant literature. In conversations with several experienced writing teachers, I found that they usually teach peer review through modelling, examples, and supervised practice. Because a goal of this study is to determine how to introduce peer review while minimizing additional instructional resources, I introduced a higher and lower-resource form of peer feedback instruction as a variable in the quasi-experimental design, termed *Peer Review Instruction:*

1. In the first type, in-class + handout, a lecture and peer-review workshop were conducted in class, and a handout with instructions on giving quality peer feedback were provided;
2. In the second, handout-only, the handout with instructions for providing quality feedback was distributed via the course management system with no in-class instruction.

These two types of peer review instruction vary in resource-utilization both in personnel and classroom hours. To establish a baseline an additional group of students received feedback from their instructor; we considered a no-feedback baseline but decided it would put some students at a disadvantage and was not justifiable in terms of ethical research practice.

3.1.2. RQ2: The Revising Process

The second question concerns the process of revising the draft to produce the final paper. Prior research finding that peer feedback could be as or more effective than expert feedback (specifically when feedback was received from multiple peers) was in the context of an upper-level social science course. Because a goal of this study is to replicate those findings in a first-year engineering context, I include *Peer Review Type* as a second variable in the quasi-experimental design. *Peer Review Type* is either a single peer, or multiple (two) peers. Another group of students received feedback from an instructor, in order to compare the effect of peer and expert feedback.

3.2. Context and Intervention

This study used existing data collected from a course-based intervention that several colleagues and I designed and implemented (Ekoniak, Scanlon, & Mohammadi-Aragh, 2013). The site for this intervention a course at a large mid-Atlantic research university required for all students intending to major in Electrical and Computer Engineering (ECE) or Computer Science

(CS). This intervention modified an existing writing assignment, the Contemporary Issue Report, that had previously not included feedback and revision. In the remainder of this section, I describe the assignment, the motivation for developing the intervention, and the intervention itself.

3.2.1. The Assignment

The Contemporary Issue Report (CIR) was a short research paper in which students were to

1. investigate and describe a contemporary issue related to electrical and computer engineering or computer science;
2. discuss the issue in terms of problem, possible solutions, and trade-offs; and
3. discuss any ethical issues or societal impacts of the issue or solutions.

Students were provided with a handout describing the assignment, an example report written by the instructors, a Microsoft Word template for the report, and a copy of the rubric used to evaluate the assignment.

The CIR assignment was aligned with a course unit on cochlear implants. In this unit, students learned about the technical working of a cochlear implant device including its microphone, microcontroller, digital signal processor, and implanted electrodes during lecture and then implemented a MATLAB model of the cochlear implant ultimately allowing them to process audio and listen to an approximation of what a user of the implant would hear. About half the lecture period was devoted to guided discussion of ethical issues surrounding cochlear implants from the perspective of both the medical engineering and deaf communities. We intended this to have students think about ethical and societal concerns going beyond the compliance and whistleblowing dilemmas they studied in the prerequisite course and to facilitate a more robust understanding of the impacts of engineering work. This unit was used explicitly to frame the CIR assignment – the example report we provided to the students was on the topic of cochlear implants and drew from the issues discussed during lecture.

Like many engineering instructors (Zhu, 2004, p. 36), we assumed that simply including a writing assignment was sufficient to support our course objectives. Not only did we not include any writing pedagogy considerations in the initial assignment design, but it never occurred to us that we should. The original CIR was assigned in a *one and done* fashion: students submitted a final version of their report for grading without receiving any feedback on a draft and did not have the opportunity for revision. Feedback was provided by the instructors but was strictly summative in nature. This approach is consistent with Zhu's (Zhu, 2004) findings on integration of writing assignments in engineering classrooms. After the first implementation of the assignment, the instructors reported anecdotally that the reports were generally of "poor quality."

3.2.2. Motivation

While we were concerned about the quality of the CIR reports, at the same time, as is common among engineering instructors (Wheeler & McDonald, 2000), we were unsure of our ability to "teach writing." After review of the composition pedagogy literature and consultation with a writing instructor, it was clear that the design of the CIR assignment was inconsistent with contemporary practices and with our own constructivist epistemologies: by simply giving students a writing assignment and then grading the result, we had unintentionally used the *product-centered* approach to writing pedagogy rather than the *process-centered* approach that is widely considered more effective by composition scholars. As discussed in Chapter 2, a key difference between these approaches is the role of teacher response: a product-oriented response aims to evaluate the text – summative assessment, whereas a process-oriented response aims to inform further improvements – formative assessment.

As is typical in large classes, resource constraints made providing feedback challenging. In contrast to most composition courses, first year engineering courses tend to have large student-faculty ratios. At the time and site of the study, a full-time instructor in the first-year engineering program had about 150 students over 5 sections whereas a composition instructor had at most 84 students over 4 sections. In other words, engineering instructors with a full teaching load have nearly twice as many students as composition instructors while also lacking the knowledge and expertise in writing pedagogy. Inspired by research showing that peer feedback on writing assignments in disciplinary classes could be as or more effective than feedback from subject matter experts (K. Cho, Schunn, & Charney, 2006; K. Cho, Schunn, & Wilson, 2006; Y. H. Cho & Cho, 2010; K. Cho & MacArthur, 2010), we set out to modify the CIR assignment to incorporate peer review.

3.2.3. Intervention Details

In order to effectively include peer review in the CIR assignment, it was necessary to understand 1) the attributes of effective (peer) feedback, and 2) strategies to train students to give effective feedback - an issue that the Cho et al. series was silent on. Another concern in the literature on peer review is that students tend to perceive peer feedback as less helpful than instructor feedback, mainly because the instructor's role as the arbiter of grades inherently adds to the perceived value of their comments (Weaver, 1995). With these factors in mind, another engineering instructor and I, along with a writing instructor, conducted a literature review and identified the following characteristics of "high-quality" peer feedback: it should be specific, encouraging, and helpful. More specifically,

1. Specific feedback is not vague, explains what the problem was and where it occurred, and is grounded in the rubric criteria. For example: "I think you should switch the order of the third and fourth paragraphs" is more specific than "You need to work on your organization." This criterion is based on Sommers' (1982) work showing that vague "rubber stamp" comments are less useful in making revision decisions than specific comments.
2. Encouraging feedback doesn't attack the peer, uses positive phrasing, and identifies strengths as well as weaknesses. For example: "Your previous transitions between paragraphs have been very smooth. This one could be stronger." is more encouraging than saying "You lost me here." While negative feedback is problematic in general, it is especially detrimental to peer feedback. When peers have a negative perception of the comments they receive those comments are less likely to impact revisions than positively-perceived comments.
3. Helpful feedback was defined as following a provided list of guiding questions, such as "Can you identify the piece's main argument?" and "What might be missing from the paper? What else, as an interested reader, would you like to see in the paper? What else does the assignment ask for?" Olds (Sharp, Olds, Miller, & Dyrud, 1999, p. 54) advocates the use of guiding questions such as these.

These criteria were used to develop both the student handout and in-class workshop on effective peer review.

Handout

The handout, included in its entirety in Appendix A, included the following:

1. Why Peer Feedback: This section explains the purpose of the peer review process, as well as communicating a rationale for the use of peer review. The purpose of the rationale was to mitigate the common perception that peer feedback is not as valid as instructor feedback.
2. Quality Peer Feedback: This section outlines the criteria for quality feedback that we identified from the literature. It includes descriptions, examples, and guiding questions for students to refer to as they review peer reports.
3. Examples: Excerpts from prior student reports are included, along with instructor-developed examples of high-quality feedback on those papers; this served to model feedback techniques for reviewers.
4. Rubric: The high-level rubric used to evaluate the reports was provided, along with weights for the various evaluation criteria.

In-class Workshop

The in-class workshop consisted of about 30 minutes of instruction conducted by an experienced writing teacher. The first 10 minutes included a presentation of the material from the handout using identical or similar instructions, justifications, and examples. Students were then asked to pull up their CIR draft and the rubric on their computers and to switch computers with the person next to them to practice giving feedback. As students wrote feedback on each other's' drafts, the writing teacher walked around the class three times to answer questions. The most common questions regarded use of the PDF markup tools. Before concluding the workshop, the writing teacher ensured that each student had made at least one comment on his or her peer's draft.

Instructor Feedback

To establish a baseline against which to compare the efficacy of peer feedback, an additional experimental group received feedback from their instructors. These instructors were PhD students in electrical and computer engineering. Each instructor provided feedback to one 30-student section in this study; these two sections were combined into a single treatment group (instructor-only feedback) for analysis. The two instructors were not involved in designing the intervention. They were given copies of the handout described above but were not explicitly told to use it and received no additional instruction or guidance on providing feedback; whether did use the recommendations from the handout when providing feedback is not known.

3.2.4. Summary of Quasi-Experimental Conditions

As noted above, in addition to varying the ways in which students received instruction in peer review, we also varied the peer review conditions themselves. Some sections received feedback only from the instructor, some from a single peer, and some from multiple peers. Tables 1 and 2 describe the treatment groups and associated quasi-experimental conditions. These groups correspond to existing sections of the course; sections were assigned instructional treatments based on availability of the writing instructor to conduct the in-class workshops.

Table 3: Treatment Groups

	Feedback Instruction		
Feedback Source	Handout Only	In-Class + Handout	None
Instructor	-	-	B[1]
Single Peer	C	D	-
Multiple Peers	E	F	-

[1] Group B received single-expert feedback but did not give any peer feedback and therefore did not receive any peer feedback instruction

Table 4: Summary of Data Collected

	Feedback Conditions		Totals		
Group	Source	Instruction	Participants[1]	Reviews[2]	Complete Revisions[3]
B	Instructor	-	47	47	47
C	Single Peer	Handout	56	46	42
D		In-class + Handout	42	29	23
E	Multiple Peers	Handout	70	110	27
F		In-class + Handout	49	81	27

[1] Subset of enrolled students with IRB consent

[2] Number of reviewed CIRs in set with IRB consent from both reviewer and author.

[3] Number of revised CIRs in set with complete draft and revision, reviews from all assigned reviewers, and research consent from all; the drop in complete revisions in the multiple peers groups (E and F) is due to each revision requiring consent from three participants rather than one or two

Other potentially-relevant details of the intervention implementation include:

- Due to our uncertainty about whether IRB approval would be received in time, students were not aware they would receive feedback and have a chance to revise their drafts when they submitted them.
- The schedule of the writing instructor determined which sections were included in the in-class instruction condition – that is, the in-class instruction occurred during sections when the writing instructor was available.
- Peers were assigned based on alphabetical roster order: each student on the roster was assigned to review drafts from the next one (for single peer) or two (for multi-peer) students on the roster; students at the end of the roster reviewed those at the beginning.
- Review was blind – names were removed, and instructors distributed the drafts to the assigned reviewers and returned reviewed drafts to their original authors using the online course management system. While this choice was made to protect students' privacy, there is evidence that blind review improves the honesty of feedback by mitigating

students' discomfort with critiquing their peers (Haaga, 1993; Loretto et al., 2016), and eliminating anxieties about showing weakness to a "competitor" in the class (Weaver, 1995).

- Feedback was written on drafts with digital inking of PDF files using tablet PCs, simulating pen-on-paper comments while allowing all documents to be handled electronically; electronic inking was also the way that students were instructed to complete other homework assignments in the course, meaning that they were already familiar with the tools.

3.3. Analytical Approach

To answer the research questions, I compared the relationships among quasi-experimental conditions, the nature of reviewer feedback (i.e. its classification based on the analytic framework described in Sections 2.5 and 3.5.2), and the nature of authorial revisions (i.e. their classification based on the codes described in Section3.6). Due to the size of the data set it was possible to make these comparisons quantitatively using statistical analysis techniques. As a result, the study design is overall quantitative in nature because it uses a quasi-experimental design in which the goal is to statistically compare the effect of various treatments across several quasi-experimental groups (Creswell, 2009, pp. 15–16). However, the data—texts—are inherently qualitative in nature (Graneheim & Lundman, 2004; Srnka & Koeszegi, 2007). Likewise, the objects of the study—feedback on and revision of writing—involve the construction and communication of meaning by the participants. The impact on the proposed study is significant: the goal of making statistical comparisons across treatment groups requires quantitative data, but the data in the set are inherently qualitative in nature. Therefore, it is necessary to transform the data from qualitative to quantitative in a way that is compatible with quantitative analysis techniques, a process well-suited to a method called Qualitative Content Analysis (Schreier, 2012).

The data set consists of draft and final reports from each participant as well as feedback from one or more of the feedback sources listed in Table 3. These were transformed in two phases:

1. Extract and classify feedback items from the drafts
2. Score draft and final reports based on the assignment criteria

In the following sections, I first describe Qualitative Content Analysis (Krippendorff, 2012; Schreier, 2012) and how it is used to transform data in general terms. In the second part, I detail how I used it to analyze the feedback items. In the third part, I describe how the reports were holistically evaluated based on the assignment rubric. Finally, I discuss how the data were used together to answer the research questions.

3.4. Qualitative Content Analysis

Qualitative Content Analysis (QCA) is "a method for systematically describing the meaning of qualitative material […] by classifying material as instances of the categories of a coding frame" (Schreier, 2012, p. 1). Crucially, the method is designed to transform qualitative data such that descriptive and inferential statistics can be used on the quantized data (Borrego, Foster, & Froyd, 2014, pp. 62–63). There is some disagreement among methodologists regarding when Content Analysis is labeled "qualitative" because it is a family of research techniques that exist on a spectrum from quantitative to qualitative (Drisko & Maschi, 2015, p. 7). The content analysis procedure used in this dissertation is based on Schreier's (2012) work, so I use her terms.

Qualitative Content Analysis uses procedures and terminology similar to grounded theory and other qualitative techniques, but there are significant differences between the methods (J. Cho & Lee, 2014; Krippendorff, 2012; Miles & Huberman, 1994; Schreier, 2012; Strauss & Corbin, 1998):

- In many forms of qualitative analysis, coding is the starting point; the researcher engages immediately with the data, looking for patterns or concepts; codes emerge from the data. In QCA, codes are developed *a priori*, usually based on existing theory or concepts.
- QCA does not employ the constant comparative method; it is a linear process.
- The purpose of QCA is data reduction through categorization. In most conceptual coding, in contrast, the purpose is to develop and interrelate the codes into concepts and theory—data elaboration.
- Codes in QCA are mutually exclusive within a dimension, whereas in conceptual coding, codes can overlap.
- In QCA, data is segmented into codeable units before coding.

The first part of Schreier's QCA method is the development of the coding frame. The second part is the process of analyzing the data. For textual data, this process consists of three stages: Unitizing, Categorizing, and Coding. These stages are represented visually in Figure 6. They are executed sequentially but can be iterative within each stage.

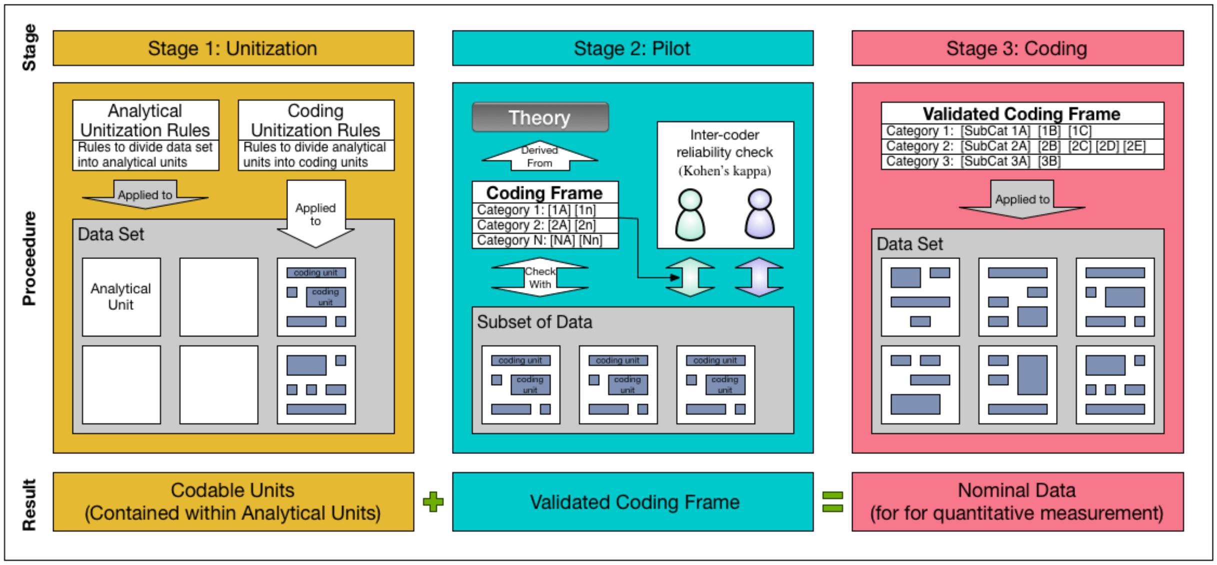


Figure 6: Visual Representation of QCA Process

Coding Frame

The coding frame is the lens through which the researcher peers at the data when conducting QCA. It is similar to a codebook, but is more highly structured and always hierarchical. The frame consists of one or more main categories (also called dimensions), each with two or more subcategories. By comparison to quantitative data, the dimensions are analogous to variables and the subcategories are analogous to levels of those variables. There are three requirements (Table 5) that apply to the structure of the coding frame: unidimensionality, mutual exclusivity, and exhaustiveness. Some indicators that unidimensionality has been violated include subcategories that are not instances of the dimensions, and subcategory names that repeat under different dimensions. (Schreier, 2012, pp. 72–78). The coding frames for this study are described in detail in Sections 3.5.2 and 3.6.1.

Table 5: Requirements for Coding Frames

Requirement	Definition
Unidimensionality	Each dimension should correspond to a distinct concept.
Mutual Exclusivity	Subcategories must not overlap.
Exhaustiveness	Each coding unit must fit into at least one subcategory in the coding frame.

Unitization

Unitization is the first stage of QCA once the data set has been compiled and the initial coding frame developed. The data set is first divided into analytical units, which are then divided into coding units. Analytical units are the highest-level units that will be compared after the QCA process is complete. Coding units are the chunks of the analytical units that will be assigned to subcategories from the coding frame. For example, if a researcher is studying how the main themes of articles in the New York Times changed over the course of a decade, the analytical units might be newspapers and the coding units articles; in interview research, interviews are usually the units of analysis and coding units might be individual turns (Schreier, 2012, pp. 130–131).

Units should be well-defined: unitization definitions and rules should be based on the nature of the data and the research question being answered. If units are based on natural divisions in the data set, formal definitions can be used; this is typical for units of analysis. Where distinctions are more ambiguous, a set of rules based on formal, thematic, or a combination of both criteria is necessary. The output of this stage is codable units. In this study, the coding units are feedback items and the analytical units are either feedback items, commented drafts, or received drafts depending on the question under analysis, as described in more detail in section 3.5.1.

Piloting and Coding

In the piloting stage, the initial coding frame is applied to a subset of data. The purpose is to ensure that the theoretically-derived coding frame makes sense with the data. The researcher first applies the coding frame to a small subset of the data to ensure that the categories are sufficiently well-defined. If there are dimensions or sub-categories that need to be added, removed, or modified, the coding frame is revised at this stage. After the coding frame has been revised, multiple coders apply the coding frame to subset of data, and an inter-coder reliability check is performed using a statistic such as Cohen's Kappa or percent agreement. The result of this stage is a validated coding frame. Finally, in the coding stage, the validated coding frame is applied to the rest of the data. The output of this stage is nominal or ordinal data that can be used for quantitative analysis.

3.5. Applying QCA to Feedback

In the first phase of data analysis, I applied the qualitative content analysis process described in the previous section to the corpus of feedback items. To assist in conducting the analysis, I used Atlas.ti 8.0 (Scientific Software Development GmbH, 2018), a computer-assisted qualitative data analysis (CAQDAS) software tool. I chose Atlas.ti over alternatives such as MaxQDA or Dedoose for two reasons that are important for this data set: robust PDF support and the software's data model.

First, the data consists of PDF files with primarily "digital ink" annotations, and Atlas.ti allows the free selection of "regions" of a PDF file and displays the selected region throughout the user interface. This makes it possible to select feedback items along with the relevant referent for both typed and handwritten feedback.

Second, the data model of Atlas.ti is especially well-suited to QCA. In most CAQDAS software, *codes* are applied to (text or region) *segments* to form *coded segments*, which are the fundamental data unit. The same text or region can have multiple codes applied, but treat these as separate coded segments internally. Atlas.ti, on the other hand, separates segment definition from code: *quotations* are created from text or region segments and codes are applied to the quotations. This can be seen in Figure 7, where the light blue boxes over the document are region quotations and may have zero or more codes applied to each quotation, which are visible in the sidebar. This means there are direct analogs between the Atlas.ti data model and the QCA process: *quotations* are analogous to QCA *coding units* and *codes* are analogous to *subcategories* of the dimensions of the coding frame.

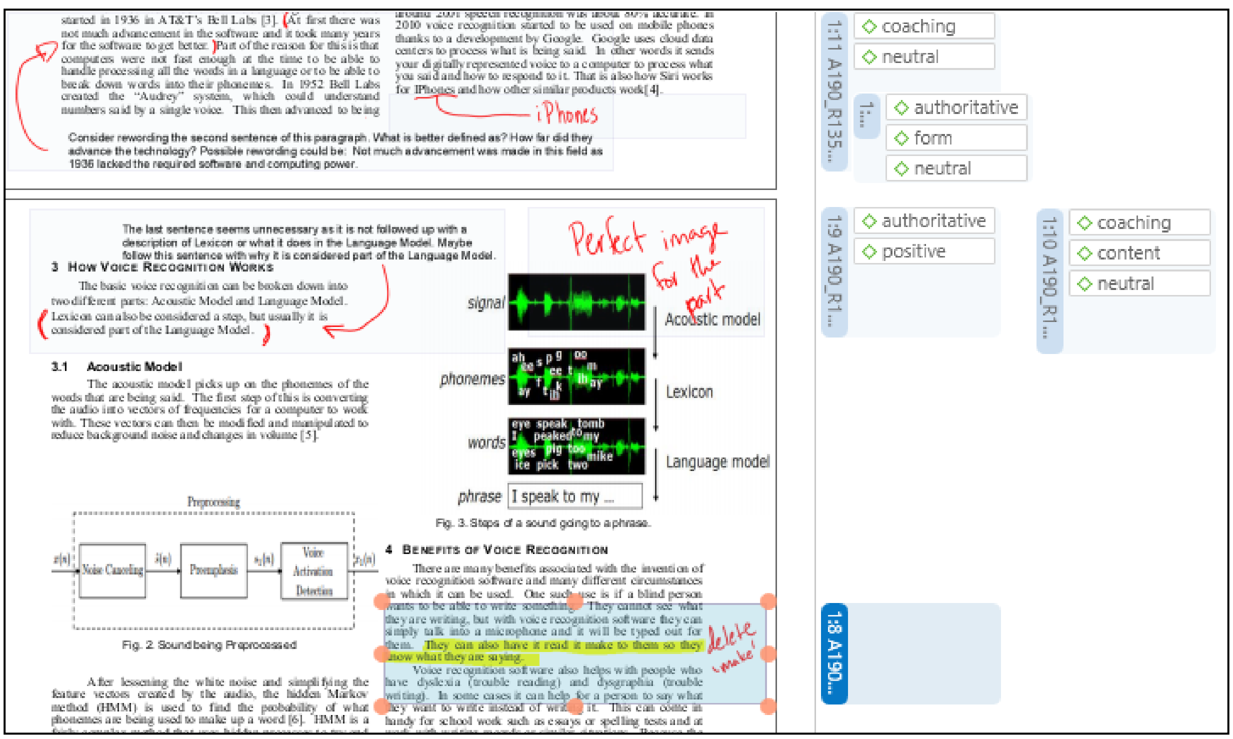


Figure 7: Quotations and Codes in Atlas.ti

3.5.1. Unitization of Feedback

The unit of coding for peer feedback is the *feedback item.* These feedback items were classified by applying the coding frame to them. In order to do this, it was necessary to define what constitutes an individual feedback item. Segmentation rules used a multi-pass procedure combining formal and thematic criteria (Schreier, 2012, pp. 134–137).

1. On the first pass, feedback items were segmented based on visual separation as seen in Figure 7.
2. Next, feedback items were subdivided if they contained more than one idea unit, a "self-contained message on a single problem" (K. Cho, Schunn, & Charney, 2006, p. 268).
3. Finally, because complete sentences are the smallest meaningful unit appropriate for the research paper genre, direct edits or proofreading marks that were within the same sentence were combined. As shown in Figure 8, four individual editing marks within a sentence were combined into a single feedback item.

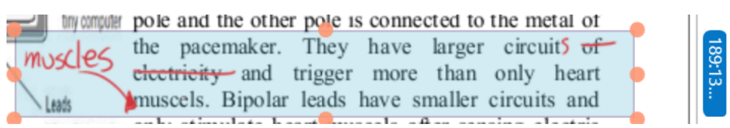


Figure 8: Combination of Within-sentence Edits

Several units of analysis were used: feedback items, commented drafts, and received drafts, which combine all feedback received by a particular author. The data are naturally divided into these units, so this simple definition is sufficient to define the units of analysis.

3.5.2. Coding Frame for Feedback Items

I developed the coding frame for feedback items based on prior empirical work classifying peer and expert feedback comments, as discussed in chapter 2, drawing specifically on Straub and Lunsford (1996a), Smith Taylor and Patton (2013), and Patchan, Schunn, and Correnti (2016). It is an extended version of the codebook I developed for my unpublished Masters of Engineering project (Ekoniak, 2014). The frame contains five dimensions: mode, tone, focus, localization, and type. Focus and localization are properties of the comment referent; mode and tone refer to the meaning-related qualities of the comment itself; type differentiates end comments vs margin comments. Edits are a unique case: they are a type of feedback and need to be captured in the coding frame, but they are not comments per se; therefore, edits are not assigned a focus, tone, or type and are inherently implementable and localized. Edits, therefore, distinguish *feedback items* from *feedback comments* (from here referred to simply as *comments*). In other words, a feedback item is either an edit or a feedback comment, and only comments are classified on the other dimensions of the coding frame.

Table 6: Overview of Coding Frame Dimensions

Dimension	Subcategories	Description
Mode	Edit, Authoritative, Coaching, Readerly	How comment wording enacts reviewer control over author's text
Tone	Negative, Neutral, Qualified, Positive	Affective language used in a comment
Focus	Form, Technical Content, Prose Content, Extra-textual	*What* a comment refers to (level of comment referent)
Localization	None, Implicit, Explicit	*Where* a comment refers to in the draft
Type	Margin, End	*Location* of comment itself

Mode

The mode dimension considers the form, content, and voice of the comment itself. Straub and Lunsford (1996b) developed the concept of mode to better capture differences in meaning, teacher role, and control that are not accounted for when considering only superficial form and voice of comments. Subcategories of mode include edit, authoritative, coaching, and readerly.

Table 7: Definitions of the Mode dimension[1]

Mode	Definition	Subject	Locus of Agency	Reviewer Role	Reviewer Control
Edit	Reviewer directly changes the text for the writer		Reviewer	Correcting	High
Authoritative	Reviewer makes an evaluation or *commands* the writer to make a change	First person; Second person (imperative)	Reviewer	Correcting Directing	High
Coaching	Reviewer *recommends* a change or asks the writer a question.	Second person; First person (conditional)	Writer	Guiding Prompting Questioning	Medium
Readerly	Reviewer takes position of audience, provides outside information, or summarizes without evaluating	First Person Third Person	Audience	Reflecting	Low

[1] adapted from Straub and Lunsford (1996b) and Smith Taylor (2007)

Mode reflects both the implied role of the reviewer and the degree of control the comment exerts over the text. This is illustrated by the following four comments about moving a paragraph to the introduction:

> "Put this in the introduction"
> "I think you should put this in the introduction"
> "You might consider putting this in the introduction"
> "Putting this in the introduction will allow the reader to more easily understand the rest of the paper."

The first comment directly commands the author to make a change; it exerts a high level of control over the writing and would be coded as authoritative. The second and third comments are both suggestions on the surface; they both appear to exert a medium level of control. However, the phrasing "I think" shifts the agency to the reviewer and in practice exerts a higher level of control while "you might" gives the writer agency; they would be coded as authoritative and coaching, respectively. The final comment takes the perspective of the audience and would be coded as readerly.

In some cases, more than one mode can be present in a single comment. In these instances, the comment was coded for the dominant mode. For example, in the comment

> Your report was very well composed, but you have many grammatical errors throughout. I would recommend reading the report aloud and when something doesn't sound quite right you should fix those issues

the middle section (a recommendation – "I would recommend reading the report aloud and when something doesn't sound quite right") gives some agency to the writer and could fit into the coaching subcategory; however, the initial evaluation (authoritative – "your report was very well composed") and final imperative (authoritative – "you should fix those issues") shift the overall mode of the comment to authoritative.

Tone

In Straub and Lunsford's (1996b) original conception of mode, there were more categories in the evaluative modes than in the definition presented above; for example: negative evaluation, qualified negative evaluation, and praise. As discussed in 2.5.2, I have separated this out into its own dimension because any comment could be phrased negatively or positively, and the effect of that phrasing may have some effect independent of mode. This separation of mode and tone is further supported by feedback models developed by cognitive psychologists which differentiate between a comment's promotion of *understanding* vs. *agreement* (Nelson & Schunn, 2009, p. 377), which correspond to the cognitive, or thinking, and affective, or feeling domains of constructivist learning theory, respectively (Lu & Law, 2012; Wadsworth, 1996). Table 7 summarizes the codes for the tone dimension used in this study.

Table 8: Definitions of Tone dimension

Tone	Definition	Example
Negative	Uses negative language	"The abstract is poorly written"
Neutral	Statements that don't contain explicitly affective language	"Talk about alternatives"
Qualified	Uses positive language to offset a criticism	"I like the topic, but I'm not sure how it relates to ECE"
Positive	Uses positive language	"Great Example!"

Focus

The subcategories of the focus dimension are also adapted from prior classification schemes (Smith Taylor, 2007; Straub & Lunsford, 1996a) and include form, content, and extra-textual. Form comments refer to features of the text at the word, sentence, between-sentence, or sub-paragraph level. Form comments are about word choice, mechanics, document design, or layout and other surface features of the text. Content comments refer to larger issues in the text such as organization, ideas, and development. The content category is further subdivided into prose content vs. technical content, which differentiates between comments that could refer to any paper and comments that are assignment-specific. If they reference specific portions of the text, it is at the sentence or paragraph level and higher. Extra-textual comments are those that do not refer to anything within the text.

Table 9: Definitions of the Focus dimension

Focus	Referent	Example
Form	Mechanics, Style, Word Choice, Document Design, Formatting	"Your text should be justified; use the template"
Content	Organization, Ideas,	
Technical	Development, Evidence, Flow	"What are the ethical implications?"
Prose		"Your examples don't support this"
Extra-Textual	Outside the document	"Reminds me of Lab 2"

Localization and Type

Localization indicates whether a comment "explicitly refers to the location of the issue" (Patchan et al., 2016, p. 1100). As discussed in the literature review, research has shown that localized comments are more likely to be understood and implemented, but are also weakly correlated with lower revision quality (Nelson & Schunn, 2009; Patchan et al., 2016). Many previous studies were restricted to end comments (K. Cho & MacArthur, 2010; K. Cho, Schunn, & Charney, 2006; Patchan et al., 2016; Smith, 1997), often due to the design of online feedback systems used. In this study, however, peer feedback comments were given using tablet PCs which allowed students to use digital inking tools that approximate the pen-on-paper feedback more commonly used by instructors. This affordance allowed for both types of comments, and students were not explicitly instructed to use either one, although the model feedback provided on the peer review handout included both.

Like end comments, margin comments can be explicitly localized in the comment text, but unlike end comments, margin comments can also be implicitly localized by proximity to the issue being commented on. On the other hand, margin comments are more likely to have a low-level (form) focus rather than a higher-level (content) focus. Because of these differences, it is possible that end and margin comments differ in implementation probability and revision quality.

To investigate any differences, comments in this study will be coded for both localization and type (margin or end), addressing a limitation of previous studies. Figure 9 presents the complete coding frame used in this study for feedback items.

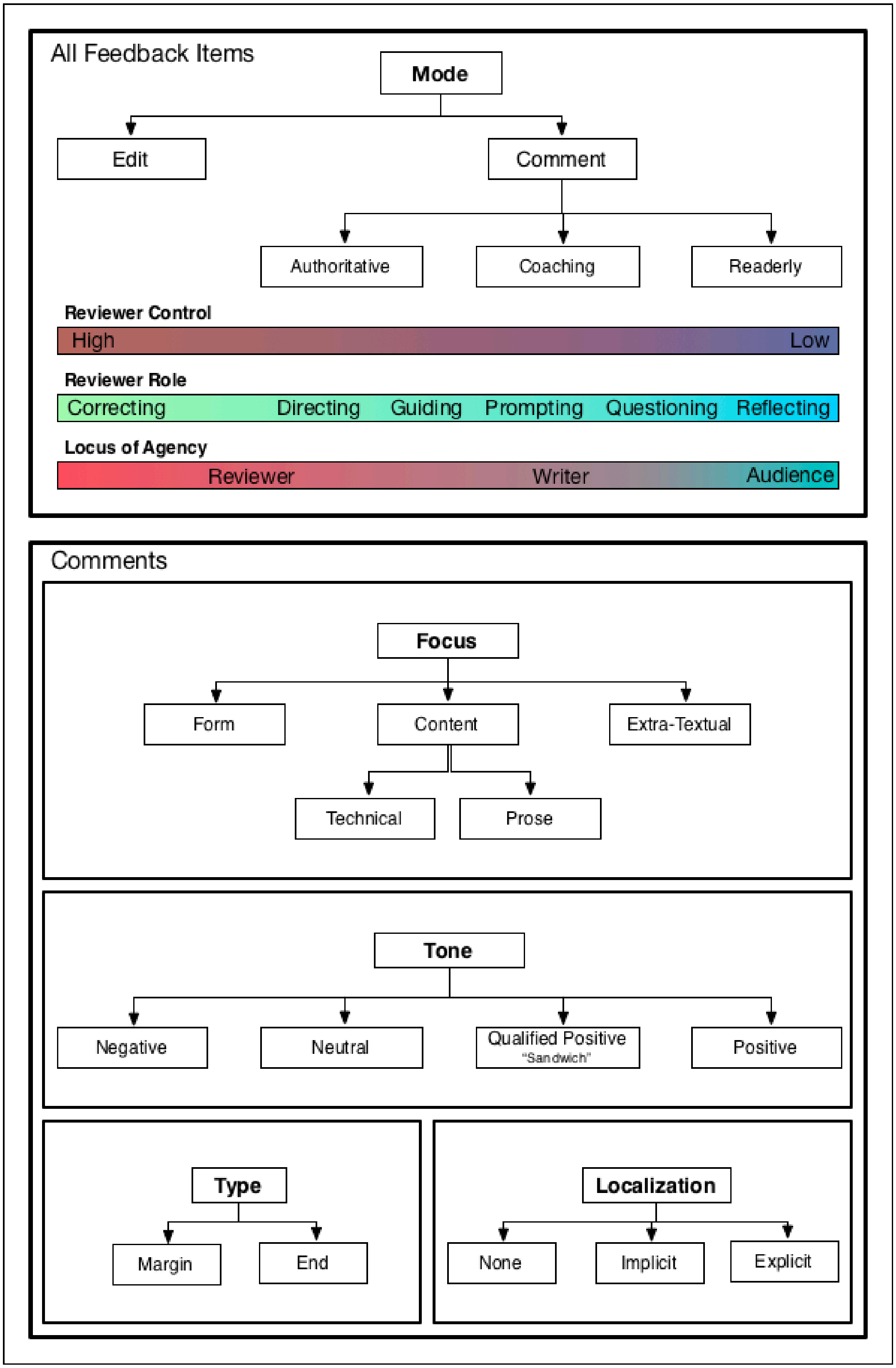


Figure 9: Visual Representation of Coding Frame for Feedback Items

3.6. Characterization of the Revision Process

After classification of feedback using qualitative content analysis, revisions were coded for both *revision quantity* and *revision quality*. The changes made between draft and final reports were scored using an instrument based on the assignment rubric to measure quality improvement along each dimension of the rubric. To assist in the identification of revisions, the draft and final reports in each unit of analysis were compared using a text-oriented *diff tool*, as seen in Figure 10. These tools are commonly used by programmers to compare software code and are conceptually similar to the "Compare Files" tools in Adobe Acrobat or Microsoft Word but eliminate much of the complexity associated with the PDF file format.

Figure 10: Sample comparison of draft and final reports using a diff tool

Direct analysis of the quantity and quality of revisions themselves enables analysis of perhaps the most salient question to improving educational practice: how and to what extent students improve – or don't improve – their drafts when revising.

Because students made their revisions within the context of this specific assignment, it was important that the rubric students used to guide their peer feedback was used to evaluate their revisions. The initial proposal for this dissertation called for each draft and final report to be evaluated independently using the exact same rubric that was used to grade the assignment, and the difference in score to be used as the measure of quality change. During pilot testing of this method, however, I determined that it lacked sufficient between-attribute discrimination and within-attribute resolution to fully capture changes in quality. To address this problem, I developed a new instrument derived from the original rubric dimensions, with several modifications:

1. Most significantly, rather than rating draft and final reports independently, the modified instrument is used to rate the difference between draft and final reports.
2. In the original assignment rubric, the "structure" attribute inappropriately combined two distinct constructs, organization and readability, which were separated for this instrument.
3. The "references" attribute, which considers general citing of sources, formatting of in-text citations and bibliography, and use of a specific citation style (IEEE or APA), was moved from the "writing quality" dimension to the "writing mechanics" dimension to clarify that the attribute is concerned only with the formal aspects of references as opposed to the "support" attribute which considers the use of sources to support the author's argument.

4. An overall assessment of the quantity and types of changes present, and an estimated letter grade were added.

In the following subsections, this instrument is described in detail. Screen captures of the full instrument, which used Google Forms, can be found in Appendix B; the assignment rubric on which it is based can be found in Appendix A: CIR Peer Feedback Handout.

3.6.1. Quantification of Changes

The first section of the instrument was used to identify the quantity of changes from draft to final. While looking at the diff tool output, each rater first estimated the overall extent of changes as either None, Minor, or Major. Next, each rater read through all of the individual changes to roughly – i.e. none, minimal, moderate, or extensive rather than an actual count – determine the number of changes falling into each of four categories, which were developed based on Faigley and Witte's (1981) taxonomy of revisions.

Faigley and Witte's taxonomy divides changes into either surface changes or text-base changes. Surface changes are further divided into either formal changes or meaning-preserving changes. Text-base revisions can either be changes to the microstructure or macrostructure of the text. Some surface changes, such as punctuation or spelling, are based on widely-accepted principles of correctness; others are based on genre-specific conventions.

Because it is at best controversial to claim that any change has zero effect on meaning, I refer to surface changes as corrections rather than revisions, a distinction made by several authors (Covill, 2010). For the purposes of this research, I distinguish between corrections that are grounded in general vs. assignment-specific criteria, such as the document template or citation style. Revisions are classified using Faigley and Witte's micro/macrostructure dichotomy. Microstructure revisions affect meaning primarily locally around the change, whereas macrostructure revisions are substantive enough to affect a summary of the paper. Figure 11 presents the coding frame used in this study for changes.

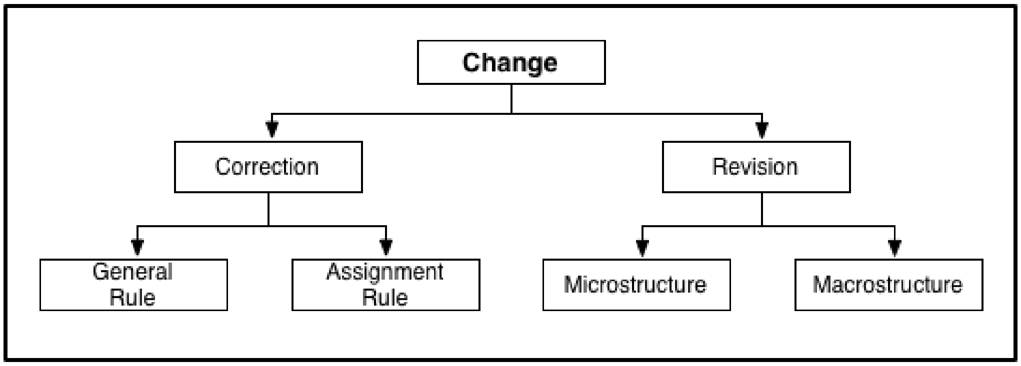


Figure 11: Categories of Changes

3.6.2. Quantification of Quality Improvement

The second section of the instrument was used to identify the impact of the changes on the quality of the report. The rubric was broken into three dimensions, with multiple attributes per dimension, as shown in Table 10. For each rubric attribute, the rater evaluated whether the changes present affected that attribute negatively, none or negligibly, positively, or very positively. Very positive is distinguished from positive in that a rating of very positive indicates that the changes present improved that rubric attribute by at least one rubric category; for example, a draft that would be rated "marginal" in organization and a final that would be rated "proficient" in organization was rated as "very positive." This relative rating system was developed after initial piloting of an absolute system. During piloting, I discovered that because multi-category improvements from draft to final were quite rare, the absolute rubric-based system lacked enough within-attribute resolution to capture improvements that were significant yet not substantial enough to change the category rating of that attribute. Finally, the rater assigned the draft report a letter grade, so analysis could control for initial quality, since a draft report with poorer-quality has more room for improvement when revising.

Once all reports were rated, an improvement score was computed for each dimension. The ordinal rating for each attribute within the dimension was transformed to a numeric score by assigning values to each category (Negative = -1, None/Negligible = 0, Positive = 1, Very Positive = 2). These scores and their ranges are summarized in Table 10.

Table 10: Quality Improvement Metrics

		Range	
Dimension / Attribute	Weight	Attribute	Score
Mechanics Quality			
Mechanics (Body)	1	{ -1 .. 2 }	
Mechanics (References)	1	{ -1 .. 2 }	
Mechanics Quality Improvement			{ -2 .. 4 }
Writing Quality			
Readability	1	{ -1 .. 2 }	
Organization	1	{ -1 .. 2 }	
Conceptual	1	{ -1 .. 2 }	
Support	2	{ -2 .. 4 }	
Writing Quality Improvement			{ -5 .. 10 }
Technical Quality			
Problem	3	{ -3 .. 6 }	
Link to ECE/CS	3	{ -3 .. 6 }	
Solutions	3	{ -3 .. 6 }	
Tradeoffs	3	{ -3 .. 6 }	
Ethics	2	{ -2 .. 4 }	
Societal Impact	2	{ -2 .. 4 }	
Technical Quality Improvement			{ -16 .. 32 }
Overall Quality Improvement			{ -23 .. 46 }

The improvement score for each dimension was computed by multiplying this value by the attribute weighting from the original assignment rubric (see Appendix A: CIR Peer Feedback

Handout) and summing the attributes within each dimension; an overall improvement score was computed by summing these scores:

$$Improvement_{Quality\ Dimension} = \sum_{\substack{for\ each \\ Attribute\ in \\ Dimension}} Weight_{Attribute} \times Rating_{Attribute}$$

For example, a paper rated as positive improvement in readability, very positive improvement in organization, negligible improvement in conceptual, and very positive improvement in support would have a writing quality improvement of 5, calculated as:

$$\begin{aligned} QualityImprovement_{Writing} \\ &= Weight_{Readability} \times Rating_{Readability} + Weight_{Organization} \times Rating_{Organization} \\ &+ Weight_{Conceptual} \times Rating_{Conceptual} + Weight_{Support} \times Rating_{Support} \\ &= 1 \times 1 + 1 \times 2 + 1 \times 0 + 2 \times 1 = \mathbf{5} \end{aligned}$$

3.7. Quantitative Analysis

Once all qualitative data were transformed into quantitative data the final step was to use statistical analysis to make comparisons relevant to the research questions. Figure 12 presents a graphical overview of the relationships among variables analyzed:

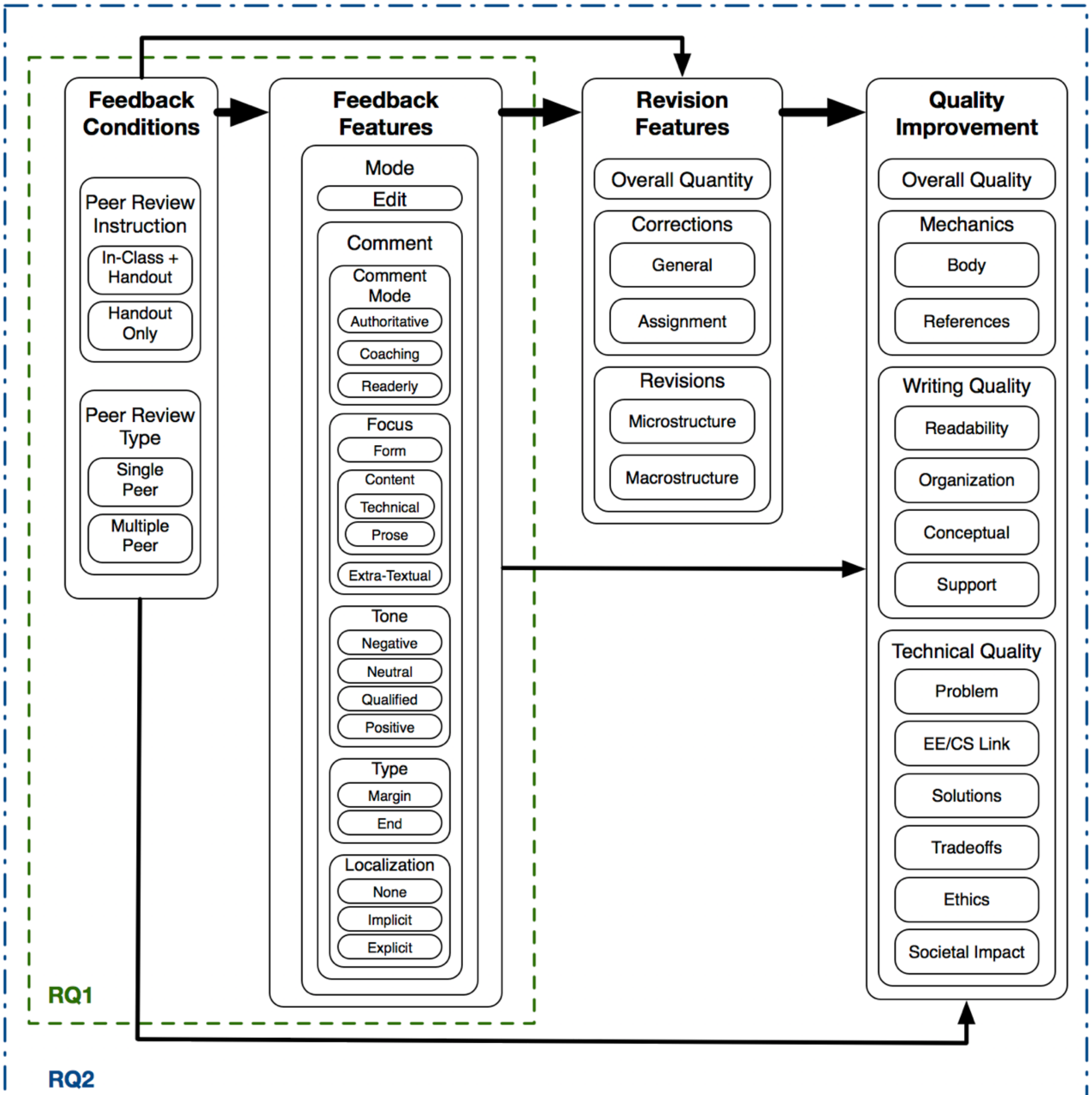


Figure 12: Conceptual Research Diagram – Variables and Research Question Boundaries

The first research question focuses on the *practice of peer review* – i.e. the giving feedback on a peer's report. The primary variable of interest for RQ1 is *peer review instruction*, but because reviewing multiple peer reports may also affect the kinds of feedback peers given, differences in *peer review type* – single and multiple review – are also analyzed. In other words, RQ1 is about the relationship between *feedback conditions* and *feedback features*. The second research question shifts the focus to the *efficacy of peer review* – i.e. revising based on received feedback. It is expected that *feedback features* will affect *revision features*, which in turn will determine *quality improvement* between draft and final reports; this is represented by the dark arrows in the diagram. There was also the possibility that lower-order variables will independently affect higher-order variables, so those effects – represented by the lighter arrows in the diagram – were also analyzed.

3.7.1. Levels of Measurement and Analysis

Data were coded at different *units of coding*: feedback features were coded at the feedback item level, as described in Section 3.5.1 whereas revision features were coded at the author level. Because of this, the *units of analysis* are at two different levels: feedback items were analyzed as their own analytical units and were also aggregated at the level of the commented draft. Furthermore, in the multiple peer conditions, authors received feedback on multiple (identical, but from different reviewers) drafts so for analysis involving revisions, commented drafts were aggregated by author. I refer to the two *levels* of analysis as the feedback item-level and the paper-level.

Additionally, which statistical techniques apply – and which questions can be asked and answered – depends on the *level of measurement* of the data. This refers to whether the variable is continuous (interval), nominal, ordinal, or dichotomous (Linneman, 2010, p. 17). For example, when I analyzed the effect of in-class peer feedback instruction, I wanted to determine if it made a difference in the focus of peer comments. First, I compared the focus of the comments from peers who received in-class instruction to those who did not at the *feedback item level of analysis* using a crosstabulation and chi-square test of significance to compare the two groups. By using these tools, I was asking the question "did in-class instruction have a statistically-significant effect on the distribution of comment focus?" My choice of statistical techniques was determined in part because comment focus has a nominal *level of measurement* (i.e. its possible values are an unordered list of categories).

While the chi-square statistic allows comparison of distribution of a nominal variable (e.g. focus), it does not indicate differences within particular sub-dimensions (e.g. form). To examine these differences, I conducted a second analysis at the *paper level of analysis* comparing number of comments per draft for each sub-dimension of the coding frame. At this level of analysis, the variables at the lower (feedback item) level are aggregated into a count (e.g. number of comments with form focus), which is no longer a nominal level of measurement but instead is continuous. To compare the in-class and handout instruction groups at this level, I wanted to compare averages: the question was now "did in-class instruction have a statistically-significant effect on the average number of form comments reviewers left on drafts?" Comparing means uses a different significance test (t-tests) than comparing distributions.

3.7.2. Selection of Statistical Techniques

In order to answer the research questions, I considered both the level of analysis and the level of measurement when choosing a statistical technique. Levels of measurement for each variable are reported in Table 11 for the feedback item level of analysis and in Table 12 for the paper level of analysis.

Table 11: Variables used in Feedback Item-Level Analysis

Variable	Measurement	Range
Feedback Conditions		
Peer Review Instruction	Nominal	{ In-Class+Handout, Handout-only }
Review Type	Nominal	{ Single Peer, Multiple Peer}
Feedback Features		
Feedback Mode	Dichotomous	(Edit, Comment)
Comment Features		
Comment Mode	Ordinal	{ Authoritative, Coaching, Readerly }
Focus	Nominal	{ Form, Content-Technical, Content-Prose, Readerly
Tone	Ordinal	(Negative, Neutral, Qualified, Positive)
Type	Dichotomous	(Margin, End)
Localization	Nominal	{ None, Implicit, Explicit }

Table 12: Variables used in Paper-Level Analysis

Variable	Measurement	Range
Feedback Conditions		
Peer Review Instruction	Nominal	{ In-Class+Handout, Handout-only }
Review Type	Nominal	{ Single Peer, Multiple Peer}
Feedback Features		
Total Feedback	Ratio	{ 0 .. *n* }
Mode		
Edits	Ratio	{ 0 .. *n* }
Comments	Ratio	{ 0 .. *n* }
Comment Features		
Comment Mode		
Authoritative	Ratio	{ 0 .. *n* }
Coaching	Ratio	{ 0 .. *n* }
Readerly	Ratio	{ 0 .. *n* }
Focus		
Form	Ratio	{ 0 .. *n* }
Content	Ratio	{ 0 .. *n* }
Technical	Ratio	{ 0 .. *n* }
Extra-Textual	Ratio	{ 0 .. *n* }
Tone		
Negative	Ratio	{ 0 .. *n* }
Neutral	Ratio	{ 0 .. *n* }
Qualified	Ratio	{ 0 .. *n* }
Positive	Ratio	{ 0 .. *n* }
Type		
Margin	Ratio	{ 0 .. *n* }
End	Ratio	{ 0 .. *n* }
Localization		
None	Ratio	{ 0 .. *n* }
Implicit	Ratio	{ 0 .. *n* }
Explicit	Ratio	{ 0 .. *n* }
Revision Features		
Overall Quantity	Ordinal	(None, Minor, Major)
Revision Type		
Correction		
General	Ordinal	(None, Minimal, Moderate, Extensive)
Assignment	Ordinal	(None, Minimal, Moderate, Extensive)
Revision		
Microstructure	Ordinal	(None, Minimal, Moderate, Extensive)
Macrostructure	Ordinal	(None, Minimal, Moderate, Extensive)
Quality Improvement		
Overall Quality Improvement	Interval	{ -23 .. 46 }
Mechanics Quality Improvement	Interval	{ -1 .. 4 }
Writing Quality Improvement	Interval	{ -5 .. 10 }
Technical Quality Improvement	Interval	{ -16 .. 32 }
Draft Rating	Ordinal	(A, B, C, D, F)

The research questions were answered by comparing groups – participants who received in-class peer feedback training to those who did not, for example. When the variable being compared between groups is categorical – nominal or ordinal level of measurement – crosstabulation and the chi-square test for significance are used. Crosstabulation is used to compare the distribution of categorical variables, while the chi-square test is used to determine the probability that differences in the distributions are attributable to chance (Linneman, 2010, p. 124). When the variable being compared is continuous, the group means are compared; the test used to determine the probability that differences in means are statistically significant across categories depends on whether two groups are being compared or three or more groups are being compared. When two groups are being compared, the t-test is used; for three or more, analysis of variance (ANOVA) is used (Linneman, 2010, pp. 193, 198).

In addition to determining whether differences in distributions or means between groups was statistically significant, I also used regression analysis when the level of measurement allowed. While the previously-discussed techniques only allow one to determine whether differences are statistically significant, regression allows one to determine the strength of relationships; in other words, regression adds explanatory power. To use regression, the both the independent and dependent variables must have meaningful order; there are regression techniques available for dichotomy, ordinal, and ratio levels of measurement.

The most common form of regression is ordinary least squares linear regression (Linneman, 2010). This type of regression fits a line that best represents the linear relationship between the independent and dependent variables by minimizing the distance between that line and the observed data. This type of regression can be used when the dependent variable is of the ratio level of measurement. The result of linear regression is the slope of that line, which represents the amount that the dependent variable changes for each unit change in the independent variable.

For non-ratio-level dependent variables, a different type of regression – logistic regression – is used (Linneman, 2010). Logistic regression is conceptually similar to a linear regression, but rather than fitting a model where the independent variable predicts the value of the dependent variable, it instead fits a model where the independent variable predicts the probability for each level of the dependent variable. A visual example is shown in Figure 13: for a particular number of feedback items, the distance between the x-axis and the lower blue line represents the fitted probability that an author receiving that number of feedback items will make no changes to their draft; similarly, the distance between the lower and upper lines represents the probability of an author making minor changes, and the distance between the upper line and the top of the chart is the probability that the author will make major changes.

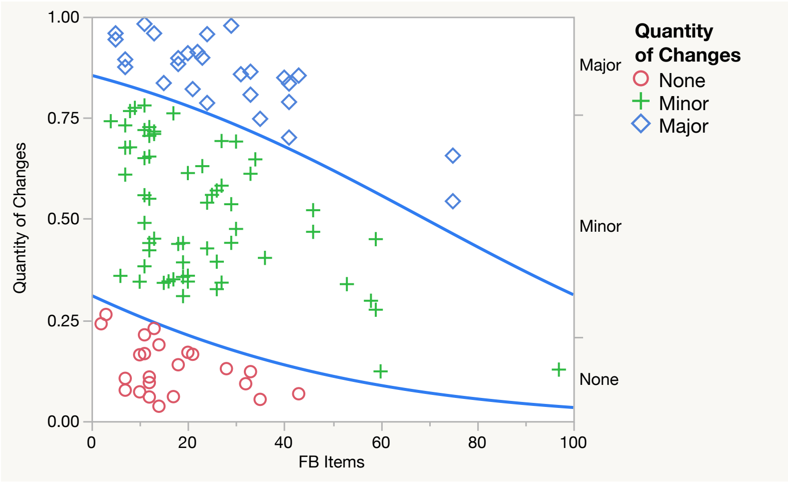


Figure 13: Logistic Fit of Quantity of Changes by Number of Feedback Items

For both linear and logistic regression, the result of analysis is a mathematical model – an equation that describes the relationship between the independent and dependent variables. For both types of regression, a coefficient of determination called R^2 can also be calculated. The coefficient of determination represents the amount of the total variation in the dependent variable is explained by the regression model and its values range from 0 meaning lack of correlation to 1 meaning perfect correlation. For example, an R^2 value of 0.42 means that 42% of the variation in the dependent variable is explained by the model. Because of the differences in linear and logistic regression, a the coefficient of determination for logistic regression is calculated differently and is referred to as the Nagelkerke R^2, which is analogous to and comparable with the linear regression R^2 – its values range from 0 meaning lack of correlation to 1 meaning perfect correlation, although perfect correlation is less likely with logistic regression (Nagelkerke, 1991).

When comparing multiple regression models for the same dependent variable, the coefficient of correlation must be adjusted for the number of independent variables because the addition of predictors artificially increases the model fit (Linneman, 2010, p. 257). When comparing models, I used this adjusted R^2 value as well as the Akaike Information Criterion (AICc), which is a fit index developed for model comparison that emphasizes model parsimony by additionally penalizing additional model variables (Akaike, 1974). Higher R^2 values indicate better model fit, whereas lower AICc indicate better model fit.

3.8. Trustworthiness

It is crucial in all research to establish the quality of the methods employed. For quantitative instruments, the accepted criteria are validity and reliability (Creswell, 2009). Because of the explicitly interpretative role of the researcher in qualitative studies, Lincoln and Guba (1985) developed the more robust concept of "trustworthiness" to ensure quality and to show that findings are "worth paying attention to." Trustworthiness is established by demonstrating credibility, dependability, confirmability, and transferability (Schwandt, Lincoln,

& Guba, 2007). Credibility, analogous to internal validity, refers to whether the research measures what it purports to measure. Dependability, analogous to reliability, refers to how consistent a measurement is. Confirmability roughly equates to refers to the scientific standard of objectivity and refers to "potential for congruence between two or more independent people about the data's accuracy, relevance, or meaning" (Elo et al., 2014, p. 2). Finally, transferability refers to the extent to which a study's findings can be extrapolated to other contexts. Although transferability of the results are limited by the single-site single-course nature of the study, the site and course were typical of first-year engineering programs at large research universities, supporting transferability to other contexts.

In this research, I used standards for each of these four criteria developed specifically for content analysis (Neuendorf, 2011; Schreier, 2012) to establish the trustworthiness of the results. First, credibility was ensured by using a codebook based on a preexisting framework that is well-established for use with similar data (Rourke & Anderson, 2004, p. 15). Dependability was established by having a subset of the data coded by two coders and using chance-corrected agreement coefficients (Neuendorf, 2011, p. 284), a process described in detail in the next paragraph. This process also helped to establish confirmability by using a second coder who was not familiar with the theoretical foundations of the study and therefore would base coding decisions more completely on the codebook rather than their expert interpretation. By using this *intercoder* reliability (where the goal is maximal agreement among coders based solely on the codebook and training) rather than *interrater* reliability (where raters are experts whose differing interpretations are valued and closely examined) the likelihood that these results could be reproduced independently is increased (Neuendorf, 2011, p. 283). The single-institution study site and course focused on ECE/CS students is a potential threat to transferability; however, the large size of the program (American Society for Engineering Education, 2017) and participant pool should help mitigate this threat.

Dependability, or reliability of measurement, is one of the more challenging aspects of qualitative research because the act of coding inherently involves some amount of subjective judgement on the part of the coder. This is somewhat simpler when using *a priori* (deductive) coding, but it is still necessary to have multiple coders independently analyze a subset of the data and determine the extent to which they agree. An appropriate *chance-corrected coefficient* of agreement or covariation should be used rather than simple percent agreement (Neuendorf, 2011, p. 283). For nominal data, the most widely-accepted coefficient is Cohen's Kappa, a measure of agreement; for ordinal, ratio, and interval data, Krippendorff's Alpha, a measure of covariation, should be used – although Kappa is also considered acceptable for ordinal data (Neuendorf, 2011, p. 284).

After defining the coding frame and unitizing the feedback items to be analyzed, I used a random number generator to select 200 (about 10% of the total) feedback comments to be coded by a research assistant. The minimum recommended threshold for Cohen's Kappa is 0.6, while some scholars recommend 0.8 (Neuendorf, 2011, p. 284); as shown in Table 13, the higher threshold was achieved for all feedback dimensions in the coding frame except mode (likely because mode category definitions are more subjective), which still met the lower threshold.

Table 13: Intercoder Reliability for Feedback

Dimension	Percent Agreement	Cohen's Kappa
Mode	85.3%	0.728
Tone	89.2%	0.800
Focus	89.2%	0.829
Localization	98.0%	0.954
Type	91.2%	0.865

For changes, 20 draft-final pairs were coded by a second research assistant. A random number generator was used to determine the sequence of analysis of the reports. This helped to ensure validity of the ratings because the rater was unaware of the treatment group the report was from. Krippendorff (2012) recommends a minimum Alpha threshold of 0.8, with values between 0.667 and 0.8 acceptable for “tentative” conclusions. As reported in

Table 14, intercoder reliability fell into this tentative range for most dimensions of the coding frame for changes. Two dimensions fell below the threshold, although only very slightly. The intercoder reliability achieved is in line with the previous work on which this analysis is based, which notes that intercoder reliability for writing is generally low (K. Cho & MacArthur, 2010, p. 333).

Table 14: Intercoder Reliability for Changes

Dimension (Level of Measurement)	Krippendorff's Alpha
Quantity of Changes (Ordinal)	
Overall	0.771
Corrections: General	0.693
Corrections: Assignment	0.664
Revisions: Microstructure	0.668
Revisions: Macrostructure	0.858
Change in Quality (Interval)	
Mechanics Quality	0.620
Writing Quality	0.793
Technical Quality	0.710
Letter Rating (Ordinal)	
Initial Letter Rating	0.745
Final Letter Rating	0.669

3.9. Limitations

Like any research design, this one has methodological limitations. Most significantly, the intervention and data collection were from a single class at a single-institution during one semester and it is possible that effects found in analysis are specific to this particular intervention in this particular assignment and are not generalizable to the larger population of first-year engineering students and writing assignments. The large population should support a claim of generalizability, but that claim is difficult to quantify without additional study. Additionally, there are factors that could affect a peer's feedback or revising practice that are not accounted for such as the peer's previous experience giving or receiving feedback, writing proficiency, etc. I assume that differences in these factors will be evenly distributed across various course sections and therefore treatment groups. Finally, the study site was a course primarily taken by students intending to major in ECE and CS. Although it was not limited to these majors there could be differences between students who select these versus other fields.

Chapter 4: Results

In the preceding chapters I presented a research study design intended to empirically test the impact of different approaches to peer review in order to answer the research questions. The utility of peer writing feedback in engineering classes is a particularly interesting question for several reasons. First, the improvement of engineering graduates' writing skills is a persistent concern of the engineering education community. Second, feedback and revision are among best practices of writing pedagogy. Third, structural barriers – large class sizes and time constraints – are common in engineering courses and constrain even pedagogically-informed instructors from providing adequate feedback. And finally, research in other disciplines has shown that peer feedback – which can address size and time barriers – can be as effective as instructor feedback.

Given this context, in this chapter, I present results of the qualitative content analysis procedures described in the previous chapter in order to answer the following research questions:

RQ1: How do different types of feedback instruction (in-class + handout, handout-only) affect the practice of peer review (i.e. the nature of reviewer feedback) for different peer review types (single peer, multi-peer)?

RQ2: What factors influence the efficacy of peer review (i.e. the nature of authorial revisions)?

The remainder of this chapter is organized as follows: First, Section 4.1 presents a comparison of quality improvement from students receiving peer feedback and instructor feedback in order to validate the assumption that multiple peer feedback can be used in place of instructor feedback. Section 4.2 addresses RQ1 and includes 1) a descriptive overview of the quantity and distribution of feedback items and the distribution and correlations between features of feedback on student drafts, and 2) the effects of peer review instruction – in-class plus handout vs. handout-only – on quantity and distribution of feedback items and the distribution and correlations between features of feedback on student drafts for both types of feedback (single and multi-peer). Section 4.3 addresses RQ2 and includes 1) the effect of feedback features on revision features, 2) the effect of revision features on quality improvement, and 3) any effects of lower-order variables on higher-order variables.

4.1. Quality Improvement Resulting from Instructor vs Peer Feedback

First, while a full analysis of instructor feedback practices is beyond the scope of this dissertation, a major assumption of this research is that, based on previous studies, it is valid to replace instructor feedback with peer feedback. Therefore, it is relevant to compare the quality improvement resulting from instructor feedback vs. peer feedback in order to empirically validate this assumption. As shown in Table 15, ANOVA analysis comparing the groups indicated that there were no statistically-significant differences in draft-final quality improvement between participants who received feedback from their instructor vs from peers. Follow-up pairwise comparison showed a statistically-significant improvement in writing quality for both multiple-peer feedback groups when compared with instructor feedback ($t(94) = 2.39$, $p = 0.0190$). These results indicate that replacing instructor feedback with multiple-peer feedback with is a reasonable choice, and may even be preferable.

Table 15: Comparison of Quality Improvement between Peer and Instructor Feedback

	Instructor		Single Peer				Multiple Peer				
			Handout Only		In-Class		Handout Only		In-Class		
	M	SD	M	SD	M	SD	M	SD	M	SD	sig
Overall Quality	4.00	4.50	3.85	5.06	4.76	4.65	6.27	6.48	4.72	4.36	0.3555
Mechanics Quality	0.69	0.73	0.54	0.68	0.71	0.78	0.85	0.92	0.72	0.94	0.6533
Technical Quality	2.38	3.72	2.21	4.31	2.95	3.71	3.46	4.27	2.40	3.75	0.7398
Writing Quality	0.93	1.36	1.10	1.39	1.10	1.30	1.96	2.47	1.60	1.44	0.0808

*p < 0.05 **p < 0.01

4.2. RQ1: The Effects of Instruction Type on Feedback

As described previously the first research question focuses on the effects of various conditions of peer review instruction on the practice of peer review. This section presents the results of analysis for RQ1 (the relationship between *feedback conditions* and *feedback features*).

4.2.1. The Nature of Peer Feedback

Before looking at how the treatment groups differ, it is useful to look at the big picture: what is the nature of peer feedback as a whole? In Chapter 3, I described a coding frame developed from prior theoretical work consisting of five dimensions (see Feedback Features in Figure 12) in which I classified feedback using the Qualitative Content Analysis (QCA) method. Feedback items were identified in each commented draft and unitized using the procedure described in Section 3.5.1. Unitized feedback items were then classified according to the coding frame for feedback items. Because feedback was analyzed strictly based on this classification, the coding frame can be considered the lens through which I understand the nature of peer feedback.

The first dimension of the coding frame, mode, initially separates feedback items into either *edits* or *comments*. The data set consists of 2,808 unitized peer feedback items, of which

28.3% are edits and 71.7% are comments. Figure 14 presents the distribution of comments versus edits.

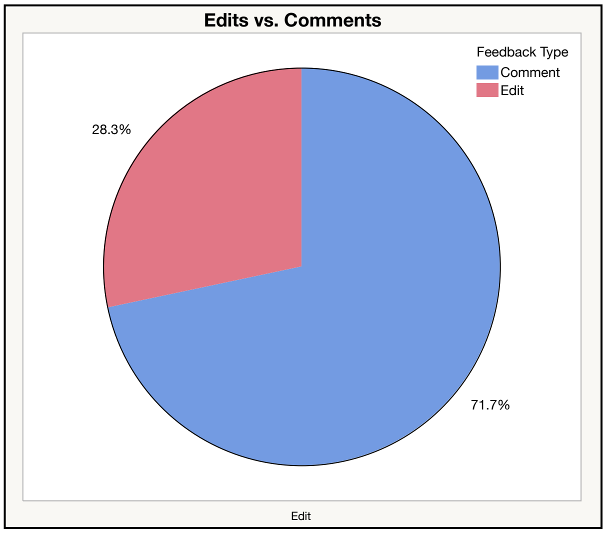


Figure 14: Feedback Items by Edit vs Comment for Instructors and Peers

Figure 15 presents a histogram of the coding results across all five dimensions. Note that because edits, by their nature, cannot differ in focus, tone, type, or localization – they are inherently localized form-focused marginalia – the focus, tone, type, and localization dimensions of the coding frame apply only to comments (i.e. feedback with a mode of authoritative, coaching, or readerly), not to edits.

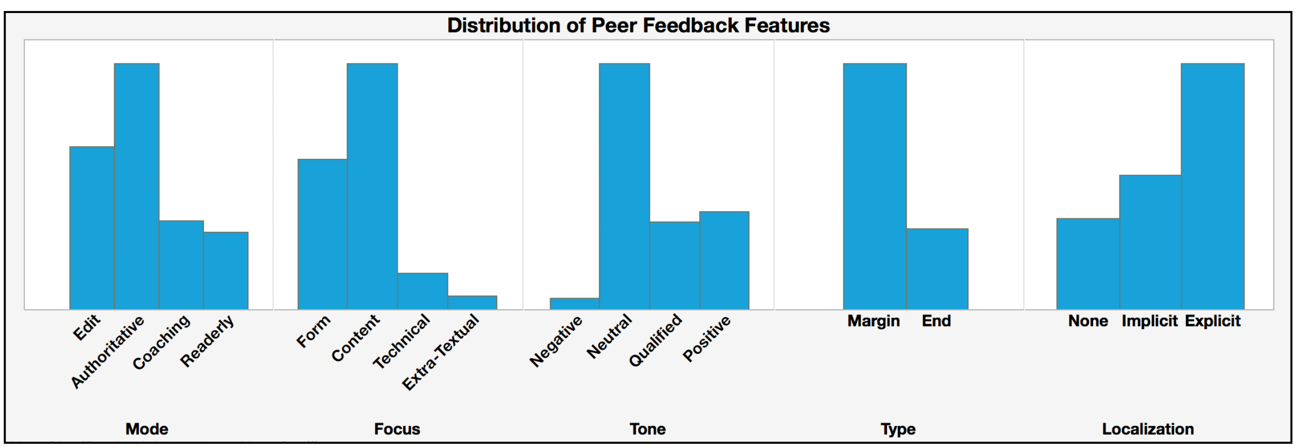


Figure 15: Histogram of Features for Peer Feedback Items
(*Focus, Tone, Type, and Localization distributions exclude Edits)

In addition to considering the distribution of comments across each of the dimensions, it is also useful to consider the correlations among dimensions, as shown in Table 16. For comments, while there are many significant correlations between subcategories of the dimensions of the coding frame, not all correlations are necessarily meaningful in the context of this study.

- Most of these correlations are weak (<0.3), meaning that it is reasonable to consider the dimensions independent when constructing predictive models.
- The interaction between Type and Localization is consistent with expectations – margin comments are more highly localized, either implicitly or explicitly, than end comments.
- End comments are more likely to be qualified (i.e. mitigating criticism with praise) with a higher-level focus than marginalia.
- The only statistically-significant relationship for negative comments is being localized, suggesting that highly critical comments are usually issue-specific rather than aimed generally at the author.

Table 16: Correlations Between Peer Comment Features (N=2013)

Dimension	1	2	3	4	5	6	7	8	9	10	11	12	13
Mode													
1 Authoritative													
2 Coaching													
3 Readerly													
Focus													
4 Form	0.205 ***	-0.048 †	-0.208 ***										
5 Content	-0.175 ***	0.054 †	0.163 ***										
6 Technical	-0.026	0.026	0.005										
7 Extra-Textual	-0.017	-0.068 *	0.093 ***										
Tone													
8 Negative	0.007	-0.030	0.022	-0.001	0.001	-0.032	0.054 †						
9 Neutral	-0.046 †	0.244 ***	-0.199 ***	0.317 ***	-0.231 ***	-0.102 ***	-0.045 †						
10 Qualified	-0.036	-0.027	0.074 **	-0.118 ***	0.059 *	0.095 ***	0.004						
11 Positive	0.087 ***	-0.255 ***	0.159 ***	-0.265 ***	0.219 ***	0.042	0.030						
Type													
12 Margin	-0.096 ***	0.124 ***	-0.010	0.107 ***	0.060 *	-0.218 ***	-0.130 ***	0.015	0.302 ***	-0.412 ***	0.027		
13 End	0.096 ***	-0.124 ***	0.010	-0.107 ***	-0.060 *	0.218 ***	0.130 ***	-0.015	-0.302 ***	0.412 ***	-0.027		
Localization													
14 None	0.096 ***	-0.115 ***	0.000	-0.096 ***	-0.049 †	0.175 ***	0.135 ***	-0.022	-0.263 ***	0.363 ***	-0.025	-0.812 ***	0.812 ***
15 Implicit	-0.112 ***	0.070 ***	0.067 *	-0.184 ***	0.190 ***	0.010	-0.063 *	-0.073 *	-0.013	-0.105 ***	0.143 ***	0.336 ***	-0.336 ***
16 Explicit	0.025	0.027 *	-0.061 *	0.242 ***	-0.134 ***	-0.146 ***	-0.050 †	0.084 **	0.219 ***	-0.191 ***	-0.109 ***	0.336 ***	-0.336 ***

†p < 0.05 *p < 0.01 **p < 0.001 ***p < 0.0001

Shading indicates a correlation strength of at least 0.1, with darker shading corresponding to stronger correlations.

Colors correspond to direction of correlation: green indicates positive correlation and red indicates negative correlation.

4.2.2. The Effect of Feedback Instruction on Peer Feedback

To answer the first research question – how different types of feedback instruction affect the practice of peer review – I analyzed the types of feedback that peers provided at two levels of analysis: the feedback item level and the draft level. At the feedback item level, I compared the overall distribution of feedback items within each dimension of the coding frame compared across the two groups (handout only and handout plus in-class instruction), using a crosstabulation and chi-square test for significance of feedback items, as described in section 3.7; the results are shown in Table 17.

As these results show, in-class instruction caused statistically-significant changes in feedback distribution in four of the five dimensions: *mode* ($\chi^2(3, N = 2808) = 16.770$, $p = 0.0008$), *focus* ($\chi^2(3, N = 2013) = 9.434$, $p = 0.0240$), *tone* ($\chi^2(3, N = 2013) = 36.544$, $p < 0.0001$), and *localization* ($\chi^2(3, N = 2013) = 12.806$, $p = 0.0017$). Differences between the two instruction types in the *type* dimension were not significant ($\chi^2(3, N = 2013) = 2.444$, $p = 0.1179$). Within the mode dimension, feedback in the in-class instruction group tended to use higher-level modes, shifting away from edits and authoritative comments toward readerly. Similarly, the in-class instruction group gave fewer form-focused comments, instead making more content-focused comments, both prose and technical. While nearly two-thirds of comments in the handout-only group were negative or neutral, nearly half of comments in the in-class group were positive or qualified – i.e., the in-class instruction group was more likely to say something positive about their peer's paper. In each of these cases, then, the in-class instruction appeared to positively impact the nature of the peer feedback by eliciting feedback more consistent with effective practices documented in writing research (see sections 2.5.3 and 2.5.2).

Table 17: Crosstabulation of Feedback Instruction and Feedback Features

	Total Items		Handout-Only		In-Class		sig
	N	%	N	%	N	%	p
Mode							0.0008**
Edit	795	28%	448	29%	347	28%	
Authoritative	1206	43%	694	45%	512	41%	
Coaching	431	15%	241	15%	190	15%	
Readerly	376	13%	172	11%	204	16%	
Focus							0.024*
Form	680	34%	400	36%	280	31%	
Content	1115	55%	598	54%	517	57%	
Technical	160	8%	75	7%	85	9%	
Extra-Textual	58	3%	34	3%	24	3%	
Tone							< 0.0001**
Negative	45	2%	31	3%	14	2%	
Neutral	1126	56%	678	61%	448	49%	
Qualified	396	20%	183	17%	213	24%	
Positive	446	22%	215	19%	231	25%	
Type							0.1179
Margin	1582	76%	882	76%	700	77%	
End	491	24%	285	24%	206	23%	
Localization							0.0017**
None	385	19%	229	21%	156	17%	
Implicit	574	29%	281	25%	293	32%	
Explicit	1055	52%	598	54%	457	50%	

*p < 0.05 **p < 0.01

The effect of peer review instruction on comment localization was more ambiguous: distribution of localization shifted away from both none and explicit towards implicit for the in-class group (i.e. the in-class instruction group made fewer non-localized comments, but they also made fewer explicitly-localized comments). Comparing non-localized comments to those with either type of localization, the in-class instruction group produced fewer non-localized comments than the handout-only group, consistent with the instructional emphasis on “being specific” and avoiding generalization when commenting; this binary comparison was also statistically-significant ($\chi^2(1, N = 2013) = 3.874$, $p = 0.0490$). Because end comments cannot have implicit localization, I also compared comment localization separately for end and margin comments. There is no statistically-significant difference between the in-class and handout-only groups for end comments; for margin comments, however, there is a statistically-significant difference ($\chi^2(2, N = 1507) = 10.553$, $p = 0.0012$): in the handout-only instruction group, 66% of marginalia had explicit localization versus 58% in the in-class instruction group.

Because of the complex relationship between localization and revision (see Figure 4), the implications of this result are difficult to interpret. In previous studies (Nelson & Schunn, 2009; Patchan et al., 2016), localization has been found to increase both feedback understanding and implementation (*what* to change is clearer), but to decrease revision quality (revision scope is narrower); however, those studies included only end comments, not marginalia. This result could also indicate that in-class instruction improved feedback practice: the higher proportion of implicitly-localized marginalia may entail the benefits of localization without the associated reduction in revision quality.

To further refine this analysis and explore whether the instructional style affected single-peer and multi-peer reviewing differently, Table 18 presents the same analysis but broken down by the two review types.

Table 18: Crosstabulation of Feedback Instruction and Feedback Features by Peer Review Type

	Single Peer					Multiple Peer				
	Handout-Only		In-Class			Handout-Only		In-Class		
	N	%	N	%	sig (p)	N	%	N	%	sig (p)
Mode					0.5710					< 0.0001 **
Edit	254	30%	140	31%		194	27%	207	26%	
Authoritative	362	43%	181	40%		332	46%	331	41%	
Coaching	110	13%	70	15%		131	18%	120	15%	
Readerly	109	13%	62	14%		63	9%	142	18%	
Focus					0.4086					0.0161 *
Form	181	31%	82	26%		219	42%	198	33%	
Content	341	59%	194	62%		257	49%	323	54%	
Technical	40	7%	27	9%		35	7%	58	10%	
Extra-Textual	19	3%	10	3%		15	3%	14	2%	
Tone					0.0310 *					< 0.0001 **
Negative	13	2%	4	1%		18	3%	10	2%	
Neutral	316	54%	154	49%		362	69%	294	50%	
Qualified	116	20%	89	28%		67	13%	124	21%	
Positive	136	23%	66	21%		79	15%	165	28%	
Type					0.8060					0.0356 *
Margin	435	75%	232	74%		387	74%	468	79%	
End	146	25%	81	26%		139	26%	125	21%	
Localization					0.3593					0.0001 **
None	106	18%	62	20%		123	23%	94	16%	
Implicit	157	27%	95	30%		124	24%	198	33%	
Explicit	319	55%	156	50%		279	53%	301	51%	

*p < 0.05 **p < 0.01

Interestingly, many of the effects seen in the full data set are not statistically significant in the feedback from students who only reviewed a single peer's report. In most cases, the difference between the instruction types is either negligible or in the same direction in both the multiple and single review groups, but only reaches significance in the case of multiple review. In the single-peer group, only the *tone* dimension had a statistically-significant difference between in-class and handout-only instruction. In contrast, the multi-peer review data set shows statistically significant differences across all dimensions. This finding suggests that reviewing multiple peer reports may have an independent effect or may enhance or be enhanced by the effect of in-class instruction.

While the chi-square statistic allows comparison of distribution within a dimension (e.g. mode), it does not indicate differences within particular sub-dimensions (e.g. readerly). To examine these differences, I conducted an analysis at the draft level, comparing the mean number of comments per draft for each sub-dimension of the coding frame pairwise using t-tests. The results for all peers are shown in Table 19.

Table 19: Comparison of Mean Number of Individual Feedback Features per Draft by Feedback Instruction (All Peers)

	Handout-Only		In-Class		sig (t)	
	M	SD	M	SD	p	
Feedback Items	15.25	11.81	17.65	13.56	0.2293	
Comments Only	10.85	6.19	12.76	6.81	0.0310	*
Mode						
Edit	4.39	8.64	4.89	11.35	0.7568	
Authoritative	6.80	4.78	7.21	5.07	0.5954	
Coaching	2.36	2.32	2.68	2.18	0.3655	
Readerly	1.69	2.24	2.87	2.48	0.0016	**
Focus						
Form	3.92	4.22	3.94	4.19	0.9729	
Content	5.86	3.89	7.28	4.44	0.0314	*
Technical	0.74	1.19	1.20	1.71	0.0514	
Extra-Textual	0.33	0.69	0.34	0.84	0.9692	
Tone						
Negative	0.30	1.02	0.20	1.13	0.5261	
Neutral	6.65	5.30	6.31	5.07	0.6734	
Qualified	1.79	2.07	3.00	3.31	0.0076	**
Positive	2.11	2.24	3.25	2.97	0.0069	**
Type						
Margin	8.06	6.03	9.86	6.22	0.0598	
End	2.79	2.48	2.90	2.09	0.7588	
Localization						
None	2.25	1.80	2.20	1.66	0.8573	
Implicit	2.75	3.05	4.13	3.22	0.0055	**
Explicit	5.86	5.34	6.44	5.39	0.4901	

*p < 0.05 **p < 0.01

Comparison of the in-class and handout-only instruction groups shows that statistically-significant differences in the mean number of feedback items per draft exist in five sub-dimensions of the coding frame: *readerly* (t(140) = 3.22, p = 0.0016), *content* (t(137) = 2.17, p =

0.0314), *qualified* ($t(107) = 2.72$, $p = 0.0076$), *positive* ($t(123) = 2.75$, $p = 0.0069$), and *implicit* ($t(145) = 2.82$, $p = 0.0055$). That is, with in-class instruction, students made more readerly comments, more content-focused comments, more comments using qualified or positive tone, and more implicitly-localized comments. As in the feedback item-level analysis, the shifts associated with in-class instruction are consistently toward more effective feedback practices.

Again, I broke the means tests down by the two review types to explore whether the instructional style affected single-peer and multi-peer reviewing differently. As shown in Table 20, only the difference in the number of *qualified* comments remains statistically significant when limited to the single-peer groups ($t(24) = 2.09$, $p = 0.0471$). Additionally, the difference in the *content* sub-dimension is no longer statistically significant in either case.

Table 20: Comparison of Mean Number of Individual Feedback Features per Draft by Feedback Instruction (Separately by Peer Review Type)

	Single Peer					Multiple Peer				
	Handout-Only		In-Class		sig (t)	Handout-Only		In-Class		sig (t)
	M	SD	M	SD	p	M	SD	M	SD	p
Feedback Items	16.06	13.34	21.57	20.10	0.2568	14.40	10.04	16.00	9.41	0.4130
Comments Only	11.17	6.45	14.90	8.48	0.0398 *	10.52	5.97	11.86	5.83	0.1294
Mode										
Edit	4.88	10.34	6.67	19.09	0.6893	3.88	6.48	4.14	5.77	0.8326
Authoritative	6.96	4.90	8.62	4.84	0.1950	6.64	4.69	6.62	5.09	0.9837
Coaching	2.12	2.22	3.33	3.01	0.1034	2.62	2.42	2.40	1.68	0.5990
Readerly	2.10	2.86	2.95	2.87	0.2556	1.26	1.23	2.84	2.33	< 0.0001 **
Focus										
Form	3.48	3.64	3.90	4.78	0.7173	4.38	4.74	3.96	3.97	0.6321
Content	6.56	4.28	9.24	5.54	0.0557	5.14	3.33	6.46	3.65	0.0617
Technical	0.77	1.28	1.29	1.27	0.1249	0.70	1.09	1.16	1.88	0.1383
Extra-Textual	0.37	0.74	0.48	1.03	0.6574	0.30	0.65	0.28	0.76	0.8874
Tone										
Negative	0.25	1.20	0.19	0.68	0.7906	0.36	0.80	0.20	1.28	0.4554
Neutral	6.08	5.34	7.33	5.99	0.4087	7.24	5.24	5.88	4.64	0.1725
Qualified	2.23	2.21	4.24	4.17	0.0471 *	1.34	1.83	2.48	2.76	0.0171 *
Positive	2.62	2.27	3.14	2.73	0.4388	1.58	2.11	3.30	3.09	0.0017 **
Type										
Margin	8.37	6.30	11.05	7.74	0.1683	7.74	5.78	9.36	5.47	0.1532
End	2.81	2.65	3.86	2.46	0.1141	2.78	2.32	2.50	1.79	0.5011
Localization										
None	2.04	1.74	2.95	1.91	0.0661	2.46	1.86	1.88	1.45	0.0860
Implicit	3.02	3.33	4.52	3.06	0.0713	2.48	2.73	3.96	3.29	0.0163 *
Explicit	6.13	5.35	7.43	6.44	0.4214	5.58	5.36	6.02	4.89	0.6692

*p < 0.05 **p < 0.01

These results again suggest that reviewing multiple peer reports may have an independent or interaction effect on the practice of peer review. Independent comparison of the multiple-peer and single-peer feedback groups confirms that there is a statistically-significant difference in the *content* (t(126) = 2.30, p = 0.0228) and *qualified* (t(133) = 2.10, p = 0.0375) sub-dimensions, as shown in Table 21. To further examine this relationship, I fit a least squares regression model on the content and qualified sub-dimensions as dependent variables with both instruction and feedback types as explanatory variables. For the *qualified* model, both independent variables have a statistically-significant relationship; in-class instruction increases the mean number of qualified comments per draft ($\beta = 0.73$, $t(172) = 3.57$, $p = 0.0005$), while multiple-peer feedback decreases the mean ($\beta = -0.60$, $t(172) = 2.96$, $p = 0.0035$). Similarly, in the *content* model, both are statistically significant and in-class instruction increases the mean number of content-focused comments per draft ($\beta = 0.91$, $t(172) = 2.87$, $p = 0.0046$), while multiple-peer feedback decreases the mean ($\beta = -0.96$, $t(172) = 3.02$, $p = 0.0029$).

Table 21: Means Comparison of Individual Feedback Features by Number of Reviews (All Peers)

	Single Peer		Multiple Peer		sig
	M	SD	M	SD	p
Feedback Items	17.64	15.64	15.20	9.71	
Mode					
Edit	5.40	13.33	4.01	6.11	
Authoritative	7.44	4.91	6.63	4.87	
Coaching	2.47	2.51	2.51	2.08	
Readerly	2.34	2.87	2.05	2.02	
Focus					
Form	3.60	3.97	4.17	4.36	
Content	7.33	4.80	5.80	3.54	0.0228*
Technical	0.92	1.29	0.93	1.55	
Extra-Textual	0.40	0.83	0.29	0.70	
Tone					
Negative	0.23	1.07	0.28	1.06	
Neutral	6.44	5.52	6.56	4.97	
Qualified	2.81	3.02	1.91	2.40	0.0375*
Positive	2.77	2.40	2.44	2.77	
Type					
Margin	9.14	6.80	8.55	5.66	
End	3.11	2.62	2.64	2.07	
Localization					
None	2.30	1.82	2.17	1.69	
Implicit	3.45	3.31	3.22	3.10	
Explicit	6.51	5.67	5.80	5.11	

*p < 0.05 **p < 0.01

While these results indicate that reviewers in the multiple peer condition gave fewer qualified and content comments, it is important to remember that these results are at the *review* level – they are the average number of each type of comment that a single reviewer wrote on a single author's draft. One obvious explanation for the decrease is that with more drafts to review, students spent less time on each. For revising, however, authors in the multiple peer condition received feedback from more than one reviewer; logically, because the differences in the review-level means are small compared to the means themselves, authors receiving feedback from multiple peers received more overall feedback than authors receiving feedback from a single peer. Table 22 summarizes the quantity of feedback received by authors – the quantity of feedback on which authors can base their revisions. As expected, authors in the multiple peer group received more feedback of every type; this difference is statistically significant for all feedback features except edit, extra-textual, negative and qualified. Therefore, the reduction in quantity of feedback given by any one reviewer in the multiple peer group is offset by the increased overall feedback received by authors across all reviewers.

Table 22: Means Comparison of Feedback Authors Received by Number of Reviews

	Single Peer		Multiple Peer		sig	
	M	SD	M	SD	p	
Feedback Items	18.35	16.48	28.68	14.83	0.0007	**
Mode						
Edit	5.97	14.28	7.57	8.82	0.4702	
Authoritative	7.42	4.94	12.51	7.35	0.0001	**
Coaching	2.60	2.62	4.74	2.84	0.0001	**
Readerly	2.37	2.86	3.87	3.17	0.0099	**
Focus						
Form	3.60	4.07	7.87	6.81	0.0001	**
Content	7.40	4.92	10.94	4.89	0.0002	**
Technical	0.93	1.18	1.75	2.43	0.0279	*
Extra-Textual	0.45	0.89	0.55	0.99	0.5869	
Tone						
Negative	0.23	1.13	0.53	1.45	0.2342	
Neutral	6.50	5.80	12.38	7.47	0.0001	**
Qualified	3.00	3.20	3.60	3.52	0.3445	
Positive	2.65	2.31	4.60	3.98	0.0024	**
Type						
Margin	9.18	7.05	16.13	8.77	0.0001	**
End	3.20	2.63	4.98	2.60	0.0005	**
Localization						
None	2.38	1.90	4.09	2.30	0.0001	**
Implicit	3.48	3.37	6.08	4.07	0.0004	**
Explicit	6.53	5.92	10.94	7.79	0.0011	**

*p < 0.05 **p < 0.01

4.3. RQ2: Factors Influencing the Efficacy of Peer Review

In the previous section, I presented the results of relationship between *feedback conditions* and *feedback features* in order to answer RQ1. In this section, I present the remainder of the analysis of peer review – considering how those *feedback features* affect *revision features* and ultimately *quality improvement,* as well as other potential relationships between variables.

4.3.1. The Nature of Revisions

Before analyzing the factors that influence the types of revisions students make to their drafts, it is helpful to understand the nature of revisions as a whole. Once a student has received feedback, they must then decide what changes to make when revising their draft. In Chapter 3, I described a method for quantifying and classifying the changes authors made when revising their reports. First, the overall extent of changes was rated as either none, minor, or major. Next, changes were classified and quantified according to a taxonomy developed from prior theoretical work consisting of two main categories, each with two subcategories. Changes can either be a correction – based on either a general or assignment-specific rule, or a revision – either at the micro or macrostructure level. The overall quantity of these four types of *revision features* was rated as none, minimal, moderate, or extensive. The distributions of these variables are shown in Figure 16.

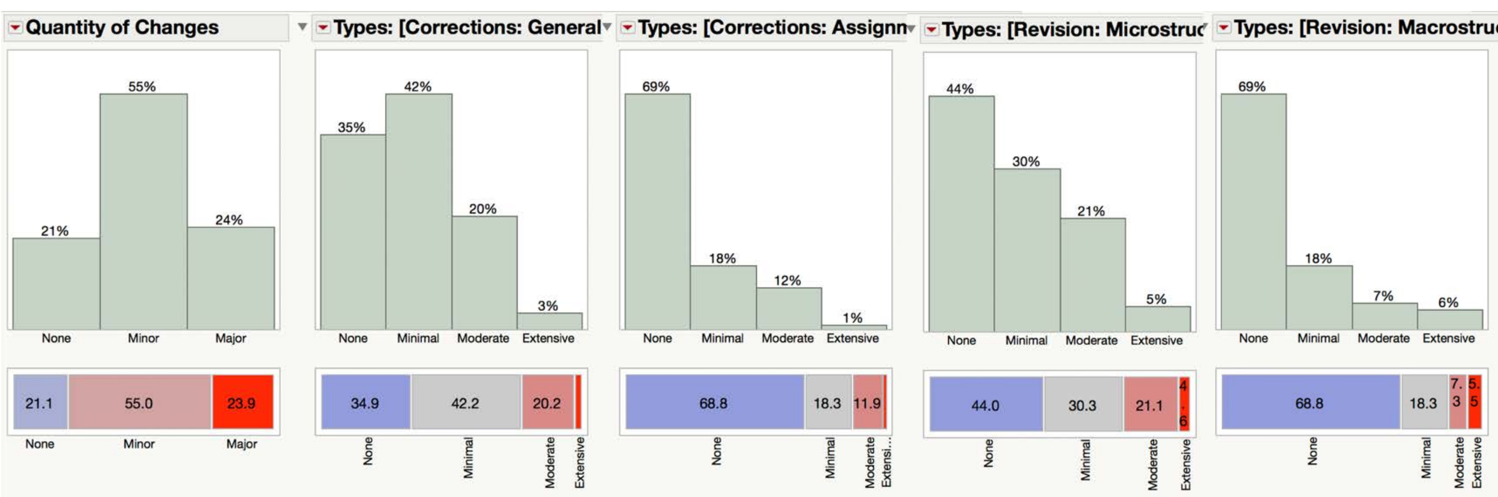


Figure 16: Histograms of Revision Features

- Over half (55%) of authors made only a minor quantity of changes when revising their drafts. The rest of the authors were roughly evenly split between making no changes (21% of authors) and making major changes (24% of authors).

- The vast majority of authors (69%) did not make any macrostructure revisions, meaning all of the changes they made were at the sub-paragraph level.
- Most authors (56%) did make some level of microstructure revisions.

4.3.2. Factors Influencing Revision Features

Relationships Between Feedback Features and Revision Features

Working at the author level, I analyzed the relationship between the *feedback features* the writer received from their peer(s) and the quantity of revisions – both overall and for each of the *revision features*. Because the independent variable in these analyses is continuous in nature and the dependent variable is ordinal, I used logistic regression to determine whether a statistically significant relationship existed between each feedback feature and each kind of revision feature. The results of the logistic regressions are reported in Table 23. Because of the very small number of students who made "extensive" revisions in any category, the regressions were re-run with the top two categories collapsed, which did not change any of the statistically-significant predictors.

Table 23: Relationships Between Feedback Features and Revision Features

Revision Feature	Significant Predictors
Overall Quantity of Revisions	Number of Feedback Items (+) (both edits and comments)
Corrections	
General Writing Rule	Number of Edits (+), Number of Comments with Form-focus (+)
Assignment-Specific Rule	None
Revisions	
Microstructure	Number of Comments with Negative Tone (-), Number of Margin Coments (+)
Macrostructure	None

Overall quantity of changes is only predicted by total number of feedback items. General Rule corrections are predicted by edits and form-focused comments. For microstructure revisions, margin comments are a positive predictor while comments with a negative tone are a negative predictor. All predictors had a Nagelkerke R^2 value well under 0.1 indicating that all are very weak predictors. There were no statistically significant predictors for assignment rule corrections or macrostructure revisions. The lack of predictors for macrostructure revisions is especially unfortunate because those were the revision type the peer review exercise was intended to elicit.

Relationships Between Feedback Conditions and Feedback Features

In addition to the direct relationship between feedback features and revision features, it was also hypothesized that there might be an indirect relationship between the feedback conditions and revision features. To test this hypothesis, I used a crosstabulation and Pearson chi-square test for significance. Because of the extremely low number of authors making "extensive" revisions of any type, I collapsed moderate and extensive into a single category. Results of these analyses are presented in Table 24 and Table 25. I also tested a further-reduced model for each revision category, combining any quantity of changes into a single category; the significance of these comparisons is reported in the "sig (reduced)" columns in Table 24 and Table 25.

Table 24: Crosstabulation of Peer Review Type and Revision Features

	Total		Single Peer		Multiple Peer		sig	sig
	N	%	N	%	N	%		(reduced)
Overall Quantity							0.5879	0.4073
Major	26	24%	12	21%	14	27%		
Minor	60	55%	32	55%	28	55%		
None	23	21%	14	24%	9	18%		
Correction: General							0.1358	0.2628
Moderate-Extensive	25	23%	9	16%	16	31%		
Minimal	46	42%	26	45%	20	39%		
None	38	35%	23	40%	15	29%		
Correction: Assignment							0.0975	0.0349 *
Moderate-Extensive	14	13%	6	10%	8	16%		
Minimal	20	18%	7	12%	13	25%		
None	75	69%	45	78%	30	59%		
Revision: Micro							0.0833	0.0348 *
Moderate-Extensive	28	26%	11	19%	17	33%		
Minimal	33	30%	16	28%	17	33%		
None	48	44%	31	53%	17	33%		
Revision: Macro							0.7896	0.7067
Moderate-Extensive	14	13%	7	12%	7	14%		
Minimal	20	18%	12	21%	8	16%		
None	75	69%	39	67%	36	71%		

*p < 0.05 **p < 0.01

Table 25: Crosstabulation of Peer Review Instruction and Revision Features

	Total		Handout Only		In Class		sig	sig
	N	%	N	%	N	%		(reduced)
Overall Quantity							0.0487	0.0956
Major	26	24%	18	28%	8	18%		
Minor	60	55%	29	45%	31	69%		
None	23	21%	17	27%	6	13%		
Correction: General							0.1406	0.2725
Moderate-Extensive	25	23%	17	27%	8	18%		
Minimal	46	42%	22	34%	24	53%		
None	38	35%	25	39%	13	29%		
Correction: Assignment							0.3589	0.2134
Moderate-Extensive	14	13%	8	13%	6	13%		
Minimal	20	18%	9	14%	11	24%		
None	75	69%	47	73%	28	62%		
Revision: Micro							0.5958	0.4765
Moderate-Extensive	28	24%	17	23%	11	24%		
Minimal	33	28%	17	23%	16	36%		
None	57	48%	39	53%	18	40%		
Revision: Macro							0.4584	0.9877
Moderate-Extensive	14	13%	10	16%	4	9%		
Minimal	20	18%	10	16%	10	22%		
None	75	69%	44	69%	31	69%		

*p < 0.05 **p < 0.01

The only statistically significant relationships between feedback conditions and revision features are for assignment rule-based corrections ($\chi^2(1, N = 109) = 4.451, p = 0.0349$) and for microstructure revisions ($\chi^2(1, N = 109) = 4.455, p = 0.0348$) under the reduced models (where revision types are combined into present or not present). Of authors in the single peer condition, only 22% made assignment rule-based corrections, whereas 41% of authors in the multiple peer group did. Similarly, 47% of authors in the single peer condition made microstructure revisions compared to 66% in the multiple peer condition.

4.3.3. Factors Influencing Quality Improvement

The final stage of analysis for understanding the efficacy of peer review is quality improvement. As shown in the research model (Figure 12), the features of revisions authors made are expected to be the primary predictor of quality improvement. In this section, I first examine this relationship and subsequently examine and independent effect of feedback features and conditions.

Relationships Between Revision Features and Quality Improvement

In order to examine and quantify the magnitude of the relationships between revision types and quality improvement, I conducted four series of linear regression models. In each series, I was interested in the relationship between the types of revisions made and one of the four outcome metrics (see Appendix A):

1. mechanics quality improvement,
2. writing quality improvement,
3. technical quality improvement, and
4. overall quality improvement.

I included two control variables in these models: initial draft rating and whether the author was in the single or multiple peer condition. Initial draft rating was important to control for because authors with a lower-rated draft have more room to increase their score. Multiple vs single peer was included because of the independent relationship it had on both feedback peers gave and revisions authors made. In preliminary analysis, I also controlled for peer review instruction, but it had no effect and was removed.

For each series of models, I first ran a baseline model consisting of only the control variables. I then ran a partial model for each type of revision plus the control variables. Next, I ran a full model with all predictors. Finally, I ran a model with predictors which were statistically insignificant in the full model removed. I compared the fit of each model using the Akaike Information Criterion (AIC). The results of these analyses are reported below.

Mechanics Quality Improvement

Table 26 shows the results of the regression for improvement in mechanics quality. The baseline model for mechanics indicated that draft quality rating had a statistically significant relationship with mechanics quality improvement, with an initial rating predicting improvement of .435 for an initial draft rating of "B" and .719 for an initial draft rating of "C." The baseline model had weak explanatory power, with Nagelkerke R^2 indicating only 7.9% of variation was explained. The best fit model was the partial model for general rule corrections, with a minimal quantity predicting improvement of .72, a moderate predicting 1.129, and extensive predicting 1.628. This model had good explanatory power, with 39% of the variation explained.

Table 26: Effect of the Types of Revisions Authors Made on Quality Improvement in Mechanics

Mechanics N=111	Control Variables			Corrections: General			Corrections: Assignment			Revisions: Microstructure			Revisions: Macrostructure			Full Model					
Author-Level Variables	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE
Mean outcome	.175	.298	.1685	-.210	.160	.1495	.096	.571	.1698	-.003	.986	.1678	.172	.310	.1693	-.232	.117	.1479	-.235	.118	.1505
Δ in outcome due to:																					
Control Variables																					
Multiple Peer (ref=SP)	.179	.222	.1470	.046	.699	.1190	.120	.416	.1479	.069	.625	.1415	.185	.209	.1477	-.023	.848	.1204	.013	.911	.1194
Draft Score (ref=A)																					
B	.435	.031	.2021	.318	.050	.1622	.465	.022	.2030	.394	.039	.1907				.352	.033	.1652			
C	.719	.000	.2038	.414	.014	.1679	.707	.000	.2018	.588	.003	.1968				.427	.013	.1726			
D	.408	.107	.2532	.308	.129	.2029	.398	.109	.2484	.345	.161	.2462				.347	.090	.2047			
Corrections: General (ref=None)																					
Extensive				1.628	.000	.2961										1.490	.000	.3047	1.508	.000	.3050
Moderate				1.129	.000	.1723										1.100	.000	.1811	1.030	.000	.1801
Minimal				.720	.000	.1352										.641	.000	.1399	.663	.000	.1390
Corrections: Assignment (ref=None)																					
Extensive							.077	.921	.7734							.268	.664	.6870			
Moderate							.264	.251	.2297							.313	.195	.1090			
Minimal							.370	.054	.1925							.030	.167	.8570			
Revisions: Micro (ref=None)																					
Extensive										.916	.009	.3493				.442	.307	.1490	.452	.137	.3033
Moderate										.378	.047	.1904				.098	.175	.5760	.086	.603	.1656
Minimal										.549	.001	.1657				.220	.149	.1390	.209	.158	.1478
Revisions: Macro (ref=None)																					
Extensive													-.233	.482	.3312	-.301	.288	.2950			
Moderate													-.183	.529	.2905	-.417	.232	.0710			
Minimal													.012	.951	.1957	-.162	.162	.3170			
Model Fit Statistics:																					
Deviance	271.085			233.547			239.48			197.359			250.228			154.431			40.563		
Akaike's Information Criterion (AIC)	268.714			224.573			233.252			260.456			273.842			233.252			227.263		
Rank by AIC	6			1			4			7			8			4			2		

Writing Quality Improvement

Table 27 shows the results of the regression for improvement in writing quality. The baseline model for writing quality indicated that draft quality rating had a statistically significant relationship with mechanics quality improvement, with an initial rating predicting improvement of 1.648 for C and 1.174.719 for D. Nagelkerke R^2 indicated that 13.4% of variation was explained by the baseline. The best fit model was the full model with the statistically insignificant variable (macrostructure revisions) removed. In this model, a moderate number of general rule corrections predicted an increase in the score of .774, extensive assignment rule corrections predicting an increase of 5.182. Microstructure revision predicted increases of 1.183, 1.737, and 2.056 points at the minimal, moderate, and extensive levels respectively. This model had high explanatory power, with 46% of the variation explained.

Table 27: Effect of the Types of Revisions Authors Made on Improvement in Writing Quality

Writing Quality N=111	Control Variables			Corrections: General			Corrections: Assignment			Revisions: Microstructure			Revisions: Macrostructure			Full Model			Insignificant Removed		
Author-Level Variables	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE
Mean outcome	.171	.618	.3424	-.023	.948	.3534	.065	.844	.3307	-.478	.123	.3102	.057	.863	.3318	-.312	.283	.2907	-.294	.321	.2956
Δ in outcome due to:																					
Control Variables																					
Multiple Peer (ref=SP)	.690	.021	.2987	.495	.078	.2812	.487	.091	.2881	.347	.184	.2616	.754	.009	.2895	.224	.344	.2367	.133	.575	.2369
Draft Score (ref=A)																					
D	1.174	.023	.5146	1.045	.029	.4797	1.175	.015	.4838	1.196	.009	.4551	1.015	.043	.5026	1.125	.005	.4023	1.202	.003	.4039
C	1.648	.000	.4142	1.242	.002	.3969	1.538	.000	.3930	1.412	.000	.3638	1.298	.002	.4160	.998	.003	.3393	1.109	.001	.3333
B	.745	.070	.4106	.633	.099	.3835	.921	.020	.3954	.555	.116	.3525	.487	.229	.4043	.452	.164	.3247	.581	.072	.3231
Corrections: General (ref=None)																					
Extensive				1.966	.005	.7000										1.270	.034	.5989	1.170	.054	.6072
Moderate				1.405	.001	.4073										.681	.056	.3560	.774	.032	.3602
Minimal				.238	.456	.3197										-.276	.316	.2750	-.312	.263	.2790
Corrections: Assignment (ref=None)																				.	
Extensive							4.910	.001	1.5063							5.476	.000	1.3043	5.182	.000	1.2378
Moderate							-.156	.727	.4475							-.256	.505	.3838	-.201	.603	.3857
Minimal							.787	.036	.3748							-.044	.893	.3283	.084	.797	.3279
Revisions: Micro (ref=None)																				.	
Extensive										2.242	.001	.6458				1.984	.001	.6025	2.056	.001	.6132
Moderate										1.810	.000	.3520				1.695	.000	.3449	1.737	.000	.3337
Minimal										1.476	.000	.3063				1.120	.000	.2919	1.183	.000	.2941
Revisions: Macro (ref=None)																					
Extensive													.626	.334	.6491	-.230	.683	.5651			
Moderate													.987	.083	.5693	.768	.091	.4551			
Minimal													1.042	.007	.3836	.426	.181	.3182			
Model Fit Statistics:																					
Deviance	271.085			233.547			239.484			197.359			250.228			148.457			154.431		
Akaike's Information Criterion (AIC)	426.117			415.572			418.358			396.884			423.230			383.280			381.659		
Rank by AIC	7			4			5			3			6			2			1		

Technical Quality Improvement

Table 28 shows the results of the regression for improvement in technical quality. The baseline model for technical quality indicated that draft quality rating had a statistically significant relationship with technical quality improvement, with an initial rating predicting improvement of 2.691 for B, 2.542 for C and 2.557 for D. Nagelkerke R^2 was weak, indicating 4.3% of variation was explained by the baseline. The best fit model was the full model with the statistically-insignificant variable (general rule corrections) removed. This model had high explanatory power, with 62% of the variation explained. Interestingly, this was the only one of the four series of models in which the best fit models predicted a revision type (assignment rule revisions) that decreased the score. It was also the only best fit model in which draft quality rating was not a statistically-significant factor.

Table 28: Effect of the Types of Revisions Authors Made on Improvement in Technical Quality

Technical Quality N=111	Control Variables			Corrections: General			Corrections: Assignment			Revisions: Microstructure			Revisions: Macrostructure			Full Model			Insignificant Removed		
Author-Level Variables	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE
Mean outcome	.471	.579	.8493	.264	.779	.9390	.501	.564	.8695	-.712	.395	.8372	.231	.701	.6018	-.120	.837	.5835	-.366	.510	.5550
Δ in outcome due to:																					
Control Variables																					
Multiple Peer (ref=SP)	.369	.618	.7407	.380	.611	.7473	.362	.632	.7575	-.172	.807	.7061	.307	.558	.5251	.254	.592	.4752	.220	.645	.4766
Draft Score (ref=A)																					
D	2.557	.045	1.2763	2.662	.037	1.2748	2.538	.046	1.2721	2.768	.024	1.2283	1.015	.266	.9115	1.296	.109	.8076	1.237	.131	.8192
C	2.542	.013	1.0273	2.564	.015	1.0548	2.450	.018	1.0335	2.331	.018	.9820	.436	.564	.7545	.920	.177	.6812	.681	.313	.6745
B	2.691	.008	1.0184	2.754	.007	1.0191	2.510	.016	1.0398	2.295	.016	.9514	1.164	.112	.7334	.945	.147	.6519	.796	.225	.6553
Corrections: General (ref=None)																					
Extensive				-1.551	.404	1.8601										-1.714	.154	1.2023			
Moderate				.382	.724	1.0824										-1.068	.135	.7148			
Minimal				.373	.661	.8496										-.791	.152	.5521			
Corrections: Assignment (ref=None)																					
Extensive							.687	.862	3.9605							-9.200	.000	2.6186	-9.139	.001	2.6485
Moderate							.926	.431	1.1765							-1.584	.040	.7705	-1.523	.041	.7471
Minimal							-.263	.790	.9855							-1.372	.037	.6591	-1.617	.013	.6536
Revisions: Micro (ref=None)																					
Extensive										2.592	.137	1.7430				2.945	.015	1.2096	2.563	.032	1.1983
Moderate										3.672	.000	.9501				3.343	.000	.6923	3.065	.000	.6876
Minimal										2.297	.005	.8268				2.075	.000	.5859	1.707	.002	.5595
Revisions: Macro (ref=None)																					
Extensive													10.122	.000	1.1773	10.862	.000	1.1345	10.896	.000	1.1403
Moderate													6.925	.000	1.0325	7.157	.000	.9137	7.156	.000	.9163
Minimal													3.191	.000	.6957	2.484	.000	.6388	2.513	.000	.6445
Model Fit Statistics:																					
Deviance	1667.701			1649.059			1655.68			1437.653			823.211			598.338			618.374		
Akaike's Information Criterion (AIC)	627.778			632.530			632.975			617.302			555.413			537.998			535.654		
Rank by AIC	5			6			7			4			3			2			1		

Overall Quality Improvement

Table 29 shows the results of the regression for improvement in overall quality. The baseline model for overall quality was able to explain 10.8% of variation in total score. The best fit model was the full model with the statistically insignificant variables (general rule and assignment corrections) removed. This model had good explanatory power, with 57% of the variation explained. In this model, lower initial draft ratings and higher numbers of micro and macrostructure revisions predicted increases in total quality improvement. Corrections were likely not significant predictors because the total quality improvement score was weighted towards technical quality by the original assignment rubric weights.

Table 29: Effect of the Types of Revisions Authors Made on Improvement in Overall Quality

Overall Quality N=111	Control Variables			Corrections: General			Corrections: Assignment			Revisions: Microstructure			Revisions: Macrostructure			Full Model			Insignificant Removed		
Author-Level Variables	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE	Coef.	*p*	SE
Mean outcome	.817	.439	1.0559	.031	.979	1.1480	.662	.538	1.0750	-1.193	.215	.9630	.460	.578	.8276	-.665	.397	.7843	-.997	.187	.7550
Δ in outcome due to:																					
Control Variables																					
Multiple Peer (ref=SP)	1.239	.179	.9210	.921	.313	.9136	.970	.300	.9366	.245	.763	.8122	1.247	.084	.7221	.455	.476	.6387	.380	.555	.6440
Draft Score (ref=A)																					
D	4.138	.009	1.5868	4.016	.010	1.5587	4.111	.009	1.5728	4.308	.002	1.4128	2.477	.048	1.2536	2.768	.011	1.0855	2.828	.012	1.1194
C	4.909	.000	1.2772	4.221	.001	1.2896	4.694	.000	1.2777	4.331	.000	1.1295	2.486	.017	1.0377	2.345	.010	.9156	2.535	.006	.9224
B	3.871	.002	1.2662	3.705	.003	1.2459	3.896	.002	1.2855	3.244	.003	1.0943	2.111	.036	1.0086	1.749	.046	.8762	1.934	.027	.8736
Corrections: General (ref=None)																					
Extensive				2.043	.369	2.2742										1.046	.518	1.6160			
Moderate				2.916	.028	1.3234										.713	.458	.9607			
Minimal				1.331	.200	1.0387										-.425	.567	.7421			
Corrections: Assignment (ref=None)																					
Extensive							5.674	.247	4.8968							-3.456	.326	3.5196			
Moderate							1.034	.477	1.4546							-1.527	.140	1.0356			
Minimal							.895	.463	1.2185							-1.386	.118	.8859			
Revisions: Micro (ref=None)																					
Extensive										5.750	.004	2.0049				5.371	.001	1.6258	4.807	.002	1.5847
Moderate										5.860	.000	1.0928				5.137	.000	.9306	4.786	.000	.9101
Minimal										4.321	.000	.9510				3.415	.000	.7875	3.311	.000	.7565
Revisions: Macro (ref=None)																					
Extensive													10.516	.000	1.6191	10.330	.000	1.5248	9.276	.000	1.4221
Moderate													7.729	.000	1.4200	7.508	.000	1.2281	7.215	.000	1.2440
Minimal													4.245	.000	.9568	2.748	.001	.8586	2.676	.002	.8742
Model Fit Statistics:																					
Deviance	2577.855			2465.152			2530.94			1901.986			1556.978			1080.948			1162.693		
Akaike's Information Criterion (AIC)	676.120			677.157			680.081			648.369			626.152			603.647			599.739		
Rank by AIC:	5			6			7			4			3			2			1		

Relationships Between Feedback Features and Quality Improvement

In order to examine any other relationships that affected quality improvement, I ran another series of regression models for each of the four quality improvement metrics, this time with feedback features as predictors. Comparing to the same baseline models as before (i.e. control variables are the same), I used the following models for this analysis:

1. Control variables only,
2. All first-order feedback features,
3. A reduced first-order model using backward stepwise analysis to remove variables; the stepwise procedure used was minimization of AICc,
4. A model including two-way interactions between feedback features, added stepwise, again minimizing AICc,
5. The best-fit revision-level model from the previous section,
6. The combination of the best-fit comment-level model with the best-fit revision-level model.

The results of this analysis for each quality improvement metric are reported below. The R^2 values reported indicate the percent of variation in the metric that is explained by the model predictors. The adjusted R^2 values are adjusted for the number of model variables. The AICc values are also a measure used for model comparison, with lower values indicating an improved model; AICc is a more conservative measure than adjusted R^2 as it more harshly penalizes additional predictor variables.

Mechanics Quality Improvement

For mechanics quality, the best comment-level model (Model 4) increased explanation of quality improvement from 11.2% to 29.5% over the baseline (control variable only) model. Regression results indicated that several comment features and interactions significantly predicted change in mechanics quality. Predictors with a positive effect included: comment features of Edit (β = 0.0203, p = 0.0009) and Extra-Textual (β = 0.1753, p = 0.0277), as well as interaction between Technical*Extra-Textual (β = 0.1057, p = 0.0073); there were no predictors with a negative effect.

While the best revision-level model explained 43.5% of the variation in mechanics quality improvement, combining the best comment-level and revision-level models (Model 6) resulted in a small additional increase, with 49.3% of variation explained. Regression results indicated that in the combined model, only the interaction effect of Technical*Extra-Textual (β =0.0960, p = 0.0050) was significant at the comment level. Comparisons of the models are summarized in Table 30.

Table 30: Model Comparison for Mechanics Quality Improvement

Model	F Ratio	R^2	R^2 (adj)	AICc	Δ R^2 (adj)	Δ AICc
1. Control Variables Only	3.35	.112	.079	269.522		
2. All Comment Features (first order)	2.41	.321	.188	276.336		
Difference from Model 1					.109	6.814
3. Comment Features Added Stepwise	5.87	.218	.181	257.679		
Improvement from Model 2					-.007	-18.658
4. Comment Features Added Stepwise With addition of interactions added stepwise	5.33	.295	.240	253.355		
Improvement from Model 3					.058	-4.324
5. Best-Fit Revision Model	11.33	.435	.397	226.356		
Difference from Model 1					.318	-43.167
6. Combined Comment and Revision Model Model 5 and Model 4 Predictors Combined	8.77	.493	.437	224.214		
Improvement from Model 5					.041	-2.142

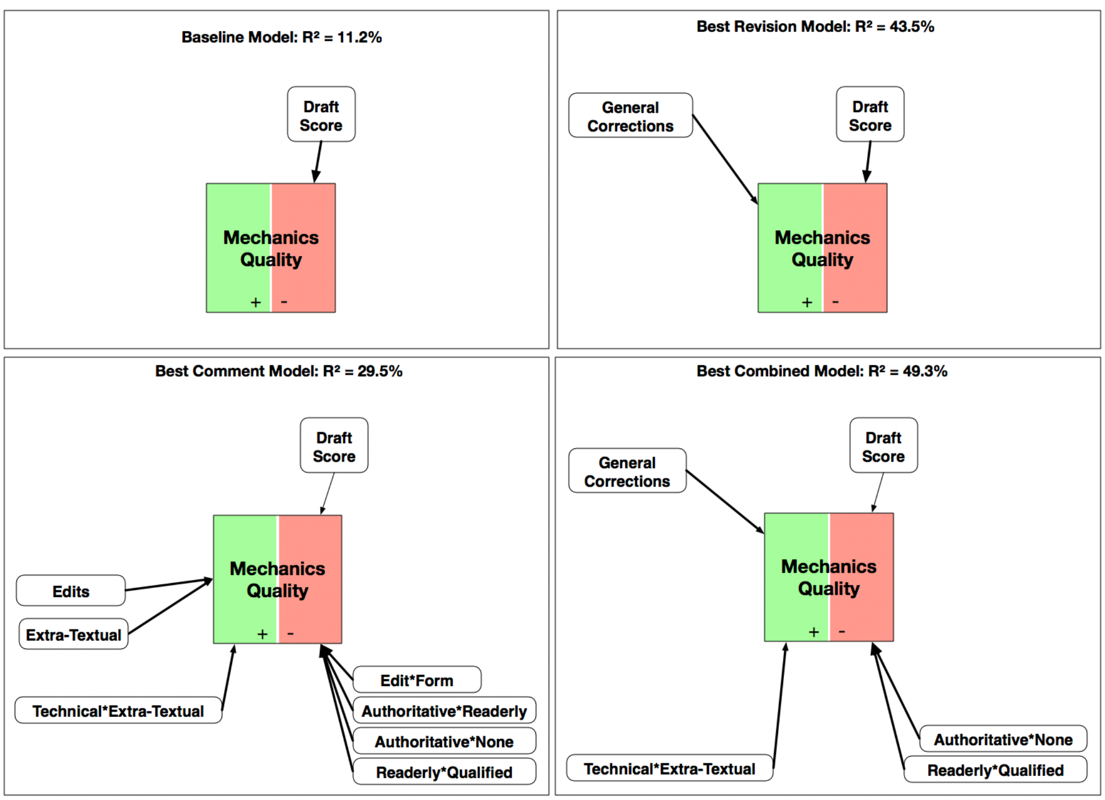


Figure 17: Graphical Comparison of Control, Direct, Indirect, and Combined Regression Models for Mechanics Quality Improvement

Writing Quality Improvement

For writing quality, the best comment-level model (Model 4) increased explanation of quality improvement from 16.6% to 48.6% over the baseline (control variable only) model. Regression results indicated that several comment features and interactions significantly predicted change in writing quality. Predictors with a positive effect included: comment features of Edit (β = 0.0501, p = 0.0002), Form (β = 0.2000, p = 0.0208), and Content (β = 0.2396, p = 0.0056), as well as interactions between Form*Content (β = 0.0233, p = 0.0095), Form*None (β = 0.0639, p = 0.0032), and Qualified*None (β = 0.1025, p = 0.0008). Predictors with a negative effect included interactions between Edit*Form (β = -0.006616, p = 0.0426), Authoritative*Readerly (β= -0.026251, p = 0.0397*), Authoritative*None (β = -0.119704, p = 0.0003*) and Readerly*Qualified (β = -0.047965, p = 0.0024*).

While the best revision-level model explained 52.5% of the variation in mechanics quality improvement, combining the best comment-level and revision-level models (Model 6) resulted in an additional increase, with 66.6% of variation explained. Regression results indicated that in the combined model, several comment features and interactions significantly predicted change in writing quality. Predictors with a positive effect included: comment features of Form (β = 0.1980, p = 0.0107), Content (β = 0.2443, p = 0.0025), and Qualified (β = 0.1450, p = 0.0065), as well as interactions between Comments*None (β = 0.0623, p = 0.0059). Predictors with a negative effect included only interactions between Authoritative*None (β = -0.0847, p = 0.0048), and Readerly*Qualified (β = -0.0387, p = 0.0104). Comparisons of the models are summarized in Table 31.

Table 31: Model Comparison for Writing Quality Improvement

Model	F Ratio	R^2	R^2 (adj)	AICc	ΔR^2 (adj)	Δ AICc
1. Control Variables Only	5.26	.166	.134	426.924		
2. All Comment Features (first order)	2.22	.302	.166	443.593		
Difference from Model 1					.032	16.668
3. Comment Features Added Stepwise	5.07	.226	.182	423.182		
Improvement from Model 2					.016	-20.411
4. Comment Features Added Stepwise With addition of interactions added stepwise	4.01	.486	.365	418.977		
Improvement from Model 3					.184	-4.205
5. Best-Fit Revision Model	8.24	.525	.461	386.711		
Difference from Model 1					.327	-40.213
6. Combined Comment and Revision Model Model 5 and Model 4 Predictors Combined	5.32	.666	.541	403.490		
Improvement from Model 5					.080	16.779

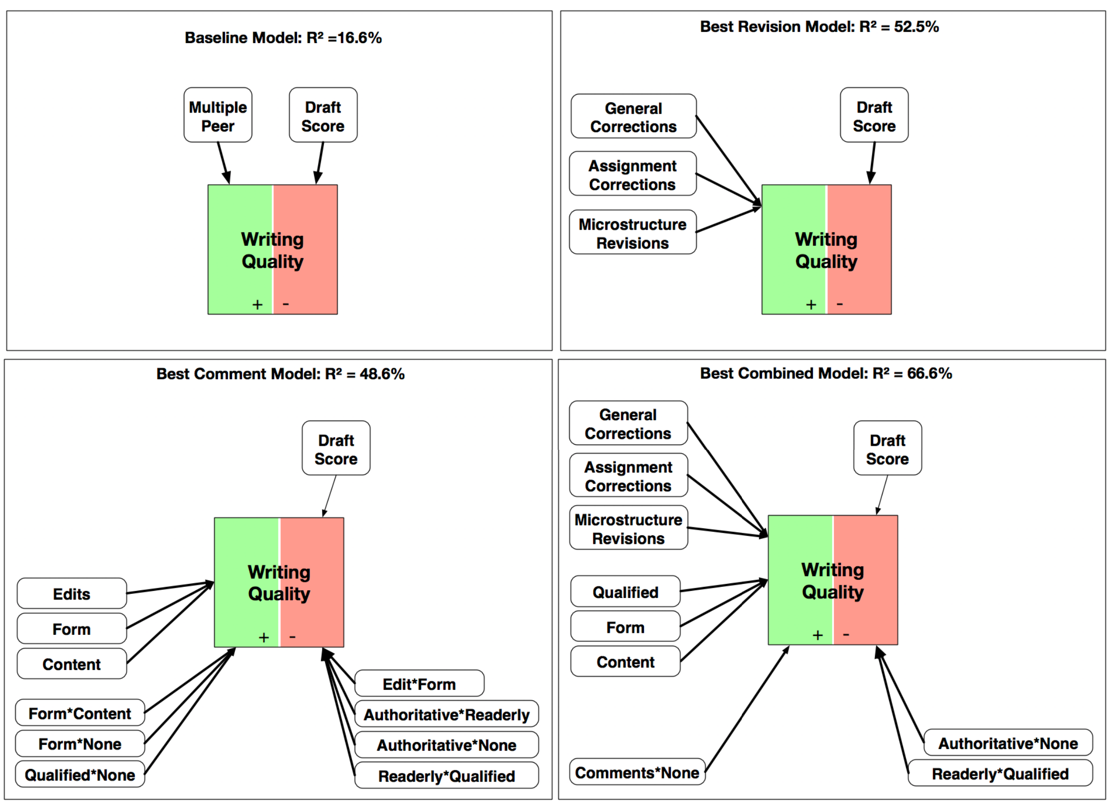


Figure 18: Graphical Comparison of Control, Direct, Indirect, and Combined Regression Models for Writing Quality Improvement

Technical Quality Improvement

For technical quality, the best comment-level model (Model 4) increased explanation of quality improvement from 7.5% to 41.2% over the baseline (control variable only) model. Regression results indicated that several comment features and interactions significantly predicted change in technical quality. Predictors with a positive effect included: comment features of Content (β = 0.5799, p < .0001) and Technical (β = 0.7898, p =0.0012), as well as interactions between Positive*Implicit (β = 0.1009, p =0.0012), and Neutral*Positive (β = 0.0489, p = 0.0313). Predictors with a negative effect included: comment features of Positive (β = -0.5711, p =0.0016), Implicit (β = -0.3863, p = 0.0025), and Neutral (β = -0.174, p = 0.0308), as well as interactions between Coaching*Technical (β = -0.3929, p < .0001) and Coaching*Positive (β = -0.1274, p =0.0118).

While the best revision-level model explained 41.2% of the variation in mechanics quality improvement, combining the best comment-level and revision-level models (Model 6) resulted in an additional increase, with 76.6% of variation explained. Regression results indicated that in the combined model, several comment features and interactions significantly predicted change in technical quality. Predictors with a positive effect included: comment features of Content (β = 0.3370, p = 0.0007) and Technical (β = 0.6331, p = 0.0002); positive interaction effects ceased to be significant in the combined model. Predictors with a negative effect included: comment features of Positive (β= -0.3051, p = 0.0167), Implicit (β = -0.1791, p = 0.0428), and Neutral (β = -0.1168, p = 0.0366), as well as interactions between Coaching*Technical (β = -0.2293, p = 0.0004) and Coaching*Positive (β = -0.0778, p = 0.0299). Comparisons of the models are summarized in Table 32.

Table 32: Model Comparison for Technical Quality Improvement

Model	F Ratio	R^2	R^2 (adj)	AICc	Δ R^2 (adj)	Δ AICc
1. Control Variables Only	2.13	.075	.040	628.586		
2. All Comment Features (first order)	1.35	.209	.054	647.696		
Difference from Model 1					.015	19.110
3. Comment Features Removed Stepwise	2.71	.135	.085	625.655		
Improvement from Model 2					.031	-22.041
4. Comment Features Removed Stepwise	4.12	.412	.312	608.822		
With addition of interactions added stepwise						
Improvement from Model 3					.227	-16.834
5. Best-Fit Revision Model	14.28	.657	.611	540.707		
Difference from Model 1					.571	-87.879
6. Combined Comment and Revision Model	11.11	.766	.697	535.529		
Model 5 and Model 4 Predictors Combined						
Improvement from Model 5					.086	-5.177

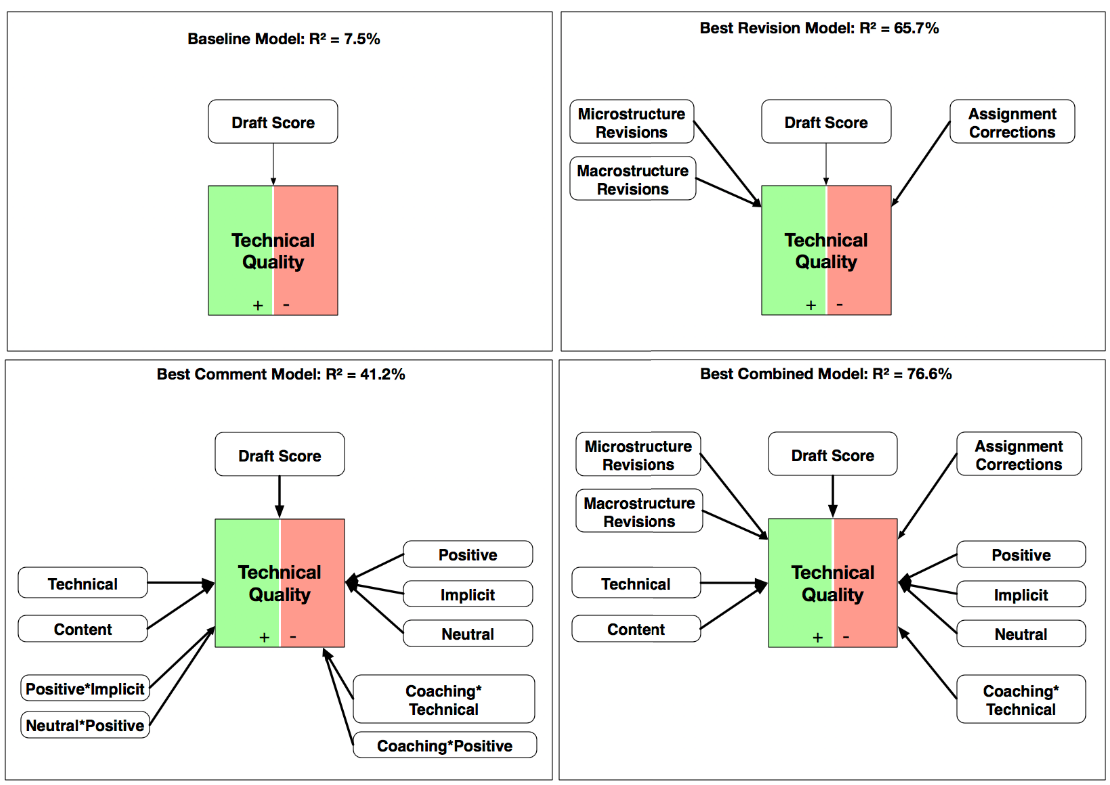


Figure 19: Graphical Comparison of Control, Direct, Indirect, and Combined Regression Models for Technical Quality Improvement

Overall Quality Improvement

For overall quality, the best comment-level model (Model 4) increased explanation of quality improvement from 14.1% to 38.3% over the baseline (control variable only) model. Regression results indicated that several comment features predicted change in overall quality. The only predictor with a positive effect on overall quality was Content ($\beta = 0.4216$, p = 0.0031); the only predictor with a negative effect was Implicit ($\beta = -0.4387$, p = 0.0066).

While the best revision-level model explained 61.3% of the variation in overall quality improvement, combining the best comment-level and revision-level models (Model 6) resulted in an additional increase, with 71.2% of variation explained. Content ($\beta = 0.3092$, p = 0.0027) and Implicit ($\beta = -0.2809$, p = 0.0159) remained statistically significant in the combined model; Technical ($\beta = 0.5171$, p = 0.0155) also reached significance as a positive predictor. Comparisons of the models are summarized in Table 33.

Table 33: Model Comparison for Overall Quality Improvement

Model	F Ratio	R^2	R^2 (adj)	AICc	Δ R^2 (adj)	Δ AICc
1. Control Variables Only	4.35 **	.141	.109	676.927		
2. All Comment Features (first order)	1.85 *	.252	.116	695.052		
Difference from Model 1					.007	18.125
3. Comment Features Added Stepwise	4.50 **	.206	.160	672.799		
Improvement from Model 2					.044	-22.254
4. Comment Features Added Stepwise With addition of interactions added stepwise	4.25 **	.383	.293	665.207		
Improvement from Model 3					.133	-7.591
5. Best-Fit Revision Model	15.81 **	.613	.574	602.923		
Difference from Model 1					.465	-74.004
6. Combined Comment and Revision Model Model 5 and Model 4 Predictors Combined	11.14 **	.712	.648	598.192		
Improvement from Model 5					.075	-4.731

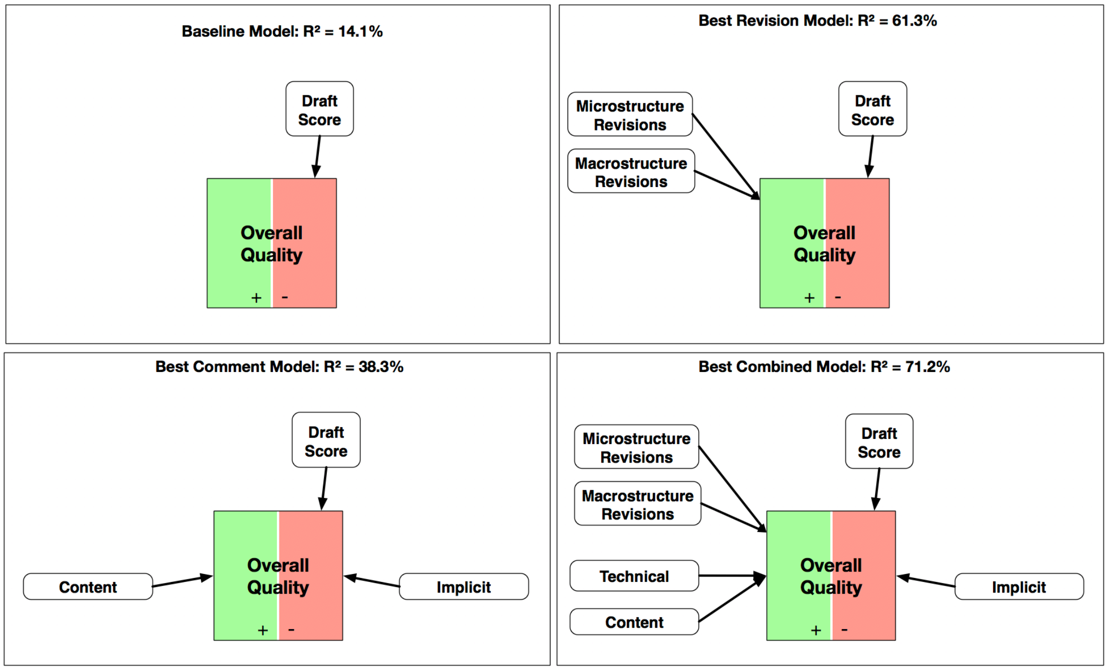


Figure 20: Graphical Comparison of Control, Direct, Indirect, and Combined Regression Models for Overall Quality Improvement

Chapter 5: Discussion

This study examined the use of peer review to enable a process-oriented approach to a writing assignment in a large first-year engineering course. Over the past 40 years, developments in writing studies research have emphasized the importance of both process-orientation – incorporation of scaffolding through feedback and revision – and situativity – recognizing that knowledge is best learned within the context it is to be applied (See Section 2.2). Engineering faculty, however, often view writing in ways that are at odds with both of these developments: treating writing as something to be taught elsewhere – such as a technical writing course offered in the English department, or when integrating writing into engineering courses doing so without incorporating feedback or revision (See Section 2.3). This mismatch is especially problematic, because improving engineering students' writing – both as a professional skill (See Section 2.1) and as a means to professional identity development (See Section 2.2.3) – is critical problem for engineering education.

Even if the aforementioned epistemic and pedagogical beliefs are overcome, a significant barrier remains: integrating writing assignments (especially those incorporating feedback and revision) into disciplinary courses is very resource-intensive in terms of instructor time. In large classes, these resource constraints could make incorporation of these assignments impossible. Recent research examining the use of peer feedback on writing assignments in disciplinary courses has shown that it can be as or more effective than instructor feedback (K. Cho & MacArthur, 2010; K. Cho, Schunn, & Charney, 2006; K. Cho, Schunn, & Wilson, 2006). Given 1) the importance of writing for engineering education, 2) the importance of situativity and process-orientation in contemporary writing pedagogy, 3) the resource barriers to integrating process-oriented writing assignments in engineering courses, and 4) the promising research on the effectiveness of peer review of writing in other disciplines, there is a need to better understand peer review within the context of engineering courses.

To address this need, I conducted an analysis of the practice and efficacy of peer feedback in a large first-year engineering course. Using data from a quasi-experimental intervention, where both the method of peer feedback instruction (either an in-class workshop or a handout) and the number of peer reviews participants engaged in were systematically manipulated, I used Qualitative Content Analysis techniques to examine how participants practiced peer review (i.e. the nature of reviewer feedback) and the efficacy of the peer review process (i.e. the nature of authorial revisions) in quality improvement between participants' draft and revised reports. In the following sections, I first discuss the major findings and situate them within the existing literature. Next, I make recommendations for engineering educators based on these findings. Finally, I discuss both the limitations of this research and how it will inform future work.

5.1. Discussion of Results

In this section, I discuss the results reported in Chapter 4 within the context of the study hypotheses and extant literature. In Chapter 3, I presented a conceptual model of the feedback and revision process based on the literature review in Chapter 2. In this model, represented at a high level in Figure 21 (see Figure 12 for a detailed model), the direct effects are expected to be through a sequence of stages:

1. feedback conditions (feedback source and feedback type) affect feedback features (as defined by the coding frame for feedback); then,
2. feedback features affect revision features (as defined by the coding frame for revisions); and finally,

3. revision features affect quality improvement.

In addition, this model allows the possibility of indirect effects on each stage resulting from prior stages (shown as gray dashed arrows in the figure). In the following subsections, I discuss results relevant to each of these stages in order. Both direct and indirect factors (where applicable) are discussed for each stage. To orient the reader, a figure displaying the relevant portion of the high-level conceptual model (Figure 21) is shown at the beginning of each subsection (Figure 22, Figure 23, Figure 24, and Figure 25). In these figures, dark solid arrows represent direct effects and gray dashed arrows represent indirect effects.

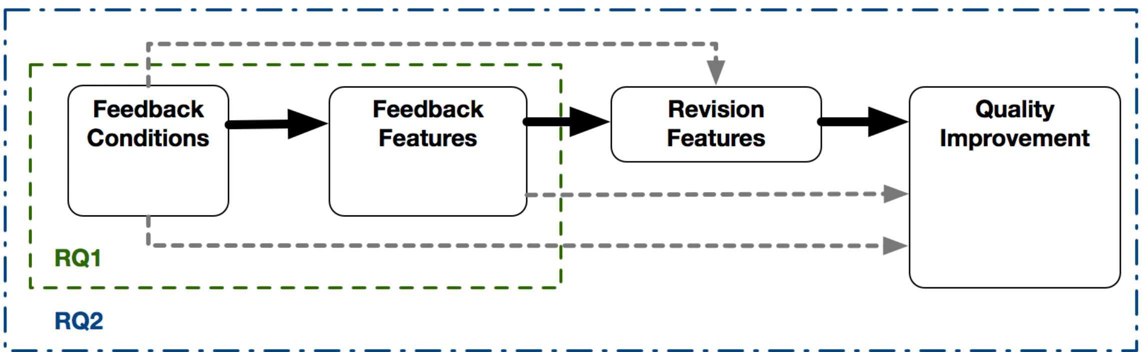


Figure 21: High-Level Conceptual Model of Feedback and Revision

There are nine major findings addressed in the discussion below:

1. Multiple-peer feedback was at least as effective as instructor feedback in terms of writing quality improvement.
2. For writing quality, multiple-peer feedback was more effective than instructor feedback.
3. In-class feedback instruction improved the practice of peer review versus handout-only.
4. Most of the improvements in the practice of peer review resulting from the in-class instruction were only statistically significant when a student reviewed multiple reports.
5. Multiple peer review was the only factor predicting corrections based on assignment rules.

6. Directive feedback was more helpful for quality improvement than anticipated.
7. There were no quality metrics for which readerly comments were a positive predictor.
8. Localization is important.
9. Macrostructure revisions, while important for quality improvement, were not predictable.

5.1.1. Quality Improvement from Peer vs Instructor Feedback

In order to validate the assumption that peer feedback is a valid substitute for instructor feedback, I compared the draft-to-revision changes in quality between students receiving instructor feedback to those receiving peer feedback. As shown in section 4.1, students who revised their drafts based on peer feedback did not statistically-significantly differ in quality improvement from those who revised based on instructor feedback overall. **Thus, the first major finding is that multiple-peer feedback was at least as effective as instructor feedback in terms of writing quality improvement**.

The exception was that both multiple-peer feedback groups showed statistically significant improvement in *writing quality* as compared to the instructor feedback group. **Thus, the second major finding is that for writing quality, multiple-peer feedback was more effective than instructor feedback**. It should be noted, however, that instructors were not explicitly trained on providing feedback; however, this lack of training makes the instructors representative of typical engineering faculty.

These findings are consistent with and extend existing literature and supports the conclusion that multiple peer feedback is a valid substitute for instructor feedback in this context. Several studies (K. Cho & MacArthur, 2010; K. Cho & Schunn, 2007) have compared quality improvement resulting from instructor, peer, and multi-peer feedback in undergraduate psychology courses with similar results: multiple-peer feedback results in the largest improvement in quality. Patchan et al. (2011) found a similar effect in an introductory physics course: students receiving peer feedback showed a statistically significant improvement in quality versus those receiving feedback from a teaching assistant. Although this result was expected, the present study is the first to empirically demonstrate this effect within the context of first-year engineering or, for that matter, any engineering course.

5.1.2. The Effect of Review Instruction on the Practice of Peer Review

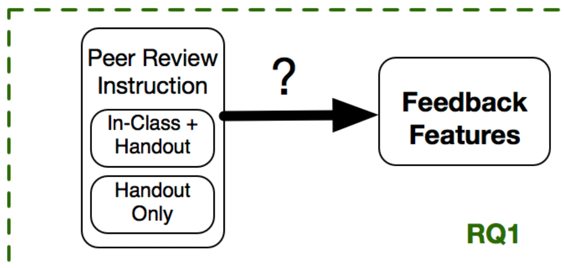


Figure 22: Effect of Peer Review Instruction on Feedback Features

The third major finding is that in-class instruction improved the practice of peer review versus handout-only. As reported in Chapter 4, analysis of the feedback corpus showed that in-class instruction appeared to positively impact the nature of the peer feedback by eliciting feedback more consistent with effective practices documented in writing research. In-class instruction improved peer feedback in several statistically significant ways: at the feedback item level of analysis, it predicted the use of higher-level modes and foci as well as more positive and qualified positive tone; at the paper level of analysis, in-class instruction increased quantity of comments per draft, more use of readerly comments, and more use of content-focused comments.

This finding is both consistent existing literature, as well as contributing a novel finding: while previous studies have examined both peer feedback training and number of peer reviewers, this is the first to report experimental results of the interaction of the two. This finding indicates that students should be trained in effective peer feedback practices. Several scholars have noted the importance of training students to provide peer feedback (Ozogul & Sullivan, 2007; Sluijsmans et al., 2002; Weaver, 1995); this instructor-provided scaffolding is crucial, as true peers lack the knowledge to intentionally work within their peer's Zone of Proximal Development when providing feedback (see 2.2.5). Of the existing literature on peer feedback training, Ozogul and Sullivan's (2007) study is the only one I identified that compared an in-class training intervention to an existing handout. Similar to the intervention developed for this study, they provided in-class instruction and practice using an existing evaluation rubric; findings indicated that after the in-class training peer evaluations more closely aligned with instructor evaluations.

It is important to remember that the both the in-class and handout-only groups in this study received the same raw information – the in-class workshop was a presentation of the material from the handout. The material differences were that in the workshop the writing teacher

1. explicitly laid out her train of thought while demonstrating her decisions regarding what feedback to give,
2. walked around the class three times to answer questions, and
3. ensured that students were able to successfully give a peer a comment using the inking feature on their tablet PC.

Cognitive Apprenticeship (Collins, 2006; Collins, Brown, & Holum, 1991) provides a possible explanation for the salience of these differences:

> In apprenticeship, learners can see the processes of work: […] they assist a tradesman as he crafts […]. But in schooling, the "practice" of problem solving, reading comprehension, and writing is not at all obvious-it is not necessarily observable to the student. In apprenticeship, the processes of the activity are visible. In schooling, the processes of thinking are often invisible to both the students and the teacher. Cognitive apprenticeship is a model of instruction that works to make thinking visible. (Collins et al., 1991, p. 1)

Several principles of the Cognitive Apprenticeship model align with the modeling and scaffolding discussed in section 2.4.2. As shown in Table 34, Cognitive Apprenticeship principles are divided into content, method, sequencing, and sociology. While the handout and workshop provided students with similar content, the handout only used two methods: scaffolding and indirect modeling through examples. The in-class workshop implemented several additional aspects of cognitive apprenticeship methods versus using the handout alone:

- students directly observed the cognitive processes of the feedback process activity (i.e. the workshop involved explicit rather than indirect modeling);
- the writing teacher observed the students performing the commenting and provided feedback (i.e. the workshop introduced coaching and extended scaffolding); and
- students observed how their neighbor performed the commenting task, comparing it to themselves (i.e. the workshop provided an opportunity for reflection).

In addition, the workshop may have made students more cognizant of the importance of the reviewing task by dedicating a class day to it, whereas students in the handout-only groups may have seen it as just another homework assignment.

Table 34: Principles of Cognitive Apprenticeship

Principle	Definition
Content	Types of knowledge required for expertise
Domain knowledge	subject matter specific concepts, facts, and procedures
Heuristic strategies	generally applicable techniques for accomplishing tasks
Control strategies	general approaches for directing one's solution process
Learning strategies	knowledge about how to learn new concepts, facts, and proceedures
Method	Ways to promote the development of expertise
Modeling	teacher performs a task so students can observe
Coaching	teacher observes and facilitates while students perform a task
Scaffolding	teacher provides supports to help the student perform a task
Articulation	teacher encourages students to verbalize their knowledge and thinking
Reflection	teacher enables students to compare their performance with others
Exploration	teacher invites students to pose and solve their own problems
Sequencing	Keys to ordering learning activities
Increasing complexity	meaningful tasks gradually increasing in difficulty
Increasing diversity	practice in a variety of situations to emphasize broad application
Global to local skills	focus on conceptualizing the whole task before executing the parts
Sociology	Social characteristics of learning environments
Situated learning	students learn in the context of working on realistic tasks
Community of practice	communication about different ways to accomplish meaningful tasks
Intrinsic motivation	students set personal goals to seek skills and solutions
Cooperation	students work together to accomplish their goals

(Collins, 2006, p. 50)

5.1.3. The Effect of Review Type on the Practice of Peer Review

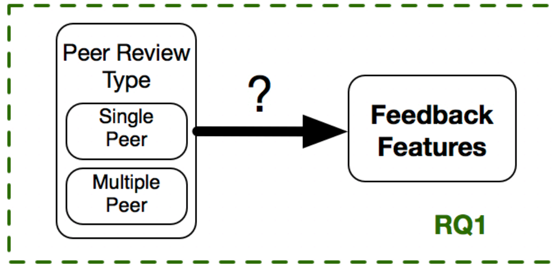


Figure 23: Effect of Peer Review Type on Feedback Features

In addition to *feedback instruction*, *feedback type* (the number of reports reviewed) was also manipulated in the quasi-experimental design. **The fourth major finding is that most of the improvements in the practice of peer review resulting from the in-class workshop were only statistically significant when a student reviewed multiple reports**; only the increases in overall number of comments and in comments with qualified tone remained significant when students reviewed only a single peer's report. This finding is a new addition to the literature, as the relationship between feedback instruction and number of reviews has not previously been examined.

Cognitive Apprenticeship also offers a potential explanation for this finding. The assignment of multiple reviews entails *increasing diversity*. This involves

> the construction of a sequence of tasks in which a wider and wider variety of strategies or skills are required [practicing] a new strategy or skill repeatedly in a sequence of (increasingly complex) tasks, […] requiring a diversity of skills and strategies […] so that the student learns to distinguish the conditions under which they do (and do not) apply [causing students to] acquire a richer net of contextual associations […] for use with unfamiliar or novel problems. (Collins et al., 1991, p. 15).

By applying the feedback skills in the context of more than one report, reviewers may be forming more robust understandings of, and therefore skills in, giving meaningful feedback. At the same time, while reviewing multiple reports and the in-class workshop both improved feedback, those improvements did not always translate to increases in draft-to-revision quality in the ways expected.

5.1.4. Factors Influencing Revision Features

The discussion in preceding sections addressed the first research question (regarding the effect of peer review instruction and type on feedback features); in the remaining sections, discussion shifts to addressing the second research question (regarding the factors that influenced revision features and quality improvement). In this section, I discuss the factors influencing revision features. As shown in Figure 24, these factors include both direct (i.e. which *feedback features* prompt certain revision features), and indirect (i.e. whether *feedback conditions* influenced revision features). The relationship between feedback conditions and revision features is considered indirect because conditions were primarily intended to influence the practice of providing feedback, not revising (e.g. the workshop and handout taught students how to provide feedback but did not address using that feedback to make revisions).

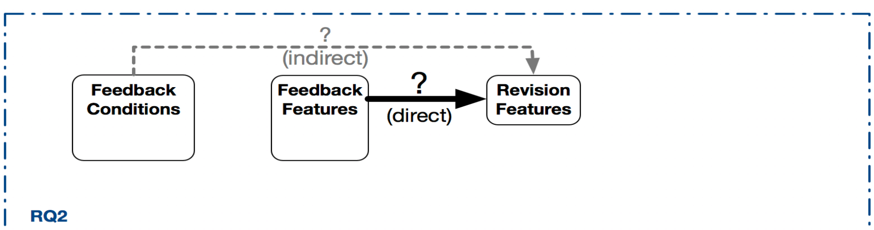


Figure 24: Factors Influencing Revision Features

In addition to the direct relationship between feedback features and revision features, I also hypothesized that there would be an independent relationship between the feedback conditions and revision features. Feedback instruction did not have any statistically-significant independent effect on revision features. However, multiple peer feedback did have a statistically-significant effect on assignment rule corrections and on microstructure revisions. As shown in Table 35, students in the multiple peer condition were more likely to make both assignment rule corrections and microstructure revisions. **Because this was the only factor that predicted assignment rule corrections, the fifth major finding is that multiple peer review was the only factor predicting corrections based on assignment rules**. Because there was no significant relationship between feedback features and assignment rule corrections, the result suggests that reviewing multiple peer reports (as opposed to receiving feedback from a peer who has reviewed multiple reports) makes students more likely to make this type of correction.

Table 35: Significant Relationships Between Review Type and Revision Features

	Single Peer		Multiple Peer	
Revision Feature	Absent	Present	Absent	Present
Assignment Rule Corrections	78%	22%	59%	41%
Microstructure Revisions	53%	47%	33%	67%

Assignment rule corrections are those based on the specific conventions – e.g. citation style, inclusion of abstract and keywords, and use of figures to illustrate major points – of the genre of the engineering research paper. The learning of these disciplinary genre conventions has been identified as an important factor supporting integration of writing assignments in disciplinary courses as opposed to separate writing courses. One concern with replacing instructor feedback with peer feedback is that peer reviewers are novices and therefore lack the perspective of experienced writers to understand the purpose and audience of their writing (Sommers, 1980). Indeed, in a study of feedback given by biology instructors on WID assignments, Szymanski (2014) found that the most effective feedback was given by instructors who viewed the purpose of feedback in just this way: to train students to write professionally in their discipline by attending to the particularities of their disciplinary genres in their responses. Given their lack of experience, it seems a reasonable concern that peer reviewers may not be as capable of providing this type of feedback.

On the other hand, the increase in assignment rule corrections under the multiple peer condition suggests reviewing multiple peer reports may help to compensate for this problem. A possible explanation for the increase is that peers were exposed to several examples of writing that varied in the way the authors operationalized their understanding of the genre. This reading of multiple examples has been found to increase novice writers' ability to revise by improving their ability to view their own writing through the lens of a reader (Holliway, 2004; Traxler & Gernsbacher, 1993), which could explain why reviewing multiple reports impacted this revision type regardless of feedback received.

5.1.5. Factors Influencing Quality Improvement

In this final discussion section, I discuss the factors influencing quality improvement. As shown in Figure 25, these factors include both direct (i.e. the extent to which certain *revision features* result in quality improvement), and indirect (i.e. the extent to which *feedback conditions* and/or certain *feedback features* result in quality improvement). Similar to the relationship between feedback conditions and revision features in the previous section, feedback conditions and feedback features both have an indirect relationship with quality improvement because 1) feedback conditions were primarily intended to influence the practice of providing feedback, not revising (e.g. the workshop and handout taught students how to provide feedback but did not address using that feedback to make revisions) and 2) feedback was expected to influence quality improvement *by eliciting certain types of revisions* rather than directly altering quality.

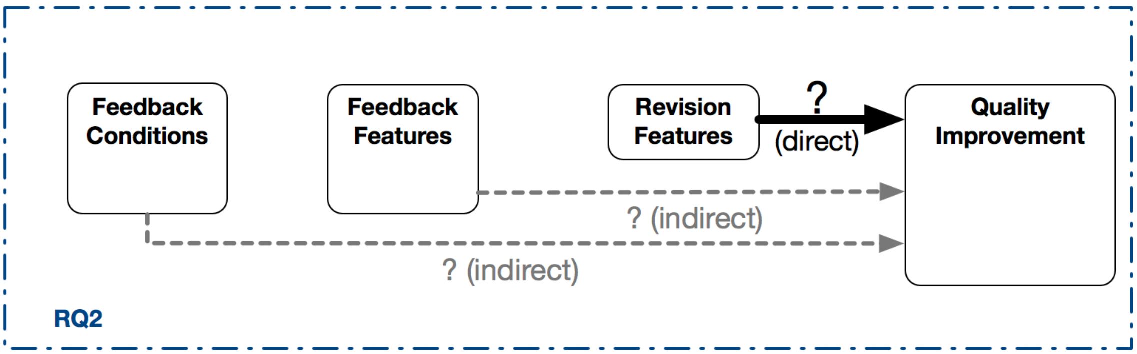


Figure 25: Factors Influencing Quality Improvement

Prior work analyzing quality improvement based on peer feedback (e.g. K. Cho & MacArthur, 2010; K. Cho & Schunn, 2007; Nelson & Schunn, 2009; Patchan & Schunn, 2015, 2016) has generally reported only overall changes in draft-revision quality whereas this study considered various dimensions of quality (mechanics, writing, technical, and overall). Additionally, as reported in Chapter 4, there were many factors influencing the various dimension of quality. For these reasons, I organize this section first by general discussion of all models, then by individual results with particular salience to this study and to the extant literature.

Results Across all Models

For each quality improvement metric, I ran a series of regression models (described in section 4.3.3) to determine which factors best predicted each metric. The baseline (control) predictors included the feedback conditions (feedback instruction and feedback type) as well as the initial draft score. Feedback instruction was never a significant predictor on its own and was therefore removed from analysis. Initial draft score was included in all models regardless of significance because drafts with a lower score by definition have more room for improvement (this fact is also why it always is a negative predictor in the models). When comparing all models, three salient results stand out: (1) revision features are a moderately strong predictor of quality improvement, (2) revision features were a stronger predictor of improvement than feedback features, and (3) most interactions between feedback features that were significant in the comment-only models lacked significance when combined with revision features.

For all metrics, revision features are a moderately strong predictor of quality improvement; this was consistent with expectations because quality improvement is by-definition a result of making revisions. This high correlation does, however, provide evidence for the validity of the coding frame for revisions. Feedback features were always a moderate (weaker than revision features) predictor, indicating that revisions are elicited by factors beyond just feedback features – a smaller gap between the percent of variation explained by revision features and feedback features would indicate a more direct correspondence between feedback and revision.

Perhaps the most interesting result that can be seen by comparing the models is that the interaction effects between feedback features seen in the comment-only models usually lack significance when combined with revision features. This finding suggests that many revisions are prompted by a complex mix of feedback features rather than certain features alone. This also indicates that the feedback features that remain significant in the combined models may be particularly salient for quality improvement, as they contribute to improvement directly rather than simply prompting a particular type of revision.

Rethinking Mode

In the writing studies literature, the level of control the reviewer exerts over the author's text – operationalized as feedback mode in this study – is considered particularly important; comments taking the position of the reader and facilitative comments are generally encouraged over directive comments and edits. (Brannon & Knoblauch, 1982; Knoblauch & Brannon, 1981; Sommers, 1982; Straub & Lunsford, 1996a). Based on this literature, the peer review instruction used in this study encouraged students to use the readerly and coaching modes. However, the results of this research only partially support, and in some ways contradict, this advice.

Comments in the authoritative mode, particularly when combined with lack of localization (i.e. are not specific), *did* have a negative impact on both mechanics quality and writing quality. Edits – a highly localized and authoritative type of feedback – on the other hand, had a positive effect on both mechanics quality and writing quality. Coaching comments – where the reviewer *recommends* a change or *asks* the writer a question rather than *telling* the writer what to do, i.e. the quintessential facilitative mode of comment – also had an unexpected negative effect: coaching comments with a technical focus were associated with a decrease in technical quality.

These results contradict the general recommendations for feedback from the writing studies literature. **The sixth major finding, then, is that directive feedback was more helpful for quality improvement than anticipated.** However, they are consistent with Patton and Smith Taylor's (2013) finding that directive commentary may actually be desirable in the context of engineering writing. They argue that the kind of writing that engineers do "typically functions in more tightly constrained genre systems than does writing in the humanities" and that "directive comments [are] necessary to foreground the factual and situational constraints." The technical quality dimension in this study was centered on exactly these constraints; the fact that the use of the coaching (non-directive) mode only had a negative impact on technical quality further supports this interpretation. However, interpretation of this finding should be also consider that there may be learning effects from non-directive comments that would require a longitudinal study to explore.

Readerly Comments Were Ineffective

It was also expected based on the writing feedback literature (g.g. Smith Taylor, 2007; Straub & Lunsford, 1996a) that readerly comments – where the reviewer takes the position of audience, provides outside information, or summarizes without evaluating – would be particularly important for quality improvement. **The results of this study, however, indicated that there were no quality metrics for which readerly comments were a positive predictor, the seventh major finding.** In fact, readerly comments with qualified tone were a negative predictor of both mechanics and writing quality improvement. While the ability to view one's writing from the perspective of a reader *is* particularly salient for effective revision, and providing readerly comments is intended to help writers gain this perspective, novice writers are less likely to be able to do so than experienced writers (Sommers, 1980).

This phenomenon is a possible explanation for the most confounding result of this study. As discussed previously in section 5.1.3:

- among students reviewing a single peer report (treatments C and D), draft-revision quality improvement was higher in every dimension when participating in the in-class workshop; and

- among students participating in the in-class workshop (treatments D and F), draft-revision quality improvement was higher in every dimension except Technical when reviewing multiple reports; however,
- among students reviewing multiple peer reports (treatments E and F), draft-revision quality improvement was highest among students who *did not* participate in the in-class workshop *even though feedback was "better" in group F.*

The feedback provided by students in this groups was "better" in that it included a higher number of readerly, qualified, positive, and implicit comments. In the regression analysis, these were precisely the factors that negatively impacted quality improvement. This result could indicate the in-class workshop increased students' use of these feedback features but that, as novices, they were unable to do so effectively. Although this result shows that peers in this study were unable to provide effective feedback from the perspective of a reader, previous studies supporting the use of multi-peer feedback had students read and respond to between four (e.g. Patchan & Schunn, 2016) and six (e.g. K. Cho & MacArthur, 2010) peer papers and also provided far more structured reading tasks, which have also been found to increase novice writers' ability to revise (Holliway, 2004; Traxler & Gernsbacher, 1993).

Localization is Important

Finally, results demonstrate the importance of feedback localization. Which describes *where* a comment refers to in the draft; when a comment is localized, it "explicitly refers to the location of the issue" (Patchan et al., 2016, p. 1100). While localization is not a component of Straub and Lunsford's feedback taxonomy, it is a particularly salient difference between *end comments* and *marginalia:* a comment such as "could use more references" could be implicitly localized due to its proximity to a paragraph, explicitly localized by an arrow pointing at a particular area, or non-localized if it were an end comment.

The regression analysis of quality improvement showed that: (1) feedback without explicit localization appears to have a negative effect on quality improvement, especially when using an authoritative mode (for writing and mechanics quality), but (2) non-localized comments have a positive effect on writing quality. **This eighth major finding is consistent with and expands upon existing evidence that the effect of localization is mixed, which showed that localized comments are more likely to be understood** (Nelson & Schunn, 2009) and more likely to be implemented (Nelson & Schunn, 2009; Patchan et al., 2016) than non-localized comments, but can reduce revision quality (Patchan et al., 2016).

5.2. Contributions to the Literature and Implications for Practice

For this study there were nine main findings, of which eight were discussed in detail above. **The ninth major finding, that macrostructure revisions, while important for quality improvement, were not predictable**, is discussed below. Table 36 below summarizes these nine major findings and their relationship to the extent literature. In this section, I discuss how these findings contribute to the extant literature in the fields of engineering education and writing studies and make recommendations for practice.

Table 36: Summary of Findings

Finding	Elaboration	Relationship to Current Literature
1. Multiple-peer feedback may be at *least as effective* as instructor feedback	There was no statistically-significant difference in overall quality improvement between students who revised their drafts based on peer feedback vs. those who revised based on instructor feedback; however this study did not consider outcomes beyond one assignment.	Consistent with current literature and expands to domain of engineering
2. For writing quality, multiple-peer feedback may be *more effective* than instructor feedback	There was a single quality dimension where there *was* a statistically-significant difference resulting from peer feedback vs instructor feedback: both multiple-peer feedback groups showed statistically-significant improvement in writing quality as compared to the instructor feedback group.	Consistent with current literature and expands to domain of engineering
3. In-class feedback instruction may improve the practice of peer review versus handout-only, however 4. This improvement is ionly significant for multi-peer groups	In-class feedback instruction improved the practice of peer review versus handout-only in terms of feedback features identified in the literature. However, this improvement was only statistically-significant when a student reviewed multiple reports.	New addition to literature: experimental comparisions of peer review instruction and interaction with number of peers not previously reported
5. Multiple-peer review was the only factor predicting corrections based on assignment rules.	Reviewing multiple reports was the only factor predicting corrections based on assignment rules, which may be important for learning disciplinary genre conventions. This was possibly due to the act of reviewing multiple peer reports, as opposed to receiving feedback from multiple peers.	New addition to literature
6. Directive feedback was more helpful for quality improvement than anticipated.	While the writing studies literature recommends against the use of directive feedbask modes, the results of this study showed that directive feedback was a significant predictor of multiple dimensions os quality improvement.	Unexpected based on writing studies literature
7. Readerly feedback was ineffective	There were no quality metrics for which readerly comments were a positive predictor, but readerly comments with qualified tone were a negative predictor of both mechanics and writing quality improvement.	Contradicts current literature
8. Localization is an important feedback feature	Feedback localization (or lack thereof), a dimension absent from Straub and Lunsford's taxonomy, had several significant impacts on quality improvement	Consistent with current literature and expands to domain of engineering
9. Macrostructure revisions were important for quality improvement, but not predictable.	There were no statistically-significant relationships between macrostructure revisions and any feedback features. However, macrostructure revisions *were* a statistically-significant positive predictor for overall quality improvement.	Unexpected based on writing studies literature

First, this study supports the extension of previous research to the context of engineering. Previous research on peer feedback has shown it to be an effective substitute for instructor feedback in non-engineering contexts (K. Cho & MacArthur, 2010; K. Cho, Schunn, & Wilson, 2006; Patchan, 2011). Findings 1 and 2 are evidence that these results are also applicable in a first-year engineering context. This suggests that several problematic status-quo aspects of engineering communication curricula – the typical separation of communication and content learning (Leydens & Schneider, 2009, p. 260), adoption of a "transmission model" of communication (Paretti et al., 2014, p. 614; Leydens, 2008) and challenges for improving these practices due to resource constraints such as limited instructor time and large student-faculty ratios (Matusovich et al., 2012) – can be addressed by using peer feedback to enable the adoption of process-oriented writing assignments in engineering courses. Addressing both the separation of "learning to write" from "learning to engineer" (i.e. treating engineering writing as a *situated activity*) and the common product-centric perspective (Swarts & Odell, 2001; Zhu, 2004) could improve both skill and identity development for engineering students.

Second, Findings 2, 4, and 5 underscore the importance of using feedback from multiple peers rather than just a single peer. Previous studies (e.g. K. Cho & MacArthur, 2010; K. Cho & Schunn, 2007; K. Cho, Schunn, & Wilson, 2006) have shown that multiple-peer feedback is more effective than single-peer both for both providing and implementing feedback. While the results of this study showed that feedback from two peers did result in the provision of "better" feedback, the change in writing quality was not sufficiently improved, suggesting that an increased number of peers – possibly the 4-6 peers used in prior work – should be used for multiple-peer feedback.

Finally, while scholars have noted the importance of training students to provide peer feedback (Ozogul & Sullivan, 2007; Sluijsmans et al., 2002; Weaver, 1995), those studies did not examine differences in efficacy for alternative types of training. This study showed (Finding 3) that training using an in-class workshop is preferable to using only a handout with respect to peers following the training recommendations. However, Findings 6 and 7 also indicate that some general recommendations for writing feedback – notably an emphasis on readerly feedback and discouraging directive feedback – that are applicable to feedback from experts may not be applicable for peers. Directive feedback is often discouraged and said to appropriate student texts and lead to primarily surface-level revisions (Brannon & Knoblauch, 1982; Knoblauch & Brannon, 1981; Sommers, 1982; Straub & Lunsford, 1996a), the results of this study showed that directive feedback was associated with significant improvements in writing quality; this finding supports Patton and Smith Taylor's (2013) position that differences between writing in engineering and the humanities mean that "directive comments [are] necessary to foreground the factual and situational constraints" present in engineering genres. While writing studies literature encourages feedback that takes the position of the reader (Straub & Lunsford, 1996a), and previous studies have supported the applicability of this recommendation to peer feedback (e.g. K. Cho & MacArthur, 2010), this study found that readerly peer feedback actually reduced revision quality. In previous studies of quality improvement from peer feedback, students received comments from 4-6 peers, whereas in this study feedback was limited to two peers; this result indicates that for *peer* feedback, response from more than two peers is necessary to effectively convey a reader's perspective. Finding 8 confirms the complex role of feedback localization, suggesting that a mix of both explicitly-localized feedback and general feedback on the entire paper is beneficial.

5.3. Recommendations for Practitioners

The findings discussed in this chapter support the following recommendations for classroom practice:

1. Writing assignments in disciplinary engineering courses should include process-oriented scaffolding; resource constraints can effectively be addressed through the use of peer feedback activities.
2. Peer feedback activities should use multiple peer review rather than single peer in order to improve their efficacy but should include feedback from more than two peers – research from other domains suggests 4-6 peers but given that this study did not consider more than two peers this specific number is a tentative recommendation.
3. Student peers should be trained in the practice of giving feedback; using an in-class peer feedback workshop can improve adherence to training recommendations over using only a handout.
4. Training should encourage a mix of directive and facilitative feedback.

5.4. Limitations and Future Work

The findings presented in this dissertation have provided valuable contributions to support the project of improving writing pedagogy in engineering. Most significantly, they support the use of directed peer review of writing in the context of engineering courses by replicating the results of prior studies had demonstrated the efficacy of peer review in other disciplinary courses. They also demonstrate the value of in-class instruction in peer review in helping students adhere to effective peer review strategies, even as they help extend research on what constitutes effective peer review practice in engineering. However, the results also point to the need for further research in this domain. In this section, I suggest possible directions for this future work.

The most significant limitation of this study was that it was limited to a single assignment in a single course at a single institution. As such, the results of this dissertation only unequivocally demonstrate that peer review was effective in this specific context, though as noted earlier the context can be considered representative of a number of other similarly situated first-year engineering programs. There is a much room to expand the scope of a similar study in order to generalize the findings: “upward” expansion would go beyond first-year engineering and explore the use of peer feedback in upper-level courses as well; “outward” expansion would go beyond the large public R1 to a more diverse set of institutions. Further generalization of the results would be beneficial to the engineering education community, but while the methods used in this study were effective to answer the research questions, the amount of data could get overwhelming very quickly if the scope were significantly expanded; any future research should capture feedback in a computer-readable format to facilitate larger-scale analysis.

A methodological limitation of this work is that because revisions were coded at the paper level, there was no direct mapping between specific feedback and specific revisions. This could be a reason for the weak correlations found between feedback and revision features. Other studies (e.g. Nelson & Schunn, 2009; Patchan et al., 2016) have examined the direct relationship between specific feedback comments and particular revisions. In those studies, peer feedback was limited to end comments; the inclusion of marginalia in the data set for this study would make mapping feedback to revisions more challenging but doing so might enhance the explanatory power of analysis and would enable more direct comparison to those results.

The results of this study indicated that peer review was at least as effective as instructor review in terms of draft-revision quality improvement. However, the goal is not really to produce a better report but to teach students to be better writers. It is possible, even likely, that instructor feedback has some other benefits to learning that are not captured in measuring quality improvement. Future work comparing instructor and peer feedback should explicitly consider not just ways in which peer feedback *can* replace instructor feedback but also explore whether there are ways in which it cannot.

Finally, although the findings from this study question the assumption that directive commentary is to be avoided, a particular concern with directive modes is that the writer simply implements such comments and that they do not learn from this simple implementation. From a constructivist perspective (see 2.2.4 for an overview) a learner must encounter a *perturbation*, or a new experience that does not meet the expectations of their existing knowledge schemata in order to assimilate the new knowledge into the schema (Rowell, 1989). As discussed in Chapter 2, one reason the use of facilitative comments is encouraged is because they are congruent with this constructivist perspective; whereas a directive comment may simply be implemented without thought (and therefore not cause a perturbation), facilitative comments by definition ask the writer to think more than superficially about the comment and its implementation. So, while directive comments did result in draft-revision quality improvement, it is far from certain whether this will transfer to future writing tasks. A future study should address this gap by collecting and analyzing similar artifacts as the present study, but over several writing assignments for the same cohort of students.

5.5. Concluding Remarks

That an ability to communicate effectively – and concerns about graduates' ability to do so – has been recognized from the earliest days of formalized engineering education (Kynell-Hunt, 2000). Despite this, contemporary research continues to demonstrate both the centrality of communication in engineering work and that graduates are lacking in their ability to communicate. Several factors contribute to this status quo, including an engineering culture that often divides skills into a dichotomy of *hard* and *soft*, with "soft" skills such as communication hard seen as separate from – and lesser than – the "hard" engineering skills (Leydens & Schneider, 2009, p. 260,261; see also Anderson et al., 2010; Trevelyan, 2010). This viewpoint, along with resource constraints, have entrenched a curricular separation of communication in many engineering curricula despite compelling arguments that communication skills are best learnt in context (Paretti et al., 2014). Even when communication is integrated, instructors often do not include revision tasks, despite its importance in the development of writing ability (Swarts & Odell, 2001; Zhu, 2004).

This study examined how these issues can be addressed through the use of peer feedback in engineering writing assignments. The results showed that peer feedback can be an effective scaffold and improve engineering students' writing while working within these common resource constraints, and in doing so overcome constraints on course and instructor time to support more effective approaches to writing instruction. These results also underscore the importance of using feedback from multiple peers rather than just a single peer and training peers to provide quality feedback. However, the results also indicate that some general recommendations for writing feedback – notably an emphasis on readerly feedback and discouraging directive feedback – that are applicable to feedback from experts may not be applicable for peers in engineering. These findings thus help nuance generalized findings from writing studies within a specific disciplinary context, suggesting that effective feedback may not be generic across disciplines or academic years, and in turn deepens our understanding of situated writing pedagogies.

By adopting the recommendations from this dissertation, engineering educators can both improve their writing pedagogy and more explicitly frame writing as integral to the work of engineering practice.

References

ABET. (2010, October). Criteria for Accrediting Engineering Programs: Effective for Evaluations During the 2011-2012 Accreditation Cycle. Retrieved November 8, 2011, from http://www.abet.org/uploadedFiles/Accreditation/Accreditation_Process/Accreditation_Documents/Current/abet-eac-criteria-2011-2012.pdf

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Allie, S., Armien, M. N., Burgoyne, N., Case, J. M., Collier-Reed, B. I., Craig, T. S., … Jacobs, C. (2009). Learning as acquiring a discursive identity through participation in a community: Improving student learning in engineering education. *European Journal of Engineering Education*, *34*(4), 359–367.

American Society for Engineering Education. (2017). *Engineering by the Numbers*.

Anderson, K. J. B., Courter, S. S., McGlamery, T., Nathans-Kelly, T. M., & Nicometo, C. G. (2010). Understanding engineering work and identity: a cross-case analysis of engineers within six firms. *Engineering Studies*, *2*(3), 153–174. https://doi.org/10.1080/19378629.2010.519772

Artemeva, N., & Logie, S. (2003). The Teaching and Practice of Peer Feedback in the Professional Communication Classroom. *Language & Learning Across the Disciplines*, *6*, 1.

Artemeva, N., Logie, S., & St- Martin, J. (1999). From page to stage: How theories of genre and situated learning help introduce engineering students to discipline- specific communication. *Technical Communication Quarterly*, *8*(3), 301–316. https://doi.org/10.1080/10572259909364670

Bazerman, C. (2015). What do sociocultural studies of writing tell us about learning to write. *Handbook of Writing Research*, 24–40.

Bielenberg, B. (2011). Academic Literacy and Engineering Education: Development through Cornerstone Design (pp. 22.130.1-22.130.18). Presented at the 2011 ASEE Annual Conference & Exposition. Retrieved from https://peer.asee.org/academic-literacy-and-engineering-education-development-through-cornerstone-design

Borrego, M., Foster, M. J., & Froyd, J. E. (2014). Systematic Literature Reviews in Engineering Education and Other Developing Interdisciplinary Fields: Systematic Literature Reviews in Engineering Education. *Journal of Engineering Education*, *103*(1), 45–76. https://doi.org/10.1002/jee.20038

Brannon, L., & Knoblauch, C. H. (1982). On Students' Rights to Their Own Texts: A Model of Teacher Response. *College Composition and Communication*, *33*(2), 157–166. https://doi.org/10.2307/357623

Bransford, J. D., Brown, A. L., & Cocking, R. (Eds.). (2000). *How people learn: brain, mind, experience, and school: expanded edition* (Vol. 24). Washington, DC: National Academies Press.

British Computer Society. (2012). *Guidelines on course accreditation* (pp. 1–54).

Broadhead, G. J. (1999). Addressing multiple goals for engineering writing: The role of course specific websites. *Language and Learning Across the Disciplines*, *3*(2), 19–43.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*(1), 32–42. https://doi.org/10.3102/0013189X018001032

Cech, E. A. (2013). The (Mis)Framing of Social Justice: Why Ideologies of Depoliticization and Meritocracy Hinder Engineers' Ability to Think About Social Injustices. In J. Lucena

(Ed.), *Engineering Education for Social Justice* (Vol. 10, pp. 67–84). Dordrecht: Springer Netherlands. Retrieved from http://link.springer.com/10.1007/978-94-007-6350-0

Cech, E. A., & Waidzunas, T. J. (2011). Navigating the heteronormativity of engineering: the experiences of lesbian, gay, and bisexual students. *Engineering Studies*, *3*(1), 1–24. https://doi.org/10.1080/19378629.2010.545065

Chandrasegaran, A. (2008). What does teaching writing as a process really mean? Retrieved from https://repository.nie.edu.sg//handle/10497/15284

Chatburn, G. R. (1907). A combined cultural and technical engineering course. In *Proceedings of the Society for the Promotion of Engineering Education* (Vol. 15, pp. 222–29).

Chi, M. T. (1996). Constructing self-explanations and scaffolded explanations in tutoring. *Applied Cognitive Psychology*, *10*(7), 33–49.

Cho, J., & Lee, E.-H. (2014). Reducing Confusion about Grounded Theory and Qualitative Content Analysis: Similarities and Differences. *The Qualitative Report*, *19*(32), 1–20.

Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction*, *20*(4), 328–338.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education*, *48*(3), 409–426. https://doi.org/10.1016/j.compedu.2005.02.004

Cho, K., Schunn, C. D., & Charney, D. (2006). Commenting on Writing: Typology and Perceived Helpfulness of Comments from Novice Peer Reviewers and Subject Matter Experts. *Written Communication*, *23*(3), 260–294.

Cho, K., Schunn, C. D., & Wilson, R. W. (2006). Validity and reliability of scaffolded peer assessment of writing from instructor and student perspectives. *Journal of Educational Psychology*, *98*(4), 891–901. https://doi.org/10.1037/0022-0663.98.4.891

Cho, Y. H., & Cho, K. (2010). Peer reviewers learn from giving comments. *Instructional Science*, *39*(5), 629–643. https://doi.org/10.1007/s11251-010-9146-1

Collins, A. (2006). Cognitive Apprenticeship. In R. K. Sawyer (Ed.) (1st ed.). Cambridge University Press.

Collins, A., Brown, J. S., & Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator*, *6*(11), 38–46.

Connors, R. (1982). The Rise of Technical Writing Instruction in America. *Journal of Technical Writing and Communication*, *12*(4), 329–52.

Connors, R., & Lunsford, A. A. (1993). Teachers' Rhetorical Comments on Student Papers. *College Composition and Communication*, *44*(2), 200–223. https://doi.org/10.2307/358839

Covill, A. E. (2010). Comparing Peer Review and Self-Review as Ways to Improve College Students' Writing. *Journal of Literacy Research*, *42*(2), 199–226. https://doi.org/10.1080/10862961003796207

Creswell, J. W. (2009). *Research design: qualitative, quantitative, and mixed methods approaches*. SAGE. Retrieved from http://books.google.com/books?id=bttwENORfhgC

Daiker, D. A. (1989). Learning to Praise. In *Writing and response: theory, practice, and research*. Urbana, Ill: National Council of Teachers of English.

DiPardo, A., & Freedman, S. W. (1988). Peer Response Groups in the Writing Classroom: Theoretic Foundations and New Directions. *Review of Educational Research*, *58*(2), 119–149. https://doi.org/10.2307/1170332

Donnell, J., Aller, B. M., Alley, M., & Kedrowicz, A. A. (2011). Why Industry Says That Engineering Graduates Have Poor Communication Skills: What the Literature Says. In *Proceedings of the 2011 American Society for Engineering Education Annual Conference & Exposition*. Vancouver, BC.

Dragga, S. (1988). The effects of praiseworthy grading on students and teachers. *Journal of Teaching Writing*, *7*(1), 41–50.

Drisko, J., & Maschi, T. (2015). *Content Analysis*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190215491.001.0001

Dudley, J. W. (1971). Writing Skills of Engineering and Science Students. *IEEE Transactions on Engineering Writing and Speech*, *14*(2), 42–46. https://doi.org/10.1109/TEWS.1971.4322468

Earle, S. C. (1911). English in the engineering school at Tufts College. In *Proceedings of the Society for the Promotion of Engineering Education* (Vol. 19, pp. 33–47). Pittsburgh, PA.

Ekoniak. (2014). *Teaching Peer Review of Writing in a Large First-Year Electrical and Computer Engineering Class: Comparison of Two Methods*. Virginia Tech, Blacksburg, VA.

Ekoniak, M., Scanlon, M. J., & Mohammadi-Aragh, M. J. (2013). Improving student writing through multiple peer feedback. In *2013 IEEE Frontiers in Education Conference (FIE)* (pp. 626–628). https://doi.org/10.1109/FIE.2013.6684901

Elbow, P. (1973). Writing without Teachers. Retrieved from http://eric.ed.gov/?id=ED078431

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *Sage Open*, *4*(1), 2158244014522633.

Evans, D. L., Beakley, G. C., Crouch, P. E., & Yamaguchi, G. T. (1993). Attributes of Engineering Graduates and Their Impact on Curriculum Design. *Journal of Engineering Education*, *82*(4), 203–211. https://doi.org/10.1002/j.2168-9830.1993.tb01075.x

Faigley, L., & Witte, S. (1981). Analyzing Revision. *College Composition and Communication*, *32*(4), 400–414. https://doi.org/10.2307/356602

Flaherty, C. (2015, June 26). Watered-Down Gen Ed for Engineers? Retrieved from https://www.insidehighered.com/news/2015/06/26/faculty-members-criticize-proposed-changes-gen-ed-accreditation-standards-engineers

Flower, L. S., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387.

Ford, J. D. (2004). Knowledge transfer across disciplines: tracking rhetorical strategies from a technical communication classroom to an engineering classroom. *IEEE Transactions on Professional Communication*, *47*(4), 301–315. https://doi.org/10.1109/TPC.2004.840486

Freedman, S. W., Greenleaf, C., & Sperling, M. (1987). *Response to student writing* (No. 23). Urbana, Ill: National Council of Teachers of English. Retrieved from http://www.jstor.org/stable/pdf/40171143.pdf

Garrison, R. H. (1974). One-to-one: Tutorial instruction in freshman composition. *New Directions for Community Colleges*, *1974*(5), 55–84. https://doi.org/10.1002/cc.36819740509

Gee, T. C. (1972). Students' Responses to Teacher Comments. *Research in the Teaching of English*, *6*(2), 212–221.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, *20*(4), 304–315. https://doi.org/10.1016/j.learninstruc.2009.08.007

Gielen, S., Tops, L., Dochy, F., Onghena, P., & Smeets, S. (2010). A comparative study of peer and teacher feedback and of various peer feedback forms in a secondary school writing curriculum. *British Educational Research Journal*, *36*(1), 143–162. https://doi.org/10.1080/01411920902894070

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112. https://doi.org/10.1016/j.nedt.2003.10.001

Greeno, G. J. (2005). Learning in Activity. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 79–96). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511816833.007

Greeno, J. G., Collins, A., & Resnick, L. B. (1996). Cognition and learning. In D. Berliner & R. Calfee (Eds.) (Vol. 77, pp. 15–46). New York: Macmillan.

Haaga, D. A. F. (1993). Peer Review of Term Papers in Graduate Psychology Courses. *Teaching of Psychology*, *20*(1), 28–32. https://doi.org/10.1207/s15328023top2001_5

Hayes, J. R., Flower, L., Schriver, K. A., Stratman, J., Carey, L., & others. (1987). Cognitive processes in revision. *Advances in Applied Psycholinguistics*, *2*, 176–240.

Hofer, B. K., & Pintrich, P. R. (1997). The Development of Epistemological Theories: Beliefs About Knowledge and Knowing and Their Relation to Learning. *Review of Educational Research*, *67*(1), 88–140. https://doi.org/10.3102/00346543067001088

Holliway, D. R. (2004). Through the eyes of my reader: A strategy for improving audience perspective in children's descriptive writing. *Journal of Research in Childhood Education*, *18*(4), 334–349.

IEEE. (2004). *Curriculum Guidelines for Undergraduate Degree Programs in Computer Engineering*.

Jensen, W., & Fischer, B. (2005). Teaching Technical Writing through Student Peer-Evaluation. *Journal of Technical Writing and Communication*, *35*(1), 95–100. https://doi.org/10.2190/MBYG-AK7L-5CT7-54DU

Jones, S. R., & McEwen, M. K. (2000). A conceptual model of multiple dimensions of identity. *Journal of College Student Development*, *41*(4), 405–414.

Kalan, A. (2015). A Practice-Oriented Definition of Post-Process Second Language Writing Theory. *TESL Canada Journal*, *32*(1), 1. https://doi.org/10.18806/tesl.v32i1.1196

Keh, C. L. (1990). Feedback in the writing process: a model and methods for implementation. *ELT Journal*, *44*(4), 294–304. https://doi.org/10.1093/elt/44.4.294

Knoblauch, C. H., & Brannon, L. (1981). Teacher Commentary on Student Writing: The State of the Art. *Freshman English News*, *10*(2), 1–4.

Krippendorff, K. (2012). *Content Analysis: An Introduction to Its Methodology*. SAGE Publications.

Kynell, T. (1995). English as an Engineering Tool: Samuel Chandler Earle and the Tufts Experiment. *Journal of Technical Writing and Communication*, *25*(1), 85–92. https://doi.org/10.2190/7L28-AQT3-PVU7-TYC5

Kynell, T. (1999). Technical communication from 1850–1950: Where have we been? *Technical Communication Quarterly*, *8*(2), 143–151. https://doi.org/10.1080/10572259909364655

Kynell-Hunt, T. (2000). *Writing in a Milieu of Utility: The Move to Technical Communication in American Engineering Programs, 1850-1950* (2nd ed.). Elsevier Science; 2 edition. Retrieved from http://www.amazon.com/Writing-Milieu-Utility-Communication-Contemporary/dp/1567504914

Lave, J., & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation*. Cambridge University Press.

Lea, M. (1998). Academic literacies and learning in higher education: constructing knowledge through texts and experience. *Studies in the Education of Adults*, *30*(2), 156–171. https://doi.org/10.1080/02660830.1998.11730680

Lea, M., & Street, B. V. (2006). The “Academic Literacies” Model: Theory and Applications. *Theory Into Practice*, *45*(4), 368–377.

Leydens, J. A. (2008). Novice and Insider Perspectives on Academic and Workplace Writing: Toward a Continuum of Rhetorical Awareness. *IEEE Transactions on Professional Communication*, *51*(3), 242–263. https://doi.org/10.1109/TPC.2008.2001249

Leydens, J. A. (2013). Integrating social justice into engineering education from the margins: Guidelines for addressing sources of faculty resistance to social justice education. In *Engineering education for social justice* (pp. 179–200). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-6350-0_9

Leydens, J. A., & Schneider, J. (2009). Innovations in Composition Programs that Educate Engineers: Drivers, Opportunities, and Challenges. *Journal of Engineering Education*, *98*(3), 255–271. https://doi.org/10.1002/j.2168-9830.2009.tb01023.x

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic Inquiry*. SAGE.

Linneman, T. J. (2010). *Social Statistics: The Basics and Beyond* (1st ed.). Routledge. Retrieved from http://www.amazon.com/dp/0415805015

Loretto, A., DeMartino, S., & Godley, A. (2016). Secondary Students' Perceptions of Peer Review of Writing. *Research in the Teaching of English*, *51*(2), 134–161.

Lu, J., & Law, N. (2012). Online peer assessment: effects of cognitive and affective feedback. *Instructional Science*, *40*(2), 257–275. https://doi.org/10.1007/s11251-011-9177-2

Marra, R. M., Rodgers, K. A., Shen, D., & Bogue, B. (2012). Leaving Engineering: A Multi-Year Single Institution Study. *Journal of Engineering Education*, *101*(1), 6.

Matusovich, H., Paretti, M., Motto, A., & Cross, K. (2012). Understanding Faculty and Student Beliefs about Teamwork & Communication Skills. In *2012 ASEE Annual Conference*.

McComiskey, B. (2000). *Teaching composition as a social process*. Retrieved from http://digitalcommons.usu.edu/usupress_pubs/127/

McComiskey, B. (2015). Introduction. In B. McComiskey, *Microhistories of Composition* (pp. 3–38). Utah State University Press. https://doi.org/10.7330/9781607324058.c000

Merriman, M. (1893). Training of students in technical literary work. In *Proceedings of the Society for the Promotion of Engineering Education* (Vol. 1). Chicago, IL.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook* (2nd edition). Thousand Oaks: SAGE Publications, Inc.

Murray, D. (1972). Teach Writing as a Process Not Product. *The Leaflet*. Retrieved from http://www.fredonia.edu/faculty/english/spangler/Renewal2009-10/12_f_Murray–Teach_Writing_as_a_Process.pdf

Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, *78*(3), 691–692.

National Research Council. (1969). *Colleges of Agriculture at the Land Grant Universities: A Profile*. Retrieved from https://www.nap.edu/catalog/4980/colleges-of-agriculture-at-the-land-grant-universities-a-profile

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: how different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375–401. https://doi.org/10.1007/s11251-008-9053-x

Neuendorf, K. (2011). Content Analysis-A Methodological Primer for Gender Research. *Sex Roles*, *64*(3–4), 276–289. https://doi.org/10.1007/s11199-010-9893-0

Newkirk, T. (1984). Direction and Misdirection in Peer Response. *College Composition and Communication*, *35*(3), 301–311. https://doi.org/10.2307/357458

Nilson, L. B. (2003). Improving student peer feedback. *College Teaching*, *51*(1), 34–38.

Nystrand, M., Greene, S., & Wiemelt, J. (1993). Where did Composition Studies Come from?: An Intellectual History. *Written Communication*, *10*(3), 267–333. https://doi.org/10.1177/0741088393010003001

O'Neill, P., & Fife, J. M. (1999). Listening to Students: Contextualizing Response to Student Writing. *Composition Studies*, *27*(2), 39–51.

Ozogul, G., & Sullivan, H. (2007). Student performance and attitudes under formative evaluation by teacher, self and peer evaluators. *Educational Technology Research and Development*, *57*(3), 393–410. https://doi.org/10.1007/s11423-007-9052-7

Paretti, M., McNair, L., & Leydens, J. A. (2014). Engineering communication. In B. M. Olds & A. Johri (Eds.), *Cambridge Handbook of Engineering Education Research* (1st ed.). Cambridge University Press.

Patchan, M. M. (2011). *Peer review of writing: Learning from revision using peer feedback and reviewing peers' texts*.

Patchan, M. M., & Schunn, C. D. (2015). Understanding the benefits of providing peer feedback: how students respond to peers' texts of varying quality. *Instructional Science*, *43*(5), 591–614. https://doi.org/10.1007/s11251-015-9353-x

Patchan, M. M., & Schunn, C. D. (2016). Understanding the Effects of Receiving Peer Feedback for Text Revision: Relations between Author and Reviewer Ability. *Journal of Writing Research*, *8*(2), 227–265. https://doi.org/10.17239/jowr-2016.08.02.03

Patchan, M. M., Schunn, C. D., & Clark, R. J. (2011). Writing in natural sciences: Understanding the effects of different types of reviewers on the writing process. *Journal of Writing Research*, *2*(3), 365–393.

Patchan, M. M., Schunn, C. D., & Correnti, R. J. (2016). The nature of feedback: How peer feedback features affect students' implementation rate and quality of revisions. *Journal of Educational Psychology*, *108*(8), 1098–1120. https://doi.org/10.1037/edu0000103

Patton, M. D., & Smith Taylor, S. (2013). Re-evaluating Directive Commentary in an Engineering Activity System. *Across the Disciplines*, *10*(1). Retrieved from http://wac.colostate.edu/ATD/articles/patton_taylor2013.cfm

Paulus, T. M. (1999). The effect of peer and teacher feedback on student writing. *Journal of Second Language Writing*, *8*(3), 265–289. https://doi.org/10.1016/S1060-3743(99)80117-9

Piaget, J. (1971). *Biology and knowledge*. University of Chicago Press.

Pinelli, T. E., Barclay, R. O., Keene, M. L., Kennedy, J. M., & Hecht, L. F. (1995). From Student to Entry-level Professional: Examining the Role of Language and Written

Communications in the Reacculturation of Aerospace Engineering Students. *Technical Communication*, *42*(3), 492–503.

Raymond, F. N. (1911). The preparation of written papers in schools of engineering. In *Proceedings of the Society for the Promotion of Engineering Education* (Vol. 19, pp. 48–90). Pittsburgh, PA.

Rourke, L., & Anderson, T. (2004). Validity in Quantitative Content Analysis. *Educational Technology Research & Development*, *52*(1), 5–18.

Rowell, J. A. (1989). Piagetian Epistemology: Equilibration and the Teaching of Science. *Synthese*, *80*(1), 141–162.

Sandler, B. R., Siiverberg, L., & Hall, R. M. (1996). The chilly classroom climate: A guide to improve the education of women.

Schneider, J., & Munakata-Marr, J. (2013). Connecting the "forgotten": Transportation engineering, poverty, and social justice in Sun Valley, Colorado. In *Engineering Education for Social Justice* (pp. 153–177). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-94-007-6350-0_8

Schreier, M. (2012). *Qualitative Content Analysis in Practice* (1 edition). Los Angeles: SAGE Publications Ltd.

Schwandt, T. A., Lincoln, Y. S., & Guba, E. G. (2007). Judging interpretations: But is it rigorous? trustworthiness and authenticity in naturalistic evaluation. *New Directions for Evaluation*, *2007*(114), 11–25. https://doi.org/10.1002/ev.223

Scientific Software Development GmbH. (2018). ATLAS.ti 8 Windows. Retrieved April 29, 2018, from https://atlasti.com/product/v8-windows/

Seely, B. E. (1999). The Other Re-engineering of Engineering Education, 1900–1965. *Journal of Engineering Education*, *88*(3), 285–294. https://doi.org/10.1002/j.2168-9830.1999.tb00449.x

Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Sharp, J. E., Olds, B. M., Miller, R. L., & Dyrud, M. A. (1999). Four effective writing strategies for engineering classes. *Journal of Engineering Education*, (January). Retrieved from http://onlinelibrary.wiley.com/doi/10.1002/j.2168-9830.1999.tb00411.x/abstract

Sluijsmans, D. M. A., Brand-Gruwel, S., & Merriënboer, J. J. G. van. (2002). Peer Assessment Training in Teacher Education: Effects on performance and perceptions. *Assessment & Evaluation in Higher Education*, *27*(5), 443–454. https://doi.org/10.1080/0260293022000009311

Smith, S. (1997). The Genre of the End Comment: Conventions in Teacher Responses to Student Writing. *College Composition and Communication*, *48*(2), 249–268.

Smith Taylor, S. (2007). Comments on Lab Reports by Mechanical Engineering Teaching Assistants: Typical Practices and Effects of Using a Grading Rubric. *Journal of Business and Technical Communication*, *21*(4), 402–424. https://doi.org/10.1177/1050651907304024

Smith Taylor, S., & Patton, M. D. (2006). Ten Engineers Reading: Disjunctions Between Preference and Practice in Civil Engineering Faculty Responses. *Journal of Technical Writing and Communication*, *36*(3), 253–271. https://doi.org/10.2190/07LL-2K2M-27KH-CX1W

Sommers, N. (1980). Revision Strategies of Student Writers and Experienced Adult Writers. *College Composition and Communication*, *31*(4), 378–388. https://doi.org/10.2307/356588

Sommers, N. (1982). Responding to Student Writing. *College Composition and Communication*, *33*(2), 148–156. https://doi.org/10.2307/357622

Srnka, K. J., & Koeszegi, S. T. (2007). From words to numbers: how to transform qualitative data into meaningful quantitative results. *Schmalenbach Business Review*, *59*(1), 29–57.

Straub, R., & Lunsford, R. (1996a). *12 Readers Reading: Responding to College Student Writing*.

Straub, R., & Lunsford, R. (1996b). *12 Readers Reading: Responding to College Student Writing*.

Strauss, A., & Corbin, J. M. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. SAGE. Retrieved from http://books.google.com/books?id=7BGtQgAACAAJ

Swarts, J., & Odell, L. (2001). Rethinking the evaluation of writing in engineering courses. In *Proceedings of the ASSS/IEEE Frontiers in Education Conference* (Vol. 1). https://doi.org/10.1109/FIE.2001.963905

Szymanski, E. A. (2014). Instructor feedback in upper-division biology courses: Moving from spelling and syntax to scientific discourse. *Across the Disciplines*, *11*(2), n2.

Topping, K. (2010). Methodological quandaries in studying process and outcomes in peer assessment. *Learning and Instruction*, *20*(4), 339–343. https://doi.org/10.1016/j.learninstruc.2009.08.003

Topping, K., & Ehly, S. (1998). *Peer-assisted Learning*. Routledge.

Traxler, M. J., & Gernsbacher, M. A. (1993). Improving written communication through perspective-taking. *Language and Cognitive Processes*, *8*(3), 311–334. https://doi.org/10.1080/01690969308406958

Trevelyan, J. (2010). Reconstructing engineering from practice. *Engineering Studies*, *2*(3), 175–195. https://doi.org/10.1080/19378629.2010.520135

Villanueva, V. (Ed.). (2003). *Cross-talk in comp theory: a reader* (2nd ed., rev. and updated). Urbana, Ill: National Council of Teachers of English.

Wadsworth, B. J. (1996). *Piaget's theory of cognitive and affective development: Foundations of constructivism, 5th ed* (Vol. xi). White Plains, NY, England: Longman Publishing.

Weaver, M. E. (1995). Using Peer Response in the Classroom: Students' Perspectives. *Research and Teaching in Developmental Education*, *12*(1), 31–37.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning, and Identity*. New York: Cambridge University Press. Retrieved from http://books.google.com/books?id=heBZpgYUKdAC

Wheeler, E., & McDonald, R. (2000). Writing in engineering courses. *Journal of Engineering Education*, 481–486.

Williams, J. M. (2001). Transformations in Technical Communication Pedagogy: Engineering, Writing, and the ABET Engineering Criteria 2000. *Technical Communication Quarterly*, *10*(2), 149–167. https://doi.org/10.1207/s15427625tcq1002_3

Williams, J. M. (2002). Technical Communication, Engineering, and ABET's Engineering Criteria 2000: What Lies Ahead? *Technical Communication*, *49*(1), 89–95.

Winsor, D. A. (1996). *Writing Like An Engineer: A Rhetorical Education*. Mahwah, N.J: Routledge.

Wolfe, J. (2009). How Technical Communication Textbooks Fail Engineering Students. *Technical Communication Quarterly*, *18*(4), 351–375. https://doi.org/10.1080/10572250903149662

Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing*, *13*(1), 29–48. https://doi.org/10.1016/j.jslw.2004.04.004

Appendix A: CIR Peer Feedback Handout

Why Peer Feedback?

Studies in education have shown that there is significant improvement in student papers when students have the opportunity to get feedback from their peers. Student's papers improve not only in terms of writing, but also in terms of content. One study (Cho & MacArthur 2010) even proved that there are times when the feedback you receive from your peers is as helpful as the feedback you receive from an instructor or subject matter expert. Why? Perhaps because your peers know what it's like to be writing the paper. They have also struggled with it and had to work through their ideas to put words to the page. You can share with one another how you overcame difficulties, found helpful sources, and corrected a certain paragraph that felt out of place in your argument. At this point, you are an expert on how to write for this assignment and you are in the best position to give your peer feedback, so long as it's quality peer feedback.

Quality Peer Feedback

There are <u>three</u> important things to remember as you read over your peer's paper.

1. **Be specific**
   Quality feedback is determined by the clarity of the feedback itself; give your peer specific comments instead of vague comments. Point to exact page numbers and paragraphs that were problematic, give examples of general problems that you found, be clear about what exactly the problem was, explain why it was a problem, etc. (Cho, Schunn & Charney 2006, Appendix B). The rubric for the CIR assignment is attached at the end of this handout. Reference it frequently as your read your peer's paper.

   For example, saying "I think you should switch the order of the third and fourth paragraphs" is more specific than saying, "You need to work on your organization."

2. **Be encouraging**
   When a peer gives you feedback on your writing, it is common to feel a little defensive, especially when you have put a lot of hard work into a piece of writing. However, try to hear feedback with an open mind. You and your peers are all trying to do the same thing: become better engineers. As you read your peer's paper, remember how this might feel. Be encouraging. You have nothing to lose and they have everything to gain from your encouragement. Remember to comment on aspects of their paper that are strong at the same time as you are suggesting areas of improvement. Write comments that you would like to receive as you work to revise your paper.

   For example, saying "Your previous transitions between paragraphs have been very smooth. This one could be stronger" is more encouraging than saying, "You lost me here."

3. **Be helpful**

Here are some questions you might ask yourself while you're reviewing a writer's paper. Notice that some of them ask you to focus on strengths and some ask you to identify weaknesses. The questions might also ask you to offer suggestions for improvement if you find areas of the paper that could use revision.

- Can you identify the piece's **main argument**?
- What are the paper's main **strengths**? What are the strongest parts of the paper?
- What sentences/ideas are **difficult** to understand? What sentences/ideas are **clear** and **easy** to understand?
- Are there **transitions** between paragraphs and ideas? If not, where might be good places for them?
- Are there sections in the paper that seem **unrelated** to the rest of the paper?
- Did you feel lost or **confused** at any point? How might the writer improve the areas that confused you?
- What might be **missing** from the paper? What else, as an interested reader, would you like to see in the paper? What else does the assignment ask for?

CIR Peer Feedback Examples

Attached is an excerpt from a student's paper from last semester. We have also provided some comments on that paper to give you examples in the context of reading a peer's paper.

Rechargeable Lithium-Ion Batteries

Fig. 1 Picture of a lithium-ion battery, Hemera/Thinkstock [1]

Abstract— With all of today's technological advances, there is a high demand for greater energy sources. The lithium-ion battery, which powers items such as cellular devices and laptops, has been undergoing research lately to improve its overall charge capacity. The desired goal is to increase the lithium-ion's maximum charge, have it charge completely at a faster rate, and maintain its charge over a long period of time. Recently there was a breakthrough when a group of researchers and scientists developed a protocol that had the ability to charge a lithium-ion battery 30-120 times faster than the simple battery. This new study has opened up multiple possibilities with battery-powered technology such as replacing the battery of small electrical devices and even using it in future electrical vehicles. Unfortunately at the present time, there are still glitches and improvements with the battery, as well as expense problems, that will most likely postpone mass production for a few more years.

Index Terms— Lithium-ion battery, anode, cathode, electrolyte, charge

1 INTRODUCTION

Lithium-ion batteries are some of the most common batteries in use today. They can be found in everyday portable electronics like iPods, laptops, and cell phones. These batteries are categorized as both primary (disposable) and secondary (rechargeable) batteries, depending on certain parameters such as voltage requirements, environmental conditions, maintenance, and the battery's lifetime [2]. The current problem with these batteries is that although they were a major breakthrough years ago, in today's fast-moving society, they take too long to charge and over a small period of time lose that charge. Researchers are currently looking for ways to speed up the charging process and capacity in order to implement the technology in small electronic devices and future cars.

2 HOW IT WORKS

In order to talk about the future, it is imperative to talk about the present state and charging process of batteries. Every battery contains an anode, cathode, and an electrolyte. When the anode and cathode are connected by an electrical conductor, an electrical current is created by the electrons flowing from the anode to the cathode. Meanwhile, the electrolyte conducts positive current in the form of cations, which are positive ions. The materials used for these components affect the battery capacity and voltage. In a

Your abstract and introduction present the problem clearly. By reading the abstract and then the introduction, I understand both the potentials and limitations of lithium-ion batteries as a contemporary issue.

Your description of how the technology works is very clear. The inclusion of a diagram complemented your description. Nicely done!

I think you do a good job of presenting examples of upgrading that would mean something to your readers. We can all relate to a cell phone that doesn't have a battery efficient enough to last all day.

lithium-ion battery, the lithium ion is the cation that travels through the electrical current. Other materials include a combination of lithium salts for the electrolyte, graphite (carbon) for the anode, and lithium cobalt oxide ($LiCoO_2$) for the cathode. All of these materials gives the battery 3.6 Volts (V), more than twice that of a disposable alkaline battery, and a better and more efficient energy density. Lithium-ion batteries are recharged by running the anode and cathode reactions in the opposing direction using an electricity source such as an outlet [3].

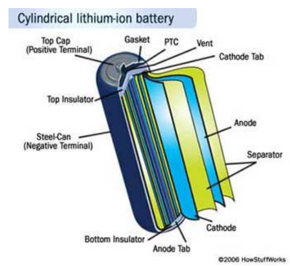


Fig 2. Parts of a lithium-ion battery [3]

3 UPGRADING

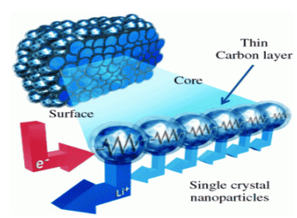
Although there may not seem like any downside to this type of battery, there are multiple glitches to the lithium-ion battery. These problems include the amount of time it takes the battery the charge, the degradation they undergo after leaving the factory, as well as a decrease in their ability to deliver current over time [3].

This engineering problem may not seem too bad on a global scheme, but when it comes to the advancement of technology in this fast-moving consumer society, without an improvement in the batteries used, certain advancements will come to a stand-still. A good example of what improving this battery could mean is having the ability to fully charge your cell phone or laptop under 15 minutes and the life of the battery lasting about a week. Another advancement that may be possible in the years to come is using the charged battery in an electric vehicle.

How does the upgraded lithium-ion battery differ from the old one? Korean scientists working at the Ulsan National Institute of Science and Technology (UNIST) have discovered a method that allows the battery to charge 30-120 times faster than a regular Li-ion battery; in other words, their research may lead to an electric car battery that can fully charge in less than a minute. The "Korean method," as it is referred to by Sebastian Anthony, soaks the cathode material, which in this case is lithium manganese oxide, in a graphite solution that, once carbonated, leaves a dense network of conductive traces throughout the cathode. This cathode, plus an electrolyte and a graphite anode, allow every part of the battery to recharge at the same time, rather than the much slower outside-in charge of a regular battery.

There are multiple lithium-ion family members that can be chosen for the cathode. The different materials are chosen based on what the battery will be used for. Determining factors include cost, stability, energy density, and voltage [2]. A detailed table outlining the different materials and their uses is located in the Appendix (Fig 3).

The reason why this new battery technology would be so useful in an electric car is mainly for size purposes. This is because the conductive trace network increases the total size of the battery, making it more probable to fit in a large car rather than a pocket cell phone.



The assignment asks you to discuss multiple/alternative solutions to the issue and the tradeoffs between them. You bring up the multiple cathode alternatives, but you don't go into detail. I'm curious what the differences are.

Fig 4. Lithium-ion cathode with carbonized graphite electrodes [4]

4 POTENTIAL PROBLEMS

These new lithium-ion batteries are not very far out of our reach. As stated before, research and successful tests have already been conducted. Now, however, comes the perfection stage of the process. One battery was able to be made, but what about mass production for consumers? FutureTimeline.net predicts that the new battery will be widely used in laptops, cell phones, other small electronics, and electric cars by the year 2015. So why wait three years? There are a few issues with making this battery more common including expenses, size, and glitches with the battery.

As far as expenses, with every new and exciting technological advance, there is always an issue with money. When scientists consider the future of this product, they envision the battery in an electric car. As stated before, a part of this is due to the fact that the size of the protocol is too large to fit in a laptop or cellular device. Secondly, the idea of a cheap, environmentally-friendly electric car has been around for years, but due to technological constraints, a lack of research, and money, the idea has never appealed to the masses. The improved lithium-ion battery, unfortunately, is no exception. The research and technology has finally caught up, but money is still a problem. By the time the battery is installed in the car, although it may be great to charge it and drive within a minute, the battery itself costs as much over time, if not more, than a car that uses gas [4]. This makes the electric car seem like a great idea, but in reality, the price would not make it a middle-class success.

The glitches with this battery are similar to those of the modern lithium-ion batteries. These problems include overheating of the anode due to the battery powering and overproduction of oxygen due to too much charge on the cathode. These two problems have the potential to start an abrupt fire. Despite the vents and automatic shut off when the battery gets too hot, the risk still exists, and there have been multiple instances where this has occurred [3]. As the size increases, more energy will be spent on keeping the battery cool to ensure that a fire does not occur, taking away from the total charge capacity the battery contains.

5 CONCLUSION

The lithium battery has recently undergone some new advances that allow it to charge 30-120 times faster than your typical Li-ion battery. This is done by soaking the cathode in a graphite solution, allowing the battery to recharge as a whole rather than from the outside-in. Hopefully by the year 2015, this battery will be implemented in small electronic devices as well as electric cars.

REFERENCES

[1] Brian, Marshall. "How Lithium-ion Batteries Work." *HowStuffWorks*. Web. Retrieved from <http://electronics.howstuffworks.com/everyday-tech/lithium-ion-battery.htm>.

[2] Byczek, Rich. "In Compliance Magazine." *In Compliance Magazine*. Same Page Publishing. Web. Retrieved from <http://www.incompliancemag.com/index.php?option=com_content>.

[3] American Physical Society. "Lithium-ion Batteries." *Physics Central*. Web. Retrieved from <http://physicscentral.org/explore/action/lithium-1.cfm>.

[4] Anthony, Sebastian. "Scientists Develop Lithium-ion Battery That Charges 120 times Faster than Normal." *ExtremeTech*. Ziff Zavis, Inc. Web. Retrieved from <http://www.extremetech.com/extreme/134635-scientists-develop-lithium-ion-battery-that-charges-120-times-faster-than-normal>.

[5] Fox, Will. "2015 Timeline of Events |2015 Timeline: Battery Technology Gets a Boosts" *2015 Timeline of Events*. Web. Retrieved from <http://www.futuretimeline.net/21stcentury/2015.htm>.

You use the term "money" in a few places, but I think you may mean affordability at times and financial backing at others. For consumers it might be an issue of affordability, but for developers it might be an issue of financial support for the research and development. I think being more specific with these terms might make your argument stronger by allowing you to save words so that you can talk about more than two ethical issues/societal impacts.

The assignment asks for a thorough discussion of Societal Impact/Ethical considerations related to proposed solution. Maybe a third topic in this section would make your discussion more complete.

CIR Peer Feedback Rubric

This rubric is identical to the rubric which will be used by graders to evaluate your final draft. Help your peer work through their strengths and weaknesses so that each aspect of their paper will meet excellent standards! Remember to offer suggestions to your peer.

Attribute	Unacceptable	Marginal	Proficient	Excellent	Weight
Writing Mechanics					
Used Template and submitted as PDF	Assignment instructions not followed	Some assignment instructions followed	Most assignment instructions followed	Report fully complies with instructions & requirements	1
Grammar, mechanics, and spelling	Consistently inadequate grammar, mechanics, and/or spelling; Errors impair meaning	Many errors, which affect writing clarity	A few errors, which do not impair meaning	Consistently correct use of grammar, mechanics, and spelling	1
Writing Quality					
Structure	Organization lacks coherency; Language and sentence structure is poor; Report is difficult to read	Organization of some sections is coherent; Report requires effort to read and understand	Distinct units of thought in paragraphs; clear transitions between developed, coherently arranged paragraphs	Apt, seemingly inevitable sequence of paragraphs; appropriate, clear and skillful transitions between sentences and paragraphs	2
Conceptual	Confuses some significant concepts; does not respond directly to the assignment	Author restates or paraphrases ideas from other sources but writing does not clearly demonstrate author's understanding	Shows good understanding of the ideas and methods of the assignment; goes beyond the obvious	Writing is original, shows clarity, and demonstrates depth of understanding; sophisticated analysis; fulfills, then exceeds the expectations of the assignment in some critical way	1
Support	Sources not cited / examples not given.	Inadequate sources: Provides some evidence/citations, but not always relevant, sufficient, or integrated into paper; undeveloped ideas or little analysis; limited use of textual evidence	Pursues thesis consistently; clearly develops a main argument with clear major points and appropriate evidence; makes effort to link rather than stack ideas;	Provides substantial, well-chosen evidence/citations (quotations or specific examples) used strategically; connections between ideas are evident; thesis consistently supported	2

Attribute	Unacceptable	Marginal	Proficient	Excellent	Weight
References	Sources not cited	Some sources cited; common reference format not sued	Some sources cited in-text AND references listed in Reference section	Sources cited using a common format; sources are cited in-text and listed in Reference section; sources are sufficient	1
Technical Quality					
Problem	The problem is simply stated, and there is no introduction or explanation of the problem.	The description of the problem is vague; no discussion of how the problem is a contemporary issue	Problem identified and introduced, but introduction should be expanded to improve clarity	Problem is clearly identified and precisely introduced; author explains how the topic is a contemporary issue	3
Problem or solution linked to ECE/CS	No link to ECE/CS for problem or solution	The author states there is an ECE/CS link but does not explain the link	The link is presented and partially explained, but link is superficial.	The strong link to ECE or CS is clearly explained.	3
Solutions	Does not discuss a solution	Only discusses a single solution		Discusses multiple solutions	3
Tradeoffs	Does not mention or discuss tradeoffs	Mentions tradeoffs but presentation does not demonstrate author's understanding of tradeoffs and/or does not compare since only one solution is included	Discusses tradeoffs and demonstrates author's understanding of tradeoffs with a limited comparison of solutions	Has comparisons of alternate solutions and/or description of tradeoffs; compares solution in terms of underlying principles (key concepts), strategies, or tradeoffs.	3
Ethics	Does not mention or discuss ethical issues	Mentions ethical issues associated with the problem or solution(s) but does not evaluate or explain	Brief discussion of ethical issues; discussion covers all relevant ethical issues but could benefit from an expanded discussion	Excellent evaluation or discussion of relevant ethical issues; conclusions are clear and logical	2
Societal Impact	Does not mention or discuss societal impacts	Mentions the impact of the problem or solution(s) on society but does not evaluate or explain	Brief discussion of societal impact; presentation is understandable, but could benefit from an expanded discussion	Excellent evaluation or discussion of relevant societal impact; conclusions are clear and logical	2

Appendix B: CIR Quality Analysis Instrument

CIR Quality Analysis

Please select your name and the Author ID (the three digits following the "A" in the file name).

* Required

Rater (That's You!) *

Choose

Author ID *

Your answer

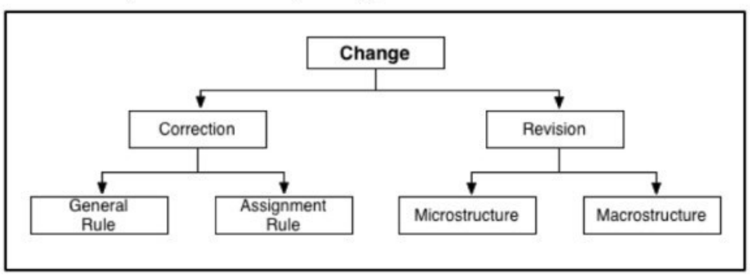
Quantification of Changes

Overall Quantity of Changes *

	None: Final and Draft are the same	Minor	Major
Quantity of Changes	○	○	○

Types of Changes Present *

For the purposes of this research, a change falls into one of four categories. First, I distinguish between changes that are corrections and those that are revisions. The main distinction is that corrections are based on a "rule" of some sort and do not substantially affect the meaning of the text. Corrections a further classified by whether they are grounded in general "good writing practice" vs. assignment-specific criteria, such as the document template or citation style. Revisions are classified using Faigley and Witte's micro/macrostructure dichotomy. Microstructure revisions affect meaning primarily locally around the change, whereas macrostructure revisions are substantive enough to affect a summary of the paper.



	None	Minimal	Moderate	Extensive
Corrections: General Rule	○	○	○	○
Corrections: Assignment Rule	○	○	○	○
Revision: Microstructure (Local Meaning Change)	○	○	○	○
Revision: Macrostructure (Global Meaning Change/Would Affect Summary)	○	○	○	○

Rating of Changes

For each rubric category, rate the difference between final and draft

Writing Mechanics *

Attribute/WEIGHTING	Unacceptable (0)	Marginal (2)	Proficient (4)	Excellent (6)
Writing Mechanics				
Grammar, mechanics, and spelling /1	Consistently inadequate grammar, mechanics, and/or spelling; Errors impair meaning	Many errors, which affect writing clarity	A few errors, which do not impair meaning	Consistently correct use of grammar, mechanics, and spelling
References /1	Sources not cited	Some sources cited; common reference format not sued	Some sources cited in-text AND references listed in Reference section	Sources cited using a common format; sources are cited in-text and listed in Reference section

	Negative	None/Negligible	Positive	Very Positive: Would change rubric score at least one category
Grammar, Mechanics, and Spelling	○	○	○	○
References	○	○	○	○

Writing Quality *

Attribute/WEIGHTING	Unacceptable (0)	Marginal (2)	Proficient (4)	Excellent (6)
Writing Quality				
Readability /1	Language and sentence structure is poor; Report is difficult to read	Report requires effort to read and understand	Report easy to read; transitions between paragraphs	Reads smoothly; appropriate, easily-followed transitions between sentences and paragraphs
Organization /1	Organization lacks coherency;	Organization of some sections is coherent	Distinct units of thought in paragraphs; coherently arranged paragraphs	Apt, seemingly inevitable sequence of paragraphs
Conceptual /1	Confuses some significant concepts; does not respond directly to the assignment	Author restates or paraphrases ideas from other sources but writing does not clearly demonstrate author's understanding	Shows good understanding of the ideas and methods of the assignment; goes beyond the obvious	Writing is original, shows clarity, and demonstrates depth of understanding; sophisticated analysis
Support /2	Sources not cited / examples not given.	Inadequate sources: Provides some evidence/citations, but not always relevant, sufficient, or integrated into paper; undeveloped ideas or little analysis; limited use of textual evidence	Pursues thesis consistently; clearly develops a main argument with clear major points and appropriate evidence; makes effort to link rather than stack ideas	Provides substantial, well-chosen evidence/citations (quotations or specific examples) used strategically; connections between ideas are evident; thesis consistently supported

	Negative	None/Negligible	Positive	Very Positive: Would change rubric score at least one category
Readability	○	○	○	○
Organization	○	○	○	○
Conceptual	○	○	○	○
Support	○	○	○	○

Technical Quality *

Attribute/WEIGHTING	Unacceptable **(0)**	Marginal **(2)**	Proficient **(4)**	Excellent **(6)**
		Technical Quality		
Problem /3	The problem is simply stated, and there is no introduction or explanation of the problem.	The description of the problem is vague; no discussion of how the problem is a contemporary issue	Problem identified and introduced, but introduction should be expanded to improve clarity	Problem is clearly identified and precisely introduced; author explains how the topic is a contemporary issue
Problem or solution linked to ECE/CS /3	No link to ECE/CS for problem or solution	The author states there is an ECE/CS link but does not explain the link	The link is presented and partially explained, but link is superficial.	The strong link to ECE or CS is clearly explained.
Solutions /3	Does not discuss a solution	Only discusses a single solution		Discusses multiple solutions
Tradeoffs /3	Does not mention or discuss tradeoffs	Mentions tradeoffs but presentation does not demonstrate author's understanding of tradeoffs and/or does not compare since only one solution is included	Discusses tradeoffs and demonstrates author's understanding of tradeoffs with a limited comparison of solutions	Has comparisons of alternate solutions and/or description of tradeoffs; compares solution in terms of underlying principles (key concepts), strategies, or tradeoffs.
Ethics /2	Does not mention or discuss ethical issues	Mentions ethical issues associated with the problem or solution(s) but does not evaluate or explain	Brief discussion of ethical issues; discussion covers all relevant ethical issues but could benefit from an expanded discussion	Excellent evaluation or discussion of relevant ethical issues; conclusions are clear and logical
Societal Impact /2	Does not mention or discuss societal impacts	Mentions the impact of the problem or solution(s) on society but does not evaluate or explain	Brief discussion of societal impact; presentation is understandable, but could benefit from an expanded discussion	Excellent evaluation or discussion of relevant societal impact; conclusions are clear and logical

	Negative	None/Negligible	Positive	Very Positive: Would change rubric score at least one category
Problem				
ECE/CS Link				
Solutions				
Tradeoffs				
Ethics				
Societal Impact				

Overall Ratings

Overall Ratings *

This is a SWAG, not a detailed analysis. Don't spend more than a minute thinking about it.

	F	D	C	B	A
Draft					
Final					

Comments

Just in case you wanted to note something...