Editorial Peer Review as a Content-Shaping Mechanism in Technical Communication Journal Scholarship

by

Suzan Flanagan

July 2019

Director of Dissertation: Dr. Michael Albers

Major Department: English

Editorial peer review serves multiple functions in academic journal publishing including gatekeeping, quality control, and mentoring. As representatives of a discipline and its body of knowledge, peer reviewers evaluate manuscripts and help determine what counts as knowledge, what methods and methodologies are acceptable in each discipline, what topics are valued, who gets published, and who gets cited. This mixed-methods study used genre theory as a framework for investigating the ways in which technical communication scholarship is shaped by editorial peer review; the relationships between peer review, editorial decision-making, and manuscript content development were examined. Analyses of 154 reviewers' reports from two technical communication journals showed limited agreement between reviewers' publication recommendations. Structural and comparative quantitative content analyses of reviewers' evaluative comments revealed that reviewers usually evaluated different aspects of manuscripts; when reviewers did evaluate the same aspects, they rarely disagreed. The results suggest that peer review operates as a type of social action in which reviewers internalize the generic conventions of journal scholarship and help authors shape content much like developmental editors do. These findings call for changes to the way we foster disciplinary knowledge-making and require actions such as defining manuscript disposition terms, reviewer roles, and tasks.

Editorial Peer Review as a Content-Shaping Mechanism in Technical Communication Journal Scholarship

A Dissertation

Presented to the Faculty of the Department of English

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Rhetoric, Writing, and Professional Communication

by

Suzan Flanagan

July 2019

EDITORIAL PEER REVIEW AS A CONTENT-SHAPING MECHANISM IN TECHNICAL COMMUNICATION JOURNAL SCHOLARSHIP

by

Suzan Flanagan

APPROVED BY:

DIRECTOR OF DISSERTATION: _____
Michael Albers, Ph.D.

COMMITTEE MEMBER: _____
Donna Kain, Ph.D.

COMMITTEE MEMBER: _____
Guiseppe Getto, Ph.D.

COMMITTEE MEMBER: _____
Margaret Bauer, Ph.D.

CHAIR OF THE
DEPARTMENT OF ENGLISH: _____
Marianne Montgomery, Ph.D.

DEAN OF THE
GRADUATE SCHOOL: _____
Paul J. Gemperline, Ph.D.

In memory of my father

Acknowledgements

I would like to thank my dissertation chair, Dr. Michael Albers; my dissertation committee members, Dr. Donna Kain, Dr. Guiseppe Getto, and Dr. Margaret Bauer, for their patience, pushback, and advice. I thank East Carolina University's Department of English and the professors who facilitated pilot testing during my coursework.

I thank the journal editors, editorial board members, editorial staff, and legal staff who made this research study possible. I thank the journal editors and editorial board members who provided (1) guidance in developing this study, (2) suggestions for follow-up studies, and (3) copies of their reviewers' guidelines, rubrics, and/or forms.

I thank the anonymous peer reviewers and authors for their labor and knowledge-making contributions to the technical communication field. In addition, I thank them for graciously allowing me editor-mediated access to the various publication-process artifacts analyzed in this research study.

I thank my research assistants for coding the reliability sample.

Finally, I thank my colleagues, friends, and family for their support and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AU: author

CA: content analysis

CATA: computer-aided text analysis

CCA: comparative content analysis

ED: editor

ES: editorial strategy

IRR: interrater reliability

J: journal

MPS: mean priority score

R: reviewer

RQI: review quality instrument

SME: subject matter expert

TC: technical communication

UX: user experience

Chapter 1

Introduction

Since the Royal Society implemented peer review around the 17th century,[1] academia has embraced the evaluative practice as the gold standard for screening and regulating journal scholarship. Shrouded in anonymity, the purportedly objective editorial practice has been used to assess scientific rigor, enforce disciplinary standards, and guide authors' manuscript revisions.

In today's publish-or-perish environment, editorial peer review continues to play a critical role in the evaluation and development of scholarship. I use the term *editorial peer review* here to distinguish the practice from peer review used in classrooms and from review by subject matter experts (SMEs) in workplaces. Unless specified otherwise, I use *peer review* throughout as shorthand for editorial peer review as practiced by most academic journals.

The International Committee of Medical Journal Editors defines peer review as "the critical assessment of manuscripts submitted to journals by experts who are not part of the editorial staff" (ICMJE, 2017; all key terms are defined in the glossary located in Appendix B). Peer review is used to evaluate factors such as accuracy, originality, writing quality, scientific methodology, and appropriateness for the publications' readers. Journals typically conduct peer review using some degree of anonymity; most commonly, peer review is arranged to be doubly anonymous, or double blind, which means neither the authors nor the reviewers know the

[1] The literature reports conflicting origins, which include the Royal Society of London, the Royal Society of Edinburgh, and the Académie Royale des Sciences of Paris, with dates spanning from 1662 to 1752. Dating to the mid-19th century, the modern incarnation of peer review is attributed to *Science* and the *Journal of the American Medical Association* (Hayhoe, 2010; Selfe & Hawisher, 2012).

identity of the other. Nonetheless, upon publication of a manuscript, the author's identity is consequently revealed to the reviewers. This study focuses on the double-blind form of editorial peer review; I briefly discuss other models of peer review in Chapter 2.

Although centuries have elapsed since peer review was adopted by the academy, limited empirical evidence exists—particularly within the technical communication literature—to support peer review's validity or reliability for evaluating manuscripts for publication in academic journals. Numerous studies comment on the lack of research on the effectiveness of peer review (e.g., Donovan, 2011; Jefferson, Alderson, Wager, & Davidoff, 2002; Jefferson, Wager, & Davidoff, 2002; Marsh & Ball, 1989; and Meruane, Vergara, & Pina-Stranger, 2016). Simply put, it's uncertain whether peer review works.

This uncertainty is problematic for scholars from all disciplines because academic committees make high-stakes, career-related decisions (e.g., hiring, promotion, tenure, research funding) based on the tenuous assumption that peer review works. Yet, meta-analyses suggest the peer review process is not only scientifically unproven but also flawed; most studies have shown high rates of disagreement between peer reviewers' evaluations, especially within the social sciences and humanities (Bornmann, Mutz, & Daniel, 2010). In studies with sample sizes ranging from 300 manuscripts to more than 6,000 manuscripts, researchers found reviewer agreement to be little better than chance (e.g., Kravitz et al., 2010; Rothwell & Martyn, 2000). Here, I mean by *reviewer agreement* that each reviewer (also called a referee) evaluated a specific manuscript and, in his or her report to the editor, the reviewer recommended the same publication outcome as the other reviewer(s). For example, all of the reviewers may have indicated that the manuscript needs major revisions before it can be published (see Figure 1.1).

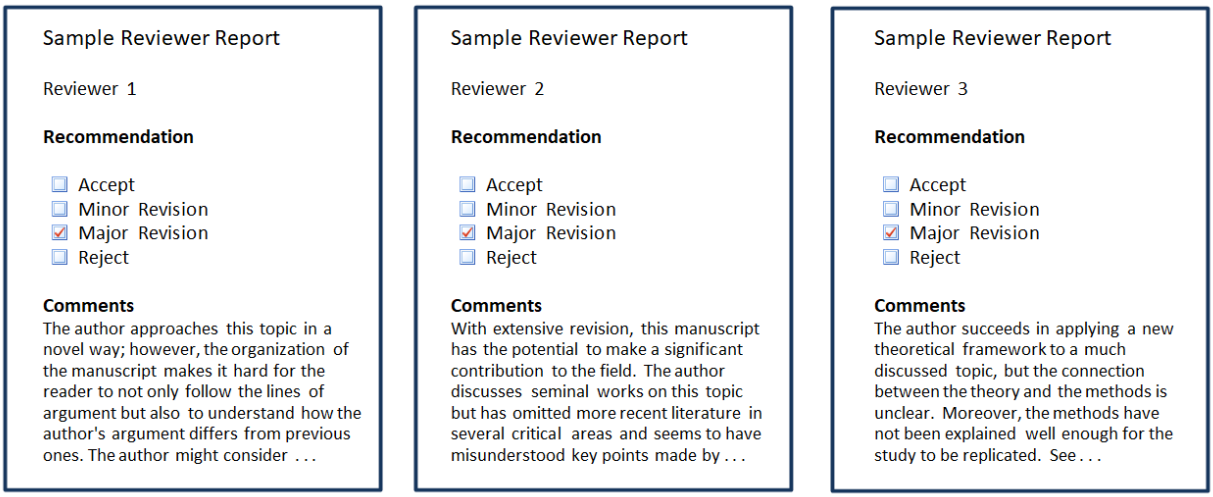


Figure 1.1: Sample reviewer reports that illustrate agreement between reviewer publication recommendations.

In contrast, the reviewers may disagree in their assessments of a manuscript. As illustrated in Figure 1.2, the reviewers' publication recommendations do not align (e.g., minor revision vs. reject) and the reviewers' evaluative comments (i.e., notes about the manuscript's strengths and weaknesses) are contradictory; in both respects, the reviewers disagree.

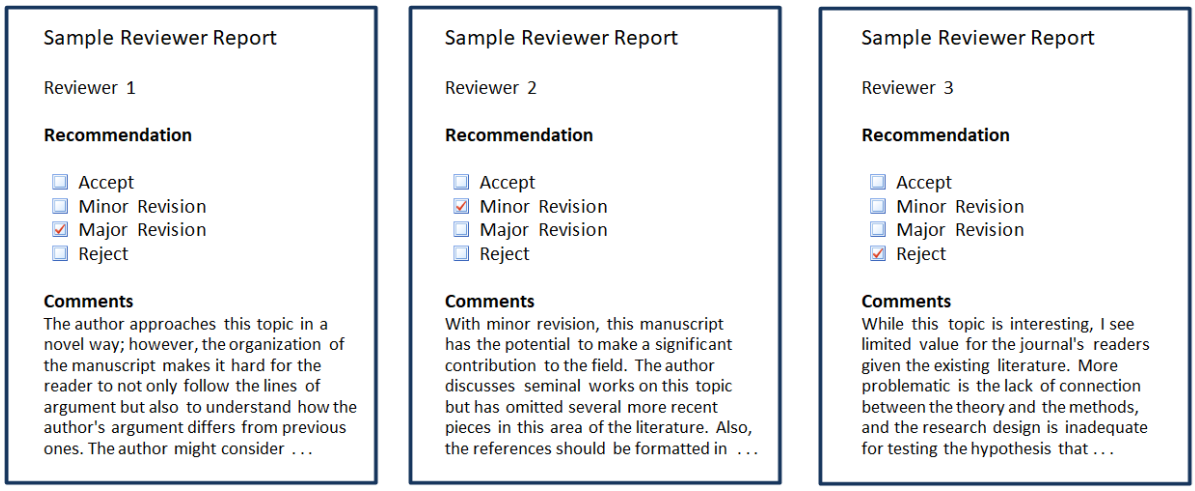


Figure 1.2: Sample reviewer reports that illustrate disagreement between reviewer recommendations and comments.

To be clear, the reviewers express their (dis)agreements indirectly through textual artifacts created as part of the publication process (e.g., reviewers' reports). In the same sense that the

literature represents academic conversations, these publication-process texts reflect editor-meditated conversations between reviewers and authors. When the texts are placed in conversation, the content may align or conflict; through their written words, the reviewers, authors, or editors may appear to (dis)agree with one another.

When reviewers do not agree in their recommendations, the editor's decision-making process may prove more difficult and less certain, which not only casts doubt on the reliability and validity of peer review as an evaluative practice but may also have implications for disciplinary knowledge-making. Multiple studies from other fields have shown that editors are influenced by reviewers' publication recommendations despite low interrater agreement rates.

Despite the problems of uncertainty and chance, the academy still relies on peer review, which suggests that peer review works in one or more capacities that are useful to editors. For instance, the literature indicates that reviewers' detailed evaluative comments influence authors' revisions and editors' publication decisions (e.g., Donovan, 2011; Kravitz et al., 2010; Lay, 2004). Kravitz et al. (2010) posited that "the information content of reviewers' narrative comments" may "drive editorial decisions in a more reliable and valid fashion than reviewers' summary recommendations" (p. 4).

Journal editors have discussed review criteria, provided examples of representative reviewer comments, and explained changes to editorial policies (e.g., Carliner, 2015; de Jong, 2009; Hayhoe, 1996, 2001, 2008; Lay, 2004). However, beyond anecdotal evidence from editors, little is known about how reviewers' comments influence the development of technical communication (TC) scholarship.

In its broadest sense, *technical communication* encompasses communication of specialized knowledge. In that respect, academic journals from myriad disciplines are technical

communication because they "communicate specialized information" (Hayhoe, 2010, p. 156). Unless specified otherwise, I use the term technical communication in the narrower sense as used by Rude and Eaton (2011): communication related to technology, technical subject matters (e.g., science, engineering), and technical procedures. This definition is consistent with those used by the Society for Technical Communication (TCBOK, n.d.) and the Bureau of Labor Statistics (2019); due to disciplinary tensions, the technical communication field itself does not have a unified definition.

Disciplines are defined by their common beliefs and their body of knowledge (Gale, 1998). Peer reviewed journals serve as repositories for each discipline's body of knowledge; the scholarly conversations occurring in each journal reflect the methodologies, frameworks, and topics favored by the disciplines and their scholars (Burbules, 2014). McNabb (2001) contends the "rules and regulations are both generated and maintained in the pages of the field's journals" (pp. 9–10). If this is true, then an analysis of publication-process artifacts from technical communication journals should help reveal how disciplinary rules are created and enforced and how scholarship develops within the TC field's disciplinary frameworks. The discipline's publication-process artifacts should contain clues about the identity, stability, and direction of the field and its scholarship.

We need to understand how peer review shapes the development and quality of TC scholarship in order to help scholars succeed in publishing their research. Hiring, promotion, and tenure decisions are determined, in part, by a scholar's publication history and the perceived quality of that peer reviewed scholarship—publications' acceptance rates and perceived quality are inversely related; the lower the acceptance rate, the higher the perceived quality of the scholarship (Hayhoe, 2010).

Academic publication opportunities are limited and partly contingent upon peer reviewers' evaluation reports and publication recommendations; their recommendations range from accept to reject (journals use different terms including revise and resubmit, major revision, minor revision, accept with revisions, and similar variations). Peer reviewers' often-contradictory recommendations inform editors' publication decisions—one reviewer may recommend that the manuscript be rejected and another reviewer may recommend that the manuscript be accepted for publication—and editors, in their decision-making process, must reconcile the apparent disagreement between reviewers' recommendations.

If peer review does not work as intended—if it indeed operates by chance—scholars may be wrongly offered or denied publication or funding opportunities (Sposato, Ovbiagele, Johnston, Fisher, & Saposnik, 2014). In turn, scholars' career progression and related opportunities may be impacted similarly; subsequently, the nature and the direction of the field and its knowledge may be altered.

Considering the high stakes, the TC field should try to confirm or refute ongoing criticisms of peer review in general. Criticisms range from "editorial capriciousness" and "poor inter-referee reliability" to "outright bias" (Selfe & Hawisher, 2012, p. 679) and elitism (Gale, 1998). Critics have identified problems with blinding (Lee, Sugimoto, Zhang, & Cronin, 2013), fairness, and transparency, and they have pointed out potential "adverse social, psychological, or financial effects" (Jefferson, Wager, & Davidoff, 2002, p. 2787). Some have criticized the peer review process for delaying the publication of research results (Goodman, Berlin, Fletcher, & Fletcher, 1994), whereas others have witnessed bias against negative research findings and ideas that are controversial or too novel (Armstrong & Hubbard, 1991).

Multiple aspects of the peer review process have been studied by fields other than technical communication. As Jefferson, Alderson, Wager, and Davidoff describe in their tabular summary of 19 studies, many research findings are mixed, inconclusive, or not generalizable (2002, Table B). Although numerous research studies seem to corroborate scholars' concerns about the reliability of the peer review process (i.e., low rates of agreement between reviewers), scholars cannot seem to agree whether reviewers should be in agreement. Some scholars view consensus as an indicator that peer review does not work (e.g., Sposato et al., 2014), whereas others hold the opposite opinion (e.g., Hirschauer, 2010). Nevertheless, none of the studies have recommended that academia abandon peer review altogether.

Scholars seem to agree that peer review plays a critical, but imperfect, role in the development of knowledge. Precisely what role peer review plays in shaping scholarship, particularly at the level of manuscript development, remains to be determined. Therefore, partly in response to Bornmann, Mutz, and Daniel's (2010) call for "comparative content analyses of reviewers' comments," this study investigates the ways in which the peer review process shapes technical communication scholarship (p. 8). In particular, this mixed-methods study examines the relationships between peer review, editorial decision-making, and manuscript content development. By *content development*, I mean the often iterative changes made to a manuscript after the author's initial submission to a peer reviewed journal; these changes include the author's revisions and the subsequent editing of accepted manuscripts—the changes may involve the text, data, visuals, multimedia, or other knowledge-related artifacts that comprise the manuscript.

To begin to determine how TC scholarship is shaped by peer review, I, in this phase of the study, analyze the content of various TC publication-process artifacts (e.g., peer reviewers'

reports and reviewers' guidelines). Among other things, I explore how peer review works on an editorial level to develop journal manuscripts and how (dis)agreements between reviewers affect editorial decisions and manuscript development. I look beyond the reviewers' publication recommendations (e.g., accept/reject) to their evaluative comments (e.g., feedback on the manuscript's merits and shortcomings), and I compare the points of (dis)agreement within and between reviewers' reports and their publication recommendations.

Research Questions & Hypotheses

My mixed-methods inquiry begins with the following research questions and hypotheses:

Research Questions

RQ_1: How well do the content and structure of the reviewers' evaluative comments align with the journal's reviewers' guidelines or scoring rubrics?

RQ_2: What role(s) do journal reviewers in the technical communication field play?

RQ_3: In what ways do reviewers' publication recommendations and evaluative comments shape editorial decisions and content development?

Hypotheses

H_1: No significant difference exists between the content of the reviewers' evaluative comments and the content of the journal's reviewers' guidelines or scoring rubrics.

H_2: No significant difference exists between the number of reviewer comments associated with higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) and the number of reviewer comments associated with lower level concerns (e.g., grammar, mechanics, style, citations).

H_3: No significant difference exists between reviewers' publication recommendations and editors' publication decisions.

H_4: For each manuscript, no significant difference exists between the types of problems each reviewer identifies.

H_5: For each manuscript, no significant difference exists between the number of manuscript problems each reviewer identifies.

H_6: For each manuscript, no significant relationship exists between the reviewer's publication recommendation and the number of manuscript problems the reviewer identified.

If we can better understand how peer review operates at the level of manuscript development, we might help scholars improve the quality of their scholarship and that of the technical communication body of knowledge as a whole.

Assumptions

The research study was designed based on the following assumptions:

- Peer review serves a useful editorial function.
- Journal editors purposefully select qualified reviewers who are familiar with the standards and conventions of their discipline and the journal.
- Journal editors expect reviewers to perform specific tasks and/or to fill specific roles.
- Reviewers fairly evaluate manuscripts to the best of their abilities.
- Reviewers evaluate manuscripts based on the standards and conventions of their discipline and the journal's reviewers' guidelines or rubrics.
- Authors make revisions based on editors' and reviewers' comments.

Limitations

Peer review is enacted slightly differently across disciplines, across publications within each discipline, across editors' tenures at each publication, and across reviewers, many of whom review for multiple publications within their discipline(s). This study design cannot control for those factors. Most of the data used for this study are publication-process materials previously collected by the academic journals as part of their standard operating procedures; therefore, it is impossible to control for possible biases in reviewer selection, decision-making, etc. Moreover, much of the data (e.g., double-blind peer review reports) must remain confidential and

anonymous; the nature of the data—and nondisclosure agreements—limits the ways in which the data can be analyzed and reported.

Permissions

This study was conducted with permission from each of the participating journals and East Carolina University's Institutional Review Board (IRB). Copies of the IRB approval letters (UMCIRB 17-001261 and UMCIRB 17-002615) are located in Appendix A; some details have been redacted.

Background

Before proceeding, I will briefly situate myself within the context of this study of technical communication journal scholarship. My interest in the relationships between peer review, editorial decision-making, and manuscript content development grew from my work as the managing editorial assistant for *Technical Communication Quarterly*, a position I held from Fall 2014 through Summer 2018.

Although I was never involved in the selection of reviewers or the manuscript publication decisions at the journal, I did have access to such information after the selections and decisions had been made. Furthermore, in performing my administrative duties, such as blinding manuscripts for review, I usually read the authors' cover letters and biographical information. As part of my copyediting process, I often read the reviewers' comments, the editors' comments and decision letters, and the authors' response letters.

My familiarity with these types of publication-process artifacts may have influenced how I interpreted the data in this study; however, having others code the data in this study should have helped curtail such biases. At the same time, my familiarity with these kinds of publication-

process artifacts and academic journal publication processes may have yielded insights that only those who have had privileged access could reach.

For the benefit of those who are unfamiliar with academic publishing conventions and the intricacies of peer review, I provide in the next section background information on academic journal publication processes, and in Chapter 2, I discuss the peer review literature.

Academic Journal Publication Processes

Academic journal publishing processes are similar across the academy. Whether it is a traditional subscription-based print publishing model or an open access digital publishing model, the publishing process involves collaborations between authors, reviewers, and editors (Burbules, 2014). Although some types of manuscripts may not be peer reviewed (e.g., editorials and book reviews), in most cases, the academic publication process includes peer review.

Figure 1.3 illustrates the basic peer review process; of course, peer review models and processes differ slightly from one journal to another. For example, one journal may use two reviewers whereas another may use three or four reviewers. At some journals, peer review is conducted by editorial board members; at other journals, peer review is conducted by outside experts and the editorial board functions in other capacities (e.g., advisory, administrative, policy-making).

Most journals use a form of anonymous peer review; the anonymous process is designed to "ensure quality control and to minimize role-conflicts among editors, referees, and contributors" (Hunter, 1995, p. 266). An increasing number of journals use other peer review models such as open review. I discuss the various peer review models in more detail in Chapter 2.

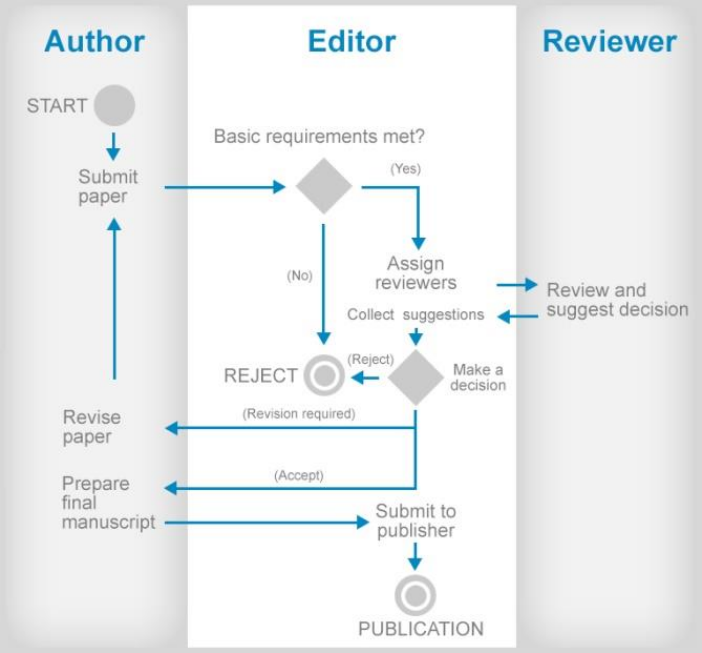


Figure 1.3: Basic peer review process (©2016 PhD on Track, Creative Commons 4.0 International License).

Regardless of the source of reviewers, the number of reviewers, or the model of peer review or publishing, manuscripts are rarely accepted without at least one round of revision. Authors may need to complete multiple major revisions before receiving an acceptance decision (Lay, 2004). In some fields where journal acceptance rates are in the single digits, the majority of manuscripts are rejected one or more times before being published, if at all (Goldberg, 2014; Jacobs, 2008; National Communication Association, 2012). Overall, the acceptance rates for the major peer reviewed journals in the technical communication field, which include *IEEE Transactions on Professional Communication*, the *Journal of Business and Technical Communication*, the *Journal of Technical Writing and Communication*, *Technical Communication*, and *Technical Communication Quarterly*, aren't as dire; prospects are best for special issues (Lay, 2004).

Journal acceptance rates reflect, among other factors, the economics and logistics of publishing. Money and space constrain the number of manuscripts that can fit within *x* pages of *y*

issues and be distributed for *z* dollars. Paradoxically, the acceptance rates for digital journals, which are not saddled with the costs of paper, ink, and postage, can be as competitive as print-based journals due to other costs that constrain output, such as labor, technology, production, and distribution. (Some of these expenses are recouped through subscriptions and other fees and/or reduced through volunteer labor.) Publication obstacles will remain regardless of whether or how peer review works.

With those obstacles in mind, I next outline the theoretical framework that I used to guide my investigation of editorial peer review within the technical communication field.

Theoretical Framework

A flexible framework is required to study peer review because the variability of peer review processes, models, functions, and policies complicates research designs.

I use genre theory as an analytic framework for examining how peer review shapes TC journal scholarship. As represented in Figure 1.4, peer review is an evaluative genre (Fortanet, 2008) that operates as a genre in the lower level sense of form and in the higher level sense of social action (Miller, 1984). As a form, peer review is understood in terms of recognizable categories or characteristics, such as types of discourse, content, and structure—all of which reflect disciplinary values and conventions (Devitt, Bawarshi, & Reiff, 2003). As a social action (i.e., an editorial practice or process), peer review is understood in terms of "typified rhetorical actions based on recurrent situations" (Miller, 1984, p. 159); in this context, peer review is an evaluative process that recurs in the realm of academic publishing based on socially constructed scholarly discourse and disciplinary knowledge. The discourse community's rules (disciplinary norms) are embedded within the genre of peer review and mediated by journal editors.

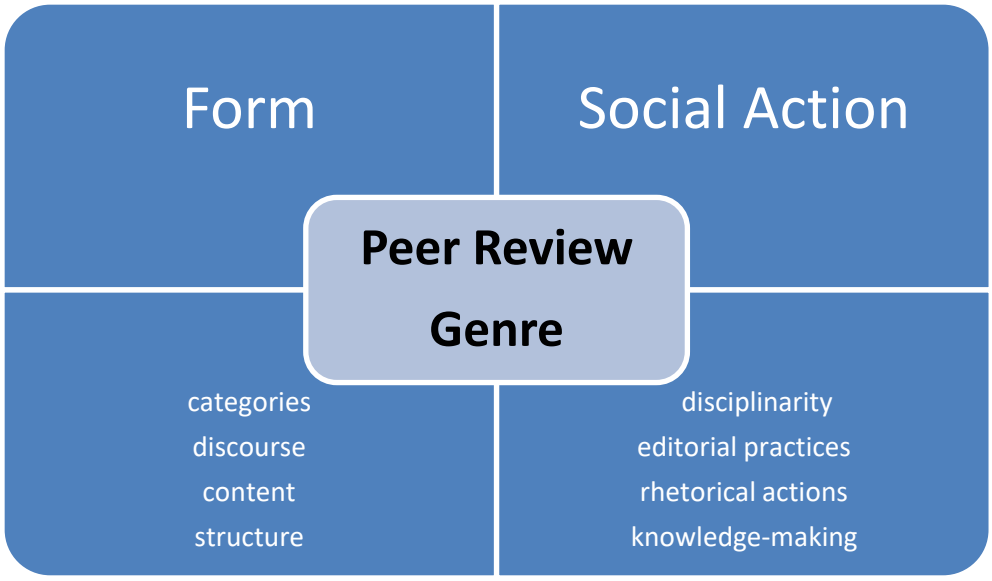


Figure 1.4. Peer review as form and social action.

My framework is informed by the genre theory work of Swales (1990), Gosden (2001, 2003), and Fortanet (2008), among others. For instance, Swales (1990) outlined a four-move genre analysis model: summary of judgment, outline of article, list of criticisms, and conclusion. Fortanet (2008) used Swales' model in her study of reviewers' reports, and Gosden (2001, 2003) took a similar approach but also incorporated Brown and Levinson's (1987) politeness theory and Halliday's (1985) metafunctional categories: ideational, interpersonal, and textual. Gosden found reviewers' responses typically followed a pattern of formulaic opening remarks, point-by-point replies, and closing remarks; he further segmented the evaluation comments into the following subcategories: discussion, technical detail, claims, references, and format—moves that are present in many peer review reports from the TC field. Using a related approach, Gonzalez (2006) based her analysis on the structure of the journal's reviewer comment form, which included criteria such as subject scope, appropriate treatment, scholarship documentation, and writing quality; these criteria resemble those in van Rooyen, Black, and Godlee's (1999) review quality instrument (RQI). A synthesis of these scholars' largely form-based approaches to genre

is helpful in answering my research questions, particularly RQ_1, which requires analyzing peer review in terms of content and structure.

However, to better understand how peer review works as an evaluative and content development mechanism, I must also approach the topic from a social action perspective because genres "affect and shape the social structure of the community or organization that uses them" (Luzón, 2005, p. 289). Genre knowledge, particularly knowledge of occluded—or blocked—publication-process genres, is critical to scholars' publishing success. Because "writers acquire and strategically deploy genre knowledge as they participate in their field's or profession's knowledge-producing activities" (Berkenkotter & Huckin, 2004, p. 285), we must study the peer review process to determine how the genre is working (or not working) to produce TC scholarship.

But to fully understand peer review, we must extend our view of the genre. Peer review consists of a constellation of subgenres (e.g., reviewers' reports, reviewers' guidelines) that form what Spinuzzi and Zachry (2000) call a *genre ecology*. In a genre ecology, the genres work together as an ecosystem to "jointly mediate the activities that allow people to accomplish complex objectives" (Spinuzzi & Zachry, 2000, p. 172). Through editor mediation, the peer review subgenres work together to shape TC scholarship. The genres operate within rhetorical contexts such as those hypothesized in Figure 1.5.

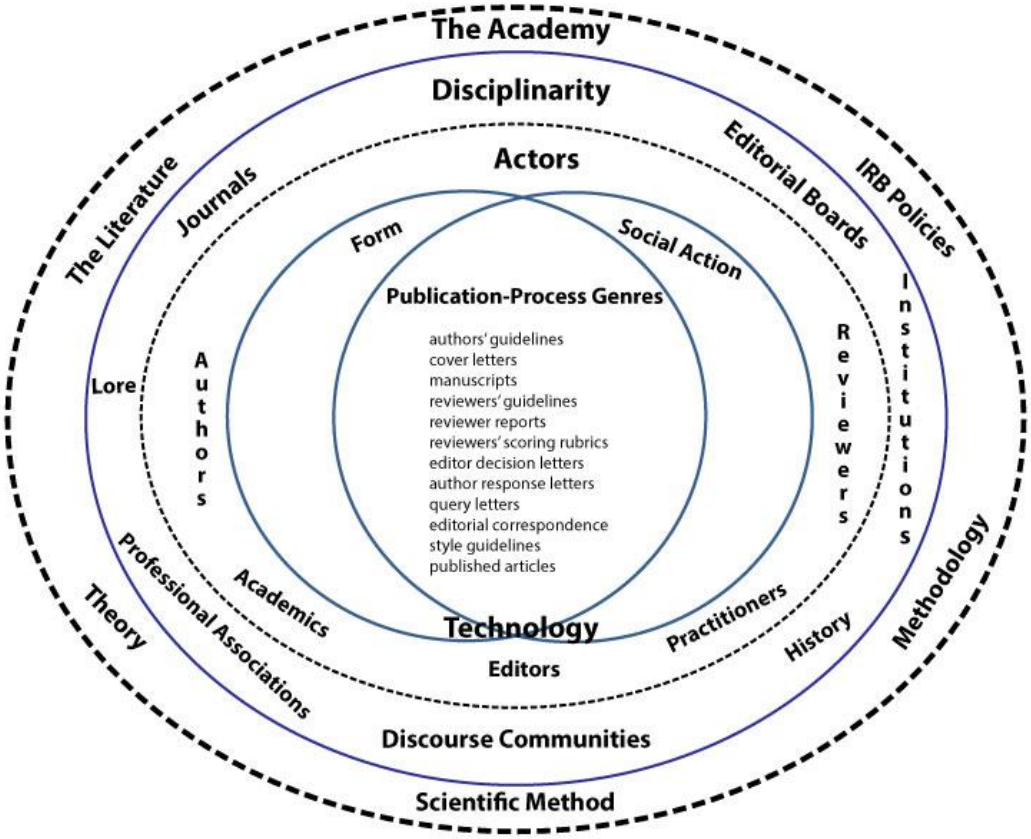


Figure 1.5. Hypothesized peer review genre ecology.

In framing peer review as a genre ecology, I look at three of the publication-process artifacts that comprise peer review (reviewers' reports, editors' decision letters, and reviewers' guidelines), and I study the following variables: review content, review evaluation criteria, types of problems identified, severity of problems identified, quantity of problems identified, reviewers' publication recommendations, reviewers' justifications, and editors' publication decisions. This analytical framework provides the flexibility to view genre as a form, social action, and ecosystem in order to explore the complex interaction of variables that shape TC scholarship. Follow-up studies will examine other elements of the peer review genre ecology.

Chapter Overview

The chapters of this dissertation follow an IMRaD structure (introduction, methods, results, and discussion). In this chapter, I have contextualized technical communication editorial practices as enacted by peer reviewed academic journals and laid the theoretical framework for examining how peer review shapes content and disciplinary knowledge. In Chapter 2, I summarize peer review literature related to academic journal publication-process genres, editorial decision-making, content development, and technical communication disciplinarity. In Chapter 3, I outline my mixed-methods procedures for conducting the study. In Chapter 4, I report the results using a combination of narrative description, descriptive statistics, and inferential statistics. In Chapter 5, I discuss the results of the content analyses and their significance, and I connect the findings to the study's research questions, hypotheses, and existing scholarship. Finally, in Chapter 6, I draw conclusions based on my interpretation of the results, note the limitations of my findings, and outline future lines of inquiry to further advance our understanding of how peer review shapes technical communication scholarship.

Chapter 2

Literature Review

In this chapter, I review portions of the peer review literature that relate to technical communication disciplinarity, academic journal publication-process genres, editorial decision-making, and content development. Voices from the technical communication (TC) field are conspicuously underrepresented in the academy's conversations about editorial peer review. A body of empirical literature—primarily from the fields of medicine and psychology—has exposed multiple flaws in the peer review system; however, across the academy, we still do not know precisely how peer review works in respect to content development and the shaping of a discipline's scholarship.

Technical Communication Disciplinarity

For decades, the technical communication field has struggled to define its identity (Dayton & Bernhardt, 2004). As a result, the field wasn't recognized by the Bureau of Labor Statistics until 2011 (Henning & Bemer, 2016). Dayton and Bernhardt (2004) attribute TC's identity crisis to its members' disciplinary backgrounds, which range from English studies and literature to composition studies and rhetoric:

> Those who identify primarily with composition studies and/or rhetoric are likely to view the field more broadly than those who identify primarily with technical/professional communication . . . the need to include both *technical* and *professional* as modifiers of *communication* points to a more subtle but still consequential difference in perspective on the boundaries of our knowledge domain. (p. 41)

Traces of these disciplinary affinities and boundaries are evident in the journals' publication-process artifacts. Some scholars believe the TC field lacks "a coherent body of knowledge" and attribute the "disjunction in knowledge bases" to the existence of "entirely different knowledge bases in the field" beyond the expected differences between academics and practitioners (Johnson-Eilola & Selber, 2001, pp. 407–408). An analysis of reviewers' reports can help us map content development within the knowledge boundaries of technical communication scholarship.

Numerous technical communication scholars have reflected upon the discipline's professional identity in edited collections (e.g., Johnson-Eilola & Selber, 2004; Kynell-Hunt & Savage, 2003, 2004), special issues (e.g., Coppola, 2011, 2012), and various other journal articles (e.g., Dayton & Bernhardt, 2004; Giammona, 2004; Savage, 1999, Schuster, 2015). Scholars have examined the discipline's body of knowledge (e.g., Coppola & Elliot, 2013), research questions and issues (e.g., Blakeslee, 2009; Rude, 2009), authorship characteristics (Lam, 2014), TC curriculum (e.g., Henschel & Meloncon, 2014; Johnson, 2009; Johnson-Eilola & Selber, 2001), TC job ads (e.g., Brumberger & Lauer, 2017), and academy–industry relationships (e.g., Boetteger & Friess, 2016).

As the field of technical communication has expanded, trends have emerged such as the use of rhetorical lenses, interdisciplinary research, and collaborative research (Schuster, 2015). Special issue topics reflect how the boundaries of the field have stretched (e.g., gaming, online teaching, data visualization, intercultural communication, social media, and social justice). In turn, the malleable boundaries lend themselves to a reexamination of the definition of the field. For example, to help legitimize and empower the field, Henning and Bemer (2016) propose expanding the definition of technical communicator listed in the labor bureau's *Occupational Outlook Handbook*:

> Technical writers, also called technical communicators, prepare instruction manuals, how-to guides, journal articles, and other supporting documents to communicate complex and technical information more easily. They also develop, gather, and disseminate technical information through an organization's communications channels. (Bureau of Labor Statistics, 2019)

Henning and Bemer's (2016) proposed revision follows:

> Technical writers, also called technical communicators, produce documents in a variety of media to communicate complex and technical information. They employ theories and conventions of communication to develop, gather, and disseminate technical usable information among specific audiences such as customers, designers, and manufacturers. (p. 328)

Several scholars have expressed concerns over the lack of disciplinary boundaries and the risk of pushing the boundaries too far beyond the field's origins in technical and scientific communication (e.g., Dayton & Bernhardt, 2004; Schuster, 2015). Henning and Bemer (2016) contend their revised definition can help the technical communication field establish a strong brand identity. Their definition is grounded in technical and scientific communication and presents a core disciplinary identity while at the same time making room for multiple perspectives and emerging technologies.

However, a definition alone is not sufficient to establish the field's identity—the field must share common discourse and beliefs upon which its body of knowledge is built. Ultimately, a discourse community's knowledge-making is contingent upon its scholars' adherence to disciplinary argumentation practices and genre conventions (Berkenkotter, 1995). But first, its members must learn the accepted discourse practices and conventions. Peer review plays a critical role in disciplinary knowledge-making and enculturation.

Hirschauer (2010), who studied judgments of peer reviewers and editors in the sociology field, views peer review as a social process of enculturation that begins with manuscript submission and concludes with publication. Similarly, Blakeslee (2001) says "newcomers to a domain learn its genres through immersion and participation in the activities of the domain" (p.

169), and Gonzalez (2006) contends "peer review serves the socialization and integration of new members into a field" (p. 127). For Gonzalez, reviewers' evaluative comments function as "pedagogical tools" (p. 127). Much can be learned from reviewers' comments in terms of publication-related genre knowledge and disciplinary knowledge—this knowledge is critical to scholars' success in a publish-or-perish academic culture. At the same time, reviewers' evaluative comments and the peer review genre itself can help us understand disciplinary power dynamics.

Disciplinary Power Dynamics

Power dynamics are often revealed through socially constructed genres; however, the perception of power varies based on one's position(s) within the hierarchy and one's motivation for participating in the discourse community (Devitt, Bawarshi, & Reiff, 2003). Theresa Enos, former editor of *Rhetoric Review*, contends "peer reviewers have a more active role in shaping the conversation than editors have" (Gale, 1998, p. 202).

Like Enos, Burbules (2014) downplays journal editors' power in shaping scholarship and argues power is distributed to and enacted by the scholarly discourse community: Authors control the topics and quality of their submissions; as volunteers, reviewers participate in the publication process—typically without knowledge of the authors' or other reviewers' identities—by evaluating manuscripts based on their disciplinary knowledge and areas of expertise. Reviewers' evaluations are "informed by generally shared objectives of the field and of the journal for which they review" (Gonzalez, 2006, p. 128).

Editors function as the audience, evaluator, mediator, and final judge, whereas, reviewers are often viewed as advisors and evaluators (Fortanet, 2008). From this viewpoint, the power dynamics place the reviewer between the editor and author(s) with the author(s) at the bottom of

the power hierarchy, in that the reviewer evaluates the author(s) and advises the editor. The author(s) must satisfy the reviewers and editor(s), yet the author has agency to reject revision suggestions outlined in peer review reports. Eden (2008) recommends complying with reviewers' advice unless "the issue is of prime importance and it would be intellectually dishonest" (p. 246). Of course, when reviewers disagree in their manuscript evaluations, the author must determine whose advice, if any, to follow. Although editors usually do not edit reviewers' reports, editors often provide guidance on how to handle conflicting reviewer advice, clarify reviewers' concerns, or note their own concerns (Eden, 2008). Occasionally, editors will omit portions of reviews that contradict the editor's publication decision.

In some instances, the anonymity of peer review may significantly alter the power dynamics and discourse of peer review, resulting in harsh criticisms instead of constructive feedback (Getchell & Amicucci, 2014). When necessary, editors can frame reviewers' criticism constructively or censor harsh remarks (Eden, 2008). An analysis of peer review reports is a key step in understanding how power manifests in occluded genres and how a discipline's scholarship is constructed.

Socially Constructed Knowledge

Peer review is an integral part of knowledge-making. As representatives of a discipline's community of scholars, peer reviewers enact the community's values, beliefs, and ways of knowing; peer reviewers may also limit "what will be admitted into a community's body of knowledge" (Thralls & Blyler, 2004, p. 129). In a social constructivist approach, knowledge is collaboratively developed through the consensus of community members (Thompson, 2001; Thralls & Blyler, 2004). Within this normative feedback loop, the community shapes the discourse, and the discourse shapes the community (Thralls & Blyler, 2004).

Consensus can be elusive, particularly for newer disciplines with scholars who hold diverse epistemological views. For example, a controversial 1988 "Burkean Parlor" column in *Rhetoric Review* led to a multiyear spat over whether narrative accounts of personal experience count as evidence. Although community members may not agree on everything, disputes like this one show how the social constructionist view of knowledge is in tension with the objective, empirically oriented positivist view that underlies peer review's origins.

Some composition scholars questioned whether "people *want*[ed] their work judged and 'validated' by the process leading to publication in refereed journals" (Gebhardt et al., 1995, p. 239). Many scholars hold the view that peer review is needed to maintain legitimacy within the academy. As it is, *College Composition and Communication* did not implement peer review until the mid-1980s.

For some technical communication scholars, peer review is a must have, if only to validate the field and its knowledge. For others, some of whom came to the field from other disciplines, peer review is standard operating procedure—that is, we've always done it that way. And for some TC scholars, peer review serves as a stopgap in knowledge development—they hope to eventually replace peer review with something better. Notwithstanding their disciplinary or epistemological stance, authors and reviewers should work toward a mutual goal: "the publication of high-quality research" (Fischer, 2011, p. 227).

Academic Journal Publication-Process Genres

Research on genre is abundant in the technical communication literature (e.g., Henze, 2004; Kain, 2005; Luzón, 2005; Miller & Fahnestock, 2013); however, the TC field has devoted little attention to academic journal publication-process genres—occluded genres unavailable for public scrutiny—and their role in scholarly publishing. Among the publication-process genres

classified as occluded are submission cover letters, peer review reports, editors' decision letters, authors' response letters, and other editorial correspondence (Fortanet, 2008; Swales, 1996).

Of particular interest is the peer review report genre because peer review plays a critical role in the production of scholarship. As a genre, peer review is black-boxed or occluded across disciplines (Baruch, Konrad, Aguinis, & Starbuck, 2008; Eden 2008; Gosden, 2003; Hirschauer, 2010; Meruane et al., 2016), which is problematic because the rhetorical patterns of genres illustrate the discourse community's "values, beliefs, and ideologies" (Devitt et al., 2003, p. 554). When conversations about publication processes and editorial decision-making remain private, scholars find it difficult to identify and evaluate the variables involved in knowledge-making. Peer review "support[s] and validate[s] the manufacture of knowledge, directly as part of the publishing process," but is "rarely part of the public record"; exemplars are "hidden . . . from public gaze by a veil of confidentiality" (Swales, 1996, pp. 45–46). As a result, scholars may fail to recognize publication-related genre conventions, and, in turn, fail to meet publishing expectations.

In the context of academic publishing, genre can function as a gatekeeping framework that disrupts professional development and enculturation into discourse communities (Swales, 1996, 2000, 2004). Based on his research on reviewers' comments, Gosden (2003) argues that novices, especially those who are nonnative English speakers, need access to gatekeeping discourse and exemplar texts to become acculturated into their discipline. Novices may misread "the rhetorical purposes behind referees' comments" (Gosden, 2003, p. 99). Such misunderstandings about the peer review genre may impede authors' revisions.

Genres do not perform the desired function for outsiders to the discourse community because the outsiders don't understand what elements of the discourse are significant (Devitt et al., 2003).

"A discourse community is built on the premise that what we know and do is connected to the language we use" (Devitt et al., p. 549). This view of genre is consistent with Gee's (2014) view of discourse as "a characteristic way of saying, doing, and being" (p. 47). Selected for their expertise, reviewers have mastered their discipline's discourse and have internalized the "purposes, values, and assumptions" embedded within the peer review genre (Devitt et al., p. 553). Unfortunately, the ethical constraints that support genre occlusion also inhibit others from developing disciplinary expertise (Guthrie, Parker, & Dumay, 2015). Occlusion is not conducive to the acquisition of tacit genre knowledge or to disciplinary enculturation.

When transparent, genres serve as heuristics or "intellectual scaffolds" for disciplinary knowledge-making that enable scholars to become vested in their discourse communities (Berkenkotter & Huckin, 2004, p. 304). Some scholars, especially digital humanities scholars, advocate open review as a solution to the black-box research problem (e.g., Ball, 2017; Selfe & Hawisher, 2012). The journals that currently use open review models tend to be ones with open access publishing models, which some people perceive as little better than vanity publishing or self-publishing.

In most academic publishing models, manuscripts must fit "editors' various conceptions of what particular knowledge their journals are making," and by extension, the manuscripts must also fit reviewers' concepts of knowledge-making (Gale, 1998, p. 200). Yet, the occluded genre of peer review thwarts scholars' efforts to analyze those knowledge-making constructs. An examination of occluded publication-process genres is needed to help us better understand the nature and boundaries of technical communication scholarship as well as to help other scholars develop content that meets the discipline's expectations—expectations enforced through editorial peer review.

Editorial Peer Review

Both as a form and as a social practice, editorial peer review is occluded or hidden from public view. As a result of the occlusion, the number of empirical studies on editorial peer review are limited, and much of the recent research is flawed or not generalizable (Fortanet, 2008; Jefferson, Wager, & Davidoff, 2002; Jefferson, Alderson, et al., 2002; Meruane et al., 2016). Table 2.1 shows a range of studies conducted by other fields; designed much like the sociology study conducted by Bakanic, McPhail, and Simon (1989), this TC study explores the relationships between reviewers' evaluative comments and their publication recommendations, among other things.

Few articles on editorial peer review exist within the English-related fields of rhetoric, writing studies, and technical communication (TC); those scarce articles are limited to commentaries and editorials. For instance, in an article from rhetoric and writing studies, Sheffield, Sparks, and Ianetta (2014) argue for transparency of editorial practices, discuss their experiences working for *Rhetoric Review* and *College English* as graduate students, and offer anecdotal evidence of peer review processes at those journals. In other articles, Selfe and Hawisher (2012) and Ball (2017) argue for open peer review, especially for digital humanities scholarship where blinding is difficult, if not impossible, because multimedia files may feature authors' images or voices, and website domain names can reveal institutional affiliations or authors' names. Also problematic are articles with numerous self-citations; the blinding process effectively marks the manuscript and calls attention to the masked author and the clues to their identity.

Table 2.1

Aspects of Peer Review Studied by Disciplines outside Technical Communication

Study	Discipline	Sample Size	Study Summary	Findings	Limitations
Bakanic, McPhail, & Simon (1989)	sociology	323 manuscripts 775 reviews	Analyzed reviews of manuscripts submitted to one sociology journal; examined the relationships between reviewers' comments & their publication recommendations	Frequent (40%) disagreement for publication outcomes, yet limited (11%) disagreement in content of evaluative comments	Positive/negative binary omits neutral comments and possibly skews results toward negative feedback
Bornmann, Weymuth, & Daniel (2010)	chemistry	98 manuscripts 223 reviews	Compared reviewers' comments of manuscripts rejected by one chemistry journal; 70% of the manuscripts were later published in other low- or high-impact journals	Negative reviewer comments about relevance & design decreased author chances of publication in high-impact journal	Only studied rejected manuscripts later published with minimal changes. Binary regression models may skew negative
Hirst & Altman (2012)	health	116 journals	Studied the availability of peer reviewer report guidelines on journal websites	Only 35% of sample had online reviewer guidelines	Single coder; sampled mostly traditional journals versus online journals
Schroter, Black, Evans, Godlee, Osario, & Smith (2008)	general medical	607 reviewers	Examined types of errors reviewers detected & whether training improved error detection	Training was minimally helpful; reviewers detected about 1/3 of the major errors	All manuscripts reported findings of randomized control trials (RCT); notable Hawthorne effect; severity of introduced errors affected completion rates
van Rooyen, Black, & Godlee (1999)	biomedicine	1868 reviews	Developed an instrument for assessing quality of peer review reports	Version 3.2 of the instrument showed good test-retest scores & good interrater reliability (IRR = .83)	Requires training to get high IRR; assesses scope of review, not the accuracy of the review; tested on editorial staff (expert users) only

From the TC field, Lay (2004) reflects on her editorship of *Technical Communication Quarterly* and offers insights on the editorial peer review process. In the same manner, Hayhoe (1996, 2001, 2008) shares anecdotes from his 12-year tenure as editor of *Technical*

Communication, outlines review criteria, and praises the journal's review process. He highlights two major editorial policy changes: (1) a requirement for manuscripts "to address specifically how the theory or research it reports can be applied by practitioners," and (2) new review procedures that required each manuscript to be evaluated by both an academic reviewer and a practitioner reviewer (2008, p. 7). Hayhoe's successor, de Jong (2009), further refined *Technical Communication's* manuscript submission guidelines and review criteria "to ensure both practical relevance and research quality" (p. 2). Reading between the lines, these editorial changes suggest shortcomings in the review process were remedied.

In a more recent piece, Carliner (2015), former editor of *IEEE Transactions on Professional Communication,* explains the rationale for changes to the journal's submission guidelines and reviewer forms. Carliner recounts how peer review data collectively led to *IEEE Transactions on Professional Communication*'s changes in author's guidelines and reviewers' forms; the journal's article structure now aligns more closely with that of *Technical Communication*. Those structural changes shape how research is reported in those empirically oriented journals and facilitate publication decisions by standardizing how writing is organized, what content is mandatory, and what criteria should be used in reviewers' evaluations. These types of editorial structures and changes can help us understand how editorial practices, such as peer review, affect content development and shape disciplinary knowledge.

As enacted by technical communication journals, such as *IEEE Transactions on Professional Communication* and the *Journal of Business and Technical Communication*, peer review is a double-blind (anonymous) manuscript evaluation practice that consists of editor-mediated private communications (e.g., peer reviewers' reports, editors' decision letters, and authors'

response letters) between the author(s), reviewer(s), and editor(s). Prior to publication, neither the author(s) nor the reviewer(s) knows the others' identities in this model of peer review.

Peer Review Models

Not only do peer review models and processes vary from publication to publication, but the terminology may vary too. For example, some journals use the term *anonymous* to mean single blind and others use it to mean double blind. In contrast, other journals avoid the term *blind* altogether because of its ableist connotations. Most of the journals in the field of technical communication use a double-blind (or doubly anonymous) peer review model. Unless noted otherwise, this study uses the term *anonymous* to indicate a double-blind review process—that is, both the reviewers' and the authors' names are unknown to one another during the review process. (In publishing, the term *double blind* is used in a different sense than in medical research, which uses the term to indicate that neither the administering researcher nor the participant know whether the intervention is a placebo.) See Table 2.2 for a description of several peer review models as I am defining them; some sources define these models in contradictory ways, perhaps due to disciplinary differences.

Each peer review model has its advantages and disadvantages, and perceptions vary by discipline. For example, medical and science disciplines tend to perceive public comments as useful to the development of science, but the quality and reliability of the crowdsourced feedback may be questionable (Ford, 2013). Across disciplines, reviewers have expressed reluctance to participate in open review; in one study, 49% said they would be less likely to review if required to sign reports (Selfe & Hawisher, 2012; van Rooyen et al., 1999). In contrast, open review eliminates the need to blind manuscripts—the cost of blinding manuscripts is perceived to be a major disadvantage, not just in terms of labor, but also in terms of limiting

reviewers' access to information needed to evaluate the research (e.g., citations to authors' previous work) and in terms of delaying distribution of research (Lee et al., 2013). As it is, reviewers' identities may be inadvertently revealed to authors if identifying metadata has not been scrubbed from the review files. For some digital scholarship, such as that published in *Kairos*, blind review is not practical because the authors' identities are embedded in the digital artifacts (Ball, 2017).

Table 2.2

A Comparison of Peer Review Models

Peer Review Model	Description	Disciplines Using Model[2]
Single Blind	Reviewers' identities are anonymous	Life sciences, physical science, engineering
Double Blind	Authors' and reviewers' identities are anonymous	Technical communication, humanities, social sciences, clinical medical, nursing
Triple Blind	Authors' and reviewers' and editors' identities are anonymous	Various medical and science fields
Anonymous	May refer to single-, double-, or triple-blind review	Most disciplines use a form of anonymous review
Open (Disclosed)	Authors' and reviewers' identities are known; reviews may be posted online with the articles	Humanities, education, various medical and science fields
Transparent	Manuscripts, anonymized reviews, and authors' responses are posted online with the article	Various medical and science fields
Signed	Reviewers' names are published but their reviews remain confidential unless the authors and reviewers agreed to open review	Various medical and science fields
Crowdsourced	Public comments	Various medical and science fields
Hybrid	Mixture of public comments and blind review	Various medical and science fields
A Priori (Prepublication)	Public comments before peer review	Various medical and science fields
A Posteriori (Postpublication)	Public comments after peer review and publication	Various medical and science fields
Synchronous	An iterative publishing model in which publication and review occur simultaneously	Various medical and science fields

[2] Compiled from Ford, 2013; Lee et al., 2013. This list is not comprehensive.

Peer Review Functions

Peer review can serve multiple functions including gatekeeping, quality control, and mentoring. First and foremost, peer review serves a critical gatekeeping function in academic publishing (Berkenkotter, 1995). As representatives of a discipline and its body of knowledge, peer reviewers evaluate manuscripts and help determine what counts as knowledge, what methods and methodologies are acceptable in each discipline, what topics are valued, who gets published, and who gets cited. Indirectly, peer reviewers help determine who gets hired, who gets tenure, who gets promoted, and who gets grant funding. In many respects, peer reviewers may help shape the direction of their field. Through their recommendation reports, peer reviewers help editors make sometimes-tough editorial decisions (Bailar, 1991).

As a quality control measure, peer review not only helps editors cull manuscripts that are not appropriate for their journals but also helps editors uphold the journal's standards and scholarly reputation (Burbules, 2014; Hirschauer, 2010). Editors rely on reviewers with specialized areas of expertise to prevent substandard scholarship from being published, to identify and cultivate potentially publishable manuscripts, and to recognize novel scholarship with merit (Armstrong & Hubbard, 1991). Occasionally, peer review seemingly malfunctions in its evaluation of novel scholarship: A notable example is *Nature*'s rejection of Stephen Hawking's seminal work on black holes (Jackson, Srinivasan, Rea, Fletcher, & Kravitz, 2011).

As a mentoring tool, peer review functions to enculturate scholars into a discipline (Gonzalez, 2006). In their mentoring role, peer reviewers discipline scholars through constructive criticism and perhaps a harsh rebuke that firmly establishes disciplinary and/or journal scholarship boundaries. Gardner, Willey, Jolly, and Tibbits (2012) contend the peer review process may be "the most important opportunity to acquire the standards and norms of

the discipline and develop researchers' judgement" (p. 1). Reviewers' qualitative feedback can help scholars understand disciplinary expectations and conventions (Burbules, 2014).

Presumably, an editor's strategy for choosing peer reviewers is related to the role(s) or function(s) that the editor expects the reviewer to play in the publishing process. Beyond roles such as gatekeeper and mentor, peer reviewers may be selected for the role of a generalist reader. In considering the questions of whether and how peer review works, one must first determine the capacities in which peer review is expected to function. In other words, how does peer review work for a specific purpose? One must also consider the policies and procedures that guide the implementation of peer review for specific purpose(s) within specific disciplinary context.

Peer Review Policies, Procedures, & Guidelines

Established by the Committee on Publication Ethics (COPE), the "COPE Ethical Guidelines for Peer Reviewers," delineate the "basic principles and standards to which all peer reviewers should adhere during the peer review process" (COPE Council, 2013, para. 1). The organization also outlines ethical guidelines and best practices for journal editors. As of this writing, about 12,000 academic journals are members of the organization, including the *Journal of Technical Writing and Communication*, the *Journal of Business and Technical Communication*, and *Technical Communication Quarterly* (COPE Council, 2017).

Depending on the journal and the discipline, editors may or may not provide reviewers with specific instructions for reviewing. For the most part, reviewers are expected to be familiar with peer review genre conventions as well as be familiar with disciplinary conventions and the specific journal's editorial aims and scope. Some reviewers' instructions include guiding questions but leave the structure of the review up to the reviewers. In contrast, some journals provide detailed reviewers' guidelines, evaluation forms, or rubrics that provide structural and

content guidance for the peer reviewers (e.g., *IEEE Transactions on Professional Communication*). Reviewers' forms usually include spaces to indicate publication recommendations (e.g., accept, minor revision, major revision, reject), the reviewer's willingness to evaluate revisions of the manuscript, confidential remarks to the editors, and evaluative comments for the authors. Reviewers may have the option to provide their feedback by attaching marked-up, commented manuscripts files. Some review systems enable reviewers to view previous versions of the manuscript, the other reviewer's comments, and the author's responses to the editor's decision letter(s). Journals without content management systems may conduct peer review by email. Before the digital age, editors mailed photocopies of manuscripts to reviewers. Regardless of the distribution method, the review copies of manuscripts are redacted for anonymous peer review models—any identifying information is removed to facilitate an impartial review process.

Bias

In all disciplines, anonymous peer review is designed to minimize bias by following Merton's norms of universalism and organized skepticism; universalism requires independent assessment of knowledge claims—knowledge is evaluated using "universal or impersonal criteria" rather than "on the basis of race, class, gender, religion, nationality, or personal attributes" (Howard, 2012, p. 323). Those outside the occluded peer review conversations often perceive the peer review process as biased; even so, several studies suggest that peer review minimizes bias as intended. For example, Zuckerman and Merton's (1971) seminal study showed "the academic status of reviewers did not affect their acceptance . . . of manuscripts submitted by authors of differing status" (Howard, 2012, p. 324).

Selfe and Hawisher (2012) contend "doubly anonymous" peer review helps "avoid bias against women scholars" (p. 679). In their meta-analysis of peer review studies, Ceci and Williams (2011) found no evidence of gender bias (Lee et al., 2012). Still, some feminist scholars (e.g., Lakoff, 1979; Lakoff & Bucholtz, 2004; Tannen, 1994, 2002; and Sawin, 1999) argue that gender can be revealed by discourse markers. In the context of peer review, it could be problematic if scholars' discourse patterns reveal their genders because earlier studies indicate that "articles supposedly written by women tended to be rated less highly than identical articles supposedly written by men" (Lloyd, 1990, p. 539). Although this study was not specifically designed to examine potential gender bias or other biases, the study should have detected any blatant patterns of bias. Of course, within the context of knowledge production, bias is not necessarily bad unless it operates to exclude specific groups, methodologies, research topics, etc.

Bias can be productive when it favors "research that is important, original, well designed, and well reported" (Meruane et al., 2016, p. 190). Reviewers value novelty (Berkenkotter & Huckin, 2004), and that value is usually embedded in reviewers' guidelines or scoring rubrics. Scholars are expected to contribute something new to the conversation. However, in breaking new ground, scholars are placing themselves "either in conflict with the existing ideas and knowledge or extending it in ways with no baseline by which to judge it" (Eberley & Warner, 1990, p. 220). Unless we study the content of reviews in the TC field, we can only speculate as to whether and how bias shapes TC scholarship. For that matter, we do not know how peer review works as a social practice to extend or constrain the discipline's body of knowledge.

Many scholars view peer review as socially constructed. Among those social constructions are boundaries. Howard (2012) perceives peer review as boundary work that "highlights the historical and sociological activities of scientists and explains how demarcations (boundaries)

can change over time" (p. 333). Falsification is used to draw the boundary "between science and non-science" (Howard, 2012, p. 326) and to determine what counts as knowledge. Rightly or wrongly, the academy perceives anonymous peer review as more objective for evaluating knowledge; "the more objective the reviews seemed in the eyes of the academy, the more certain were tenure and promotion committees that an accepted article represented scholarship that was independently vetted by the scholarly community and thus appropriate for publication" (Selfe & Hawisher, 2012, pp. 673–674). Given the high stakes, it is important that we understand how peer review shapes TC scholarship.

Meta-Analyses

Numerous aspects of peer review have been studied, primarily by editors or former editors who have access to the data. The definition of peer review—or lack thereof—is problematic in some of the studies. For instance, some studies (e.g., Goodman et al., 1994) don't differentiate between peer review and technical editing, yet peer review, author revision, and technical editing impact manuscript quality in different ways at different stages of the publication process.

By and large, comparisons of peer review studies are difficult because the studies use different instruments with different scales, and most instruments are not validated, or have low reliability; nevertheless, several meta-analyses have been conducted. The details of five meta-analyses conducted since 2002 are compiled in Table 2.3. Four of the five meta-analyses are from the biomedicine field. Notably, the results of the largest study (i.e., Bornmann, Mutz, & Daniel's 2010 study of 19,443 manuscripts) suggest that high interrater reliability (IRR) might be undesirable because IRR seems to correlate inversely with sample size.

Table 2.3

Meta-Analyses of Peer Review Studies outside Technical Communication

Study	Discipline	Sample Size	Study Summary	Findings
Bornmann, Mutz, & Daniel (2010)	economics, law; natural sciences; medical sciences; social sciences	48 studies 28 to 1983 manuscripts	Examined studies' interrater reliability (IRR) coefficients, including Cohen's Kappa, intraclass correlation (ICC), and Pearson correlation (r)	Interrater reliability tends to correlate with sample size: the smaller the sample, the higher the level of agreement; studies with high IRR are less credible because the sample sizes are usually small
Bruce, Chauvin, Trinquart, Ravaud, & Boutron (2016)	biomedicine	22 studies 50 to 609 (units of analysis were manuscripts and/or peer reviewers)	Examined randomly controlled trial studies to determine impacts of interventions on peer review quality; study topics included • reviewer training (e.g., Houry et al., 2012) • statistical peer review (e.g., Cobo et al., 2007) • reviewer checklists (e.g., Cobo et al., 2011) • open peer review (e.g., Vinther et al., 2012) • blinded peer review (e.g., Alam et al., 2011) • accelerated peer review (e.g., Johnston et al., 2007)	Inconclusive findings; Bruce et al. (2016) recommend establishing clear goals, guidelines, and outcomes for peer review
Godlee (2002)	biomedicine	5 studies 109 to 558 reviews	Examined effects of signed and unsigned reviews on two or more variables: • review quality • publication recommendations • review completion time • reviewer participation	Open review is feasible and ethically superior, lacks major adverse effects, recognizes reviewers' labor, and increases accountability
Jefferson, Alderson, Wager, & Davidoff (2002)	biomedicine	19 studies 36 to 568 reviews	Examined various topics: • effects of blinding on review quality (e.g., Das Sinha et al., 1999) • reviewer bias (e.g., Ernst & Resch, 1999; Fisher et al., 1994) • effects of reviewer guidelines, checklists, evaluation scales, and/or training (e.g., Callaham et al., 1998) • effects of peer review and editing on manuscript quality (e.g., Goodman et al., 1994; Pierie et al., 1996) • open review (e.g., Walsh et al., 2000) • peer review validity (e.g., Elvik, 1998)	Most of the research designs were flawed; many studies had problems with one or more of the following: • randomization • representativeness • generalizability • self-selection bias • response rates • blinding
Jefferson, Wager, & Davidoff (2002)	biomedicine	19 studies 36 to 568 reviews	Examined effects of peer review and technical editing on quality of published articles	Need established peer review goals to assess and improve process

Editorial Decision-Making

While peer review helps mitigate problems with bias, peer review seldom simplifies editorial decision-making. As previous studies have shown (e.g., Berkenkotter, 1995), editors regularly encounter contradictory feedback from reviewers.

Reviewer (Dis)agreement

At the root of reviewer disagreements may be the uncertainty over the purpose of peer review. Jefferson, Wager, and Davidoff (2002) contend we cannot evaluate the effectiveness of peer review, much less address its shortcomings, until we establish its purposes. The call for clear peer review goals is echoed in multiple research studies. Perhaps the equivalent of Van Buren and Buehler's (1980) levels of edit is needed for peer review.

Without a mutual understanding of peer review, reviewers are bound to disagree in their manuscript evaluations. Reasons for reviewer disagreement run the gamut:

- Reviewers use different evaluation criteria or scoring scales (Price & Flach, 2017).
- Reviewers have different areas of expertise or disciplinary backgrounds (Marsh & Ball, 1989).
- Reviewers represent different audience perspectives, such as lay reader, generalist, or specialist (Hirschauer, 2010).
- Reviewers examine different aspects of the manuscript (Fiske & Fogg, 1990).
- Journals lack standardized instructions and/or standardized forms for reviewing (Gosden, 2003; Hirst & Altman, 2012).
- Editors select reviewers who complement one another (Marsh & Ball, 1989).

Studies suggest that standardized assessment forms improve reviewer agreement and help reviewers evaluate research reports' methods, data analysis, and results (Rothwell & Martyn, 2000). Even so, some fear reviewers' rubrics might be misused as criteria to justify rejections (Hirst & Altman, 2012). Notwithstanding how reviewer rubrics are used, the lack of standardized

reviewer rating systems partly accounts for high between-study variations of interrater reliability (Bornmann, Mutz, et al., 2010). In their meta-analysis, Bornmann, Mutz, et al. (2010) confirmed previous studies' findings of low IRR (.34 using a random effect model and .17 using Cohen's kappa, which accounts for chance agreement).

Whatever the reasons for disagreements, the extent of disagreement varies considerably between the reviewers' qualitative feedback and their corresponding publication recommendations. Consensus tends to occur more frequently between reviewers' evaluative comments than between reviewers' publication recommendations (Goodman et al., 1994). Irrespective of the reviewers' publication recommendations, the number of negative reviewer comments usually exceeds positive comments (Bakanic et al., 1989).

The majority of the literature reports on the high rates of disagreement between reviewers' publication recommendations, whereas the relatively few studies devoted to reviewers' reports and reviewers' qualitative feedback point to more complexity in the evaluations. For example, a study from the sociology field, in which researchers analyzed 323 initial manuscript submissions and the corresponding reviews, found only 11% of the manuscripts received contradictory qualitative feedback from reviewers; however, 40% of the reviewers' publication recommendations differed (Bakanic et al., 1989). Bakanic et al. found three areas accounted for nearly 60% of the disagreements in their sample: theory, style, and results. Analysis, measurement, and general comments accounted for about 25% of the disagreements, and the remaining 15% of the disagreements were attributed to review, sample, data, design, and topic.

According to another study, when the publication recommendations differ, editors consult additional reviewers 43% of the time and reach their own decision 40% of the time (Grod, Lortie, & Budden, 2010). Moreover, it is not uncommon for editors to make decisions that go

against reviewers publication recommendations (i.e., publish manuscripts that reviewers have recommended be rejected or reject manuscripts that reviewers have recommended be accepted) because reviewers' evaluative comments influence editors' decisions (Bailar, 1991). In publication decision processes, the roles of reviewers and editors are complementary. Although tough reviewers may influence editors' decisions more, editors are tougher overall in their evaluations than reviewers, and editors' decisions are probably more reliable because they have more information on which to base their decisions (Eden, 2008; Marsh & Ball, 1989; Sposato et al., 2014).

One study that compared reviewers' evaluative comments found few instances of explicit disagreements between reviewers or between reviewers and editors but provided few details regarding the areas or spread of disagreement (Fiske & Fogg, 1990); the researchers attributed reviewers' publication recommendation disagreements to the minimal overlap in review content. Although the reviewers seldom focused on the same topics, the researchers believed the reviewers' comments were appropriate (Fiske & Fogg, 1990). Their findings are compatible with Sposato et al.'s (2014) observation that different reviewers detect different flaws; the researchers speculated that the types of flaws detected align with the reviewers' areas of expertise. The nature of these findings resembles those of usability studies; in usability studies, different users detect different problems—together, through one or more rounds of testing, the users typically detect most of the serious problems. Peer review operates in a similar fashion. Through one or more rounds, reviewers identify manuscript weaknesses that can be characterized in simple, reductive terms:

- "something not done or omitted,"
- "something done incompletely,"
- "something done poorly," or
- "something done wrong" (Fiske & Fogg, 1990, p. 593).

In their analysis of 153 manuscripts (and the corresponding reviews and decision letters) submitted to seven American Psychological Association journals, Fiske and Fogg (1990) identified 10 categories of manuscript weaknesses (conceptual: pre-execution; conceptual: linkage to execution; design; procedures; measurement; statistical analyses; results; interpretations and conclusions; editorial and writing; and general). They attributed the weaknesses to two subcategories: (1) planning and execution and (2) presentation.

Weaknesses in the first subcategory align with higher level problems, and weaknesses in the second subcategory align with lower level problems. The distribution and implications of these weaknesses in technical communication manuscripts is unknown. Also unknown is whether reviewers are in agreement on these types of weaknesses, some of which pertain only to TC manuscripts with quantitative data—a small fraction of the largely qualitative TC body of knowledge.

Validity & Reliability

From a statistical standpoint, higher levels of agreement are perceived as indicators of good scientific rigor, yet from an applied science standpoint, higher levels of agreement may be perceived as "detrimental to the review process" and as "a sign that the review process is *not* working well, that reviewers are not properly selected for diversity, and that some are redundant" (Bailar, 1991, p. 138, as cited in Bornmann, Mutz, et al., 2010, p. 6). Low levels of reviewer agreement can be beneficial to the development of scholarship when multiple perspectives are represented (Meruane et al., 2016). Of course, when editors select reviewers for their different perspectives single-reviewer reliability decreases while validity increases (Bornmann, Mutz, et al., 2010; Marsh & Ball, 1989).

Researchers have analyzed peer review using metrics such as single-reviewer reliability, rater response bias, and reviewer toughness. Marsh and Ball (1989) contend these metrics can help "standardize the review process" (p. 167); *single-reviewer reliability* is "the correlation between two independent reviewers of the same manuscript across a large number of manuscripts submitted for publication" (p. 152), and *rater response bias* is "systematic differences in the leniency or harshness of ratings that are idiosyncratic to a particular reviewer and that generalize across the reviews of different manuscripts by the same reviewer" (p. 154). Reviewer toughness scores are calculated using large data sets of reviewers who evaluate multiple manuscripts or who comprise a limited pool of reviewers (e.g., editorial board members; Sposato et al., 2014).

Researchers have also developed metrics for evaluating peer review reports and manuscript quality. For example, van Rooyen et al. (1999) developed the review quality instrument (RQI), which determines "the extent to which a peer reviewer has considered key aspects of a manuscript," but does not verify the accuracy of reviewers' comments (p. 628). With modifications to accommodate theoretical manuscripts, the instrument could be useful to the TC field in that reviewers must substantiate their comments and provide comprehensive, specific, constructive feedback to receive high scores, and in doing so, the reviewers can facilitate manuscript revision and editing.

Similarly, Sposato et al. (2014) used a mean priority score metric to represent manuscript quality. Manuscript priority score points are assigned as follows: rejection (1), major revision (2), minor revision (3), and acceptance (4). The cumulative scores are divided by the number of reviews to determine the mean priority score. In a study of more than 31,000 peer review reports, Sposato et al. (2014) found mean priority scores to be the best predictor of manuscript acceptance. On the surface, their finding suggests that this metric might be useful for checking

for correlations between reviewers' recommendations and editors' publication decisions; however, their calculations are problematic in that they involve averaging interval data.

Irrespective of the rates of reviewer agreement or the quality of reviewers' reports, editors "synthesize [reviewers'] comments and ratings and arrive at decisions that are more accurate than would be suggested by the low relationship between individual reviewer [manuscript] quality ratings or recommendations" (Jackson et al., 2011, p. 6). In other words, for manuscript selection, the peer review process seems to work despite low interrater agreement. Nevertheless, some scholars believe reviewer agreement should "exceed chance by at least 20%," but, by some accounts, approximately 7–18 reviewers would be needed to achieve that goal (Kravitz et al., 2010, p. 2). In contrast, other studies found that agreement levels decreased with each additional reviewer; moreover, studies with smaller sample sizes tend to produce higher agreement rates (Jackson et al., 2011; Sposato et al., 2014).

Most scholars would agree that we need "a shared understanding of required standards" (Gardner et al., 2012, p. 5); however, agreeing on review standards can be difficult, especially in fields where disciplinary norms are not well defined (Eberley & Warner, 1990). Considering the technical communication field still struggles to define itself, one would expect limited agreement on the discipline's norms, review standards, and the ways in which content should be developed.

Content Development

Journal content is developed through collaborations between authors, reviewers, and editors (Burbules, 2014). The collaborations involve editorial processes such as peer review, revision, and editing. Some authors prepare for the collaborative process at the research stage and design their study with a target journal in mind, while other authors complete their research—and

perhaps even the research article—before contemplating a suitable publication venue. Both strategies affect how content is developed.

Each publication has different aims, scopes, and target audiences, all of which are reflected in the publication's authors' guidelines and reviewers' guidelines. Scholars who are aware of these differences and who follow the guidelines should find the content development process less arduous; by proactively shaping their content to meet the journal's expectations, scholars can improve their odds of publication. Scholars who are initially less cognizant of the publication's expectations will become familiar with them during the peer review process, if not sooner (e.g., a bench rejection).

Publication criteria are also reflected in peer reviewers' and editors' evaluative comments (Fortanet, 2008; Lay, 2004); this qualitative feedback can be useful in guiding authors' revisions and in helping them understand disciplinary expectations and conventions. However, to improve their work, authors must critically analyze the comments and decide what advice to follow and what advice to ignore. Bakanic et al. (1989) characterize peer review as "research collaboration" with peer review functioning as "expert feedback"; they hint at the developmental nature of reviewer comments, particularly for manuscripts that were eventually accepted (p. 651). Some editors advocate developmental editing and view peer review as a mentoring tool, particularly for emerging scholars (e.g., Burbules, 2014; Gale, 1998).

Scholars seem to agree that peer reviewers' detailed comments influence authors' revisions and editors' publication decisions (e.g., Donovan, 2011; Lay, 2004; Marsh & Ball, 1989), which suggests that, as an editorial practice, peer review does indeed shape content and knowledge. For example, peer review influences what content is included or excluded, and how data, ideas, and arguments are presented. Many editors find the reviewers' evaluative comments more useful

than their publication recommendations (Goodman et al., 1994). According to Lay (2004), the reviewers' "comments have more impact on the nature and quality of the manuscripts published and the ways in which the manuscripts are revised" than the reviewers' formal guidelines (p. 112). However, we do not know precisely how these comments are shaping manuscript revisions and subsequently the nature of technical communication scholarship.

In terms of content development, most journal editors "make a significant contribution to the conversation but do not direct or control it: the members of the community do" (Gale, 1998, p. 199). Some scholars perceive "the editors' role in shaping the disciplinary discourse" to require active intervention "in the discourse by encouraging and promoting, or even by recruiting articles that lead the disciplinary conversation in new directions," while other scholars believe journal content "should reflect the interests of the readers" (Gale, 1998, p. 202). These different content development philosophies align with curation models (e.g., special issues and invited articles) and filtration models (e.g., gatekeeping) of content acquisition (Vardi, 2017).

In journals with developmental editing cultures, editors are "wary of rejecting a potential gem" (Eden, 2008, p. 243). Likewise, some reviewers are hesitant to recommend manuscripts be rejected and instead of risking mistakenly rejecting a paper (an action similar to a Type I error, i.e., mistakenly rejecting a null hypothesis), they recommend revision (Chrisman, Sharma, & Chua, 2017). In these instances, a comparison of the reviewer's qualitative feedback to the reviewer's publication recommendation would probably reveal the hesitations. In their confidential remarks to editors, reviewers sometimes acknowledge their reluctance to reject manuscripts. Either way, reviewers' actions or inactions along with their qualitative feedback impact manuscript content and disciplinary knowledge (Burbules, 2014).

We know these mediated peer review conversations occur as a part of technical communication journals' standard operating procedures; however, we do not know the particular details of the conversations or how influential they are in respect to content development and the shaping of disciplinary knowledge. We know reviewers disagree, but we do not know how reviewers' often-conflicting comments and publication recommendations affect content. Moreover, we do not know what authors do with the evaluative feedback, how they reconcile conflicting advice, or how peer review shapes their revisions. Nor do we know how accepted manuscripts are further shaped by editing. In short, we do not know how these editorial practices shape technical communication scholarship.

To address this research gap, I conducted a mixed-method empirical study designed not only to help demystify scholarly publication practices but also to help us understand how peer review works as a content-shaping mechanism in technical communication journal scholarship. I outline my research methods in the next chapter.

Chapter 3

Methods

In this chapter, I summarize my mixed-methods approach to conducting this empirical study of technical communication scholarship. I begin by recapping my research questions and hypotheses. Next, I outline my methodology and methods, including the theoretical framework and variables. Then, I describe the sampling protocols and blinding procedures for the publication-process artifacts studied. Finally, I explain my coding procedures and data analysis methods.

Research Questions & Hypotheses

My mixed-methods inquiry begins with the following research questions and hypotheses:

Research Questions

RQ_1: How well do the content and structure of the reviewers' evaluative comments align with the journal's reviewers' guidelines or scoring rubrics?

RQ_2: What role(s) do journal reviewers in the technical communication field play?

RQ_3: In what ways do reviewers' publication recommendations and evaluative comments shape editorial decisions and content development?

Hypotheses

H_1: No significant difference exists between the content of the reviewers' evaluative comments and the content of the journal's reviewers' guidelines or scoring rubrics.

H_2: No significant difference exists between the number of reviewer comments associated with higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) and the number of reviewer comments associated with lower level concerns (e.g., grammar, mechanics, style, citations).

H_3: No significant difference exists between reviewers' publication recommendations and editors' publication decisions.

H_4: For each manuscript, no significant difference exists between the types of problems each reviewer identifies.

H_5: For each manuscript, no significant difference exists between the number of manuscript problems each reviewer identifies.

H_6: For each manuscript, no significant relationship exists between the reviewer's publication recommendation and the number of manuscript problems the reviewer identified.

Methodology & Methods

In this section I discuss my methodology and methods. Methodology refers to the rationale for using a particular theoretical framework and system of methods for answering the study's research questions, whereas methods refer to the tools used for collecting the research data.

Methodology

This empirical study on editorial peer review uses a mixed-methods research design to provide a fuller understanding than what could be learned from a quantitative or a qualitative study alone. Deductive in nature (reasoning from general to specific), quantitative research designs are useful for answering questions of whether or to what degree something happened and for testing hypotheses. In contrast, qualitative research designs, which are usually inductive (reasoning from specific to general), are useful for answering who, what, when, where, why, or how.

In this study, the quantitative and qualitative design elements are mixed in a manner that Cresswell (2014) refers to as convergent parallel mixed methods, which means the quantitative research and the qualitative research are conducted in tandem and the data are merged "to provide a comprehensive analysis of the research problem" (p. 15). The findings are integrated and contradictions are explained or examined further. In mixing quantitative and qualitative approaches, the perceived advantages and disadvantages (e.g., objectivity, subjectivity,

reliability, validity, generalizability) of each approach are counterbalanced within the theoretical framework.

Theoretical framework & variables.

As discussed in Chapter 1, genre theory is used as an analytic framework for examining how peer review shapes technical communication journal scholarship. Genre theory provides a conceptual model for measuring the study's variables and examining the relationship between them. This study examines the primary and secondary variables shown in Tables 3.1 and 3.2.

Table 3.1

Primary Variables Studied

Primary Variables
Reviewers' publication recommendations (independent variables)
Reasons for reviewers' publication recommendations (mediating variables)
Editors' publication decisions (dependent variables)

Table 3.2

Secondary Variables Studied

Secondary Variables
Manuscript evaluation criteria, i.e., reviewer guidelines or scoring rubrics (control variable)
Review content (independent variables)
Types of problems reviewers identified in manuscripts (mediating variables)
Severity of problems reviewers identified in manuscripts (mediating variables)
Quantity of problems reviewers identified in manuscripts (mediating variables)
Associate editors' publication recommendations (mediating variable)
Reviewer role (mediating variable)

Publication-Process Artifacts

The area of inquiry for this research study is technical communication scholarship as represented by data collected from existing records that were created as part of the publication process at technical communication journals. I refer to these records as publication-process artifacts; the artifacts may include authors' manuscripts, authors' cover letters, authors' response letters, journals' guidelines for authors and reviewers, journals' style guides, reviewers' scoring rubrics, reviewers' recommendation reports, editors' decision letters, and published articles.

All artifacts collected for this project are related to manuscripts that have been through the publication decision process; in other words, no decisions were pending. Appropriate permissions were obtained from the journals that provided artifacts for use as research data in this IRB-approved research study. The artifact data was anonymized and has been reported in aggregate and descriptive forms only. No information that might reveal the identities of authors, reviewers, or manuscripts is included in the results.

In the next section, I outline my sampling methods and criteria and justify my sample size.

Sampling methods and sample sizes.

The literature outlines several sampling methods for studies of editorial peer review, ranging from random samples to samples selected by journal impact factors. This study used a purposeful sampling method that tempered the realities of limited resources with the difficulties of accessing occluded publication-process genres and the labor-intensive data collection process, which involved editorial staff spending several weeks compiling and redacting data. A purposeful sample is selected for its potential to answer the research questions through the use of representative data (Cresswell, 2014). I describe my purposeful sampling method below.

Based on the literature (e.g., Davy & Valecillos, 2011; Lowry, Humpherys, Malwitz, & Nix, 2007; Smith, 2000) and my knowledge of the field, I first selected five major peer reviewed journals in the technical communication (TC) field that were appropriate for answering the study's research questions (see Table 3.3).

Table 3.3

Major Technical Communication Journals

Journal	**Impact Factor***	
(published quarterly)	**2016**	**2017**
IEEE Transactions on Professional Communication	0.899/2.184	0.84/0.756
Journal of Business and Technical Communication	1.020/1.062	0.87/0.750
Journal of Technical Writing and Communication	0.367	0.64
Technical Communication	0.621	0.77
Technical Communication Quarterly	1.529	1.12

*As reported by Scimago Journal and Country Rank (www.scimagojr.com) for the respective years; where two impact factor numbers are listed, the first number is from Scimago, and the second number comes from Thomson Reuters' Journal Citation Reports (JCR) as reported on the individual journal's website.

Then, I estimated the potential number of manuscripts[3] by looking at the publication history of each journal. Table 3.4 shows the approximate number of articles published by these journals between 2007 and 2018. (The range of years is somewhat arbitrary because the information had originally been compiled for other purposes.) Extrapolating from these publication figures, I estimated that between 2,500 and 12,000 unique manuscripts had been submitted to one or more of these journals during that 12-year period; the lower figure assumes an average acceptance rate of 50% and the higher figure assumes an average acceptance rate of 10%.

[3] I refer to the data sample in terms of manuscripts because they are the primary publication process artifact to which all of the other artifacts are connected; however, the manuscripts were not analyzed in this phase of the long-term study. This phase of the study focuses on the journals' guidelines for reviewers, reviewers' scoring rubrics, and reviewers' reports. The manuscripts and other remaining artifacts will be analyzed in follow-up studies.

The literature provides limited guidance on sample sizes for studies of peer review. Olson et al. (2002) calculated that they would need about 750 manuscripts to obtain significant results in their study on bias toward negative research results. Bornmann, Mutz, and Daniel's (2010) meta-analysis indicated that previous studies of peer review in the medical field used samples ranging from 28 to 1,983 manuscripts, with an average of 311 manuscripts. In contrast, Kravitz et al. (2010) analyzed more than 6,000 manuscripts submitted to a medical journal; and a more recent longitudinal study analyzed approximately 13,000 manuscripts and 31,000 corresponding peer review reports (Sposato et al., 2014).

Table 3.4

Number of Articles Published in Major TC Journals by Year (2007–2018)

Journal / Year	IEEE	JBTC	JTWC	TC	TCQ	Yearly Totals*
2007	28	23	24	31	18	124
2008	26	20	16	24	19	105
2009	26	17	26	25	19	113
2010	31	17	26	19	20	113
2011	25	18	24	15	20	102
2012	23	16	24	19	19	101
2013	20	17	23	17	18	95
2014	18	15	24	17	18	92
2015	22	17	27	19	16	101
2016	28	16	22	20	21	107
2017	26	16	23	22	25	112
2018	26	14	22	22	24	108
Running Totals	**299**	**206**	**281**	**250**	**237**	**1273**

***** These numbers include some editors' introductions and special issues.

Some of the larger studies randomly sampled 10% of the available manuscripts collected over periods extending up to 30 years—a sampling strategy often used when the amount of data exceeds available resources (Koerber & McMichael, 2008). Using a process much like stratified sampling, Gonzalez (2006) supplemented random sampling with purposeful sampling to achieve

data saturation over a broad time span. According to Thayer, Evans, McBride, Queen, and Spyridakis (2007), 384 units of observation are needed to obtain results with a 95% confidence level.

Although large sample sizes, such as those used in other fields, improve a study's reliability and validity, the data in Table 3.4 suggest that those large sample sizes not realistically achievable for studies of the technical communication scholarship. Even a more modest sample size of 384 units presents a challenge, especially after factors such as journals' archival practices and other sampling criteria are factored in.

Weighing those limitations, I requested 30 initial[4] manuscript submissions (and the corresponding publication-process artifacts) from the five purposefully selected TC journals listed in Table 3.3 for a maximum combined sample of 150 manuscripts, which equates to approximately half the average sample size used in other studies (Bornmann, Mutz, & Daniel, 2010). I provided the editors with my study's sample criteria, which resemble those used in other studies (e.g., Bakanic et al., 1989). I asked that several categories of manuscripts be excluded from the sample:

- Bench rejects (manuscripts rejected without being sent for review)
- Resubmissions (previously reviewed, revised manuscripts resubmitted as new manuscripts)
- Revisions (previously reviewed, revised manuscripts resubmitted as revisions)[5]
- Manuscripts designated for special issues[6] (manuscripts typically prescreened by special issue editors)

[4] The publication-process artifacts were collected for a multiphase study of TC scholarship. To fully understand how editorial peer shapes TC scholarship, I needed a complete data set that would allow me to study the publication process from start to finish. This multiphase study examines reviewers' first impressions of initial manuscript submissions and traces the evolution of the manuscripts at various points in the publication process (i.e., through authors' revisions and the editing of accepted manuscripts).

[5] Resubmissions and revisions will later be analyzed if they accompanied the original manuscript submissions.

- Manuscripts not sent out for review (e.g., manuscripts reviewed internally or exempt from review)
- Manuscripts sent to only one reviewer
- Manuscripts with missing reviewer reports[7]
- Manuscripts pending publication decisions.

After consulting their editorial boards and/or legal departments, two of the five editors permitted me to study publication-process artifacts from their respective journals. The other editors suggested alternative data collection methods that I will use in follow-up studies.

Journal 1 provided artifacts associated with 34 manuscripts; four of those were excluded from the study because they did not meet the sampling criteria (two manuscripts were book reviews, one manuscript was a bench rejection, and another manuscript had only one reviewer).

Journal 2 provided artifacts associated with 35 manuscripts; three of those manuscripts were excluded because they appeared to be revised manuscripts that had been resubmitted as new manuscripts. One manuscript from Journal 2 appeared to be a revision of another manuscript in the sample; the artifacts from those two manuscripts were combined in order to represent the full publication history of the initial manuscript.

The final sample comprised 61 manuscripts—30 manuscripts from Journal 1 and 31 manuscripts from Journal 2.

This purposeful sample is a convenience sample in that I could not access all the publication-process artifacts in the population that met my sampling criteria due to the confidential nature of occluded genres. The artifacts may not be representative of the full set and may be biased in

---

[6] Acceptance rates for special issues tend to be higher than regular issues (Lay, 2004), which suggests that something about the decision-making and/or publication processes differs and therefore should be analyzed separately.

[7] I received some manuscripts that did not meet all of the sampling criteria. Given the difficulty in obtaining the materials, I did use a couple of manuscripts that were sent to one reviewer or that had missing reports for some analyses and noted as much in the results and limitations.

ways that aren't apparent (Krippendorff, 2019). Nevertheless, within the context of occluded genres, I would argue that the available sample provides rich data that offset the potential problems with representativeness and generalizability (Koerber & McMichael, 2008).

Blinding.

Editorial staff at the respective journals redacted identifiable details in the publication-process artifacts such as authors' names, reviewers' names, and institution names. I further anonymized the artifacts prior to analysis. For example, the manuscript ID numbers were replaced with randomly generated numbers and any identifying details in the documents were blinded (e.g., the actual journal name was replaced with JOURNAL NAME, JOURNAL #, or J# (to reflect reviewer's use of abbreviations in their discourse) in a manner that fit the context. Other potentially identifiable information that was not needed for analysis purposes was relabeled in a similar fashion or redacted with XXX. Details, such as manuscript topics, were retained for analysis so that coders could understand the context of the reviewers' comments; however, any details that could potentially be connected to authors were blinded, redacted, or stop listed for word frequency counts.

The data from the publication-process artifacts are reported in aggregate and descriptive forms in accordance with my nondisclosure agreements. To further protect the anonymity of the authors, reviewers, and editors, I have not reported the precise time period from which the manuscripts (and the associated publication-process artifacts) were sampled, merely that the sample includes manuscripts submitted from various periods within the last 10 years. Although I cannot provide certain details, including excerpts of the text, I have, to the extent possible, been transparent in explaining my methods because transparency improves validity, reliability, replicability, trust, and credibility (Cresswell, 2014).

Methods

To begin to understand how peer review shapes technical communication journal scholarship, I used two textual analysis methods: structural analysis and content analysis. I briefly describe the methods in this section and provide additional details in the Procedures and Analyses section.

Textual analysis of publication-process artifacts.

Textual analysis is an appropriate—yet labor intensive—method for studying communication phenomena, such as anonymous peer review, that cannot be observed directly or no longer exists (Boettger & Palmer, 2010; MacNealy, 1999). Textual analysis is a broad term that encompasses discourse analysis, narrative analysis, genre analysis, and structural analysis—among other qualitative methods—as well as content analysis (CA), a method that can take either a qualitative or a quantitative form (Lockyer, 2012). Content analysis and structural analysis were selected as methods because the study's research questions explicitly address content and structure.

Structural analysis.

Structural analysis is used to describe genres (e.g., lengths, metawriting, topic strings) and to identify document design elements (e.g., headings, numbered or bulleted lists) and organizational patterns (MacNealy, 1999). The method overlaps with genre analysis, particularly in respect to Swales' (1990) analysis of rhetorical structures and Gosden's (2001, 2003) response pattern norms.

Swales (1990) refers to rhetorical structures as moves; his four-move genre analysis model consists of a summary of judgment, an outline of the article, a list of criticisms, and a conclusion. Similarly, Gosden's response pattern norms include formulaic opening remarks, point-by-point replies, and closing remarks. These genre analysis models were used along with the reviewers'

guidelines and rubrics as benchmarks for evaluating the structure of reviewers' reports; the model elements were incorporated into part two of the coding form (Appendix C).

Content analysis.

Content analysis (CA) was used in this study because it enables a researcher to make "replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use" (Krippendorff, 2019, p. 24). The research method involves "tallying the number of specific communication phenomena in a given text . . . and then categorizing those tallies into a taxonomy" (Thayer et al., 2007, p. 268).

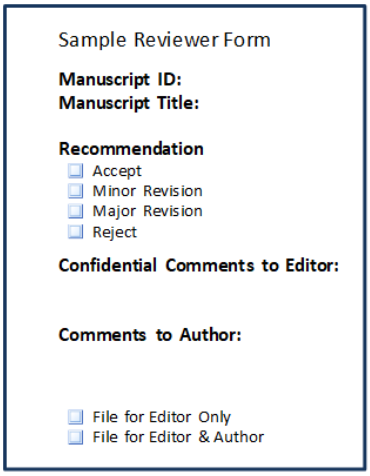
In its qualitative form, CA uses emergent codes—an inductive measurement—to describe the text's latent, or underlying, meaning; in its quantitative form, CA uses a priori, or predetermined, codes—a deductive measurement—to describe the text's manifest, or surface, features (Saldaña, 2016; Thayer et al., 2007). The role of the researcher also differentiates the two forms. "In quantitative content analysis, the empirical process is independent of the particular scholar; in qualitative or critical message analyses, it is not" (Neuendorf, 2017, p. 9). Boettger and Palmer (2010) contend quantitative content analysis is "more powerful than surveys and interviews because of its unobtrusive nature and its lack of reliance on subjective perceptions" (p. 346). (Of course, quantitative research is not entirely objective because humans are involved in the process—humans design the research study, the computer algorithms, etc.)

Since the technical communication literature tends to conflate the two forms of CA (Boettger & Palmer, 2010), I will explicitly state that this study features the quantitative form of content analysis; unless noted otherwise, subsequent discussions of content analysis refer to quantitative content analysis.

Given this study's methodology of convergent parallel mixed methods, the content analysis and structural analysis were performed simultaneously as part of the coding procedures that are described in more detail in the Procedures and Analyses section.

Procedures & Analyses

The transcribed and blinded publication-process artifacts were entered into NVivo 12.4 Pro to code for specific purposes (e.g., content, review structure, error severity). NVivo is used primarily for coding qualitative data; however, the software includes quantitative tools that can be used for both qualitative and quantitative content analysis (Neuendorf, 2017). Artifacts such as reviewer rubrics were entered as individual files. However, the reviewer report forms (Figure 3.1) were split into two audience-segmented files (one file for confidential comments to the editor—the editor (ED) file—and one file for comments to the author—the author (AU) file) because reviewers may provide confidential comments to the editor and/or comments to the author; this feedback can be entered in the text fields of reviewer forms, submitted as file attachments, or both (see Figure 3.2). (Editors' decision letters usually include the reviewers' comments to the authors; the reviewers' comments to the editor remain confidential.)


Sample Reviewer Form

Manuscript ID:
Manuscript Title:

Recommendation
- ☐ Accept
- ☐ Minor Revision
- ☐ Major Revision
- ☐ Reject

Confidential Comments to Editor:

Comments to Author:

- ☐ File for Editor Only
- ☐ File for Editor & Author

Figure 3.1: Sample reviewer report form.

When file attachments accompanied the reviewer reports, the file text was analyzed with the text entered into the report form's comment field (see Figure 3.2). Any annotated file attachments (i.e., PDF reviewer proofs with inserted comments or marked-up text) were transcribed and the locations of each comment were recorded (e.g., page, paragraph, line, or section number). As transcription artifacts, these location notations were treated as stop words when running word frequency count queries; several reviewers used similar notations within the review form itself. The origins of these notations (i.e., author or researcher transcription) were accounted for during the manual coding process.

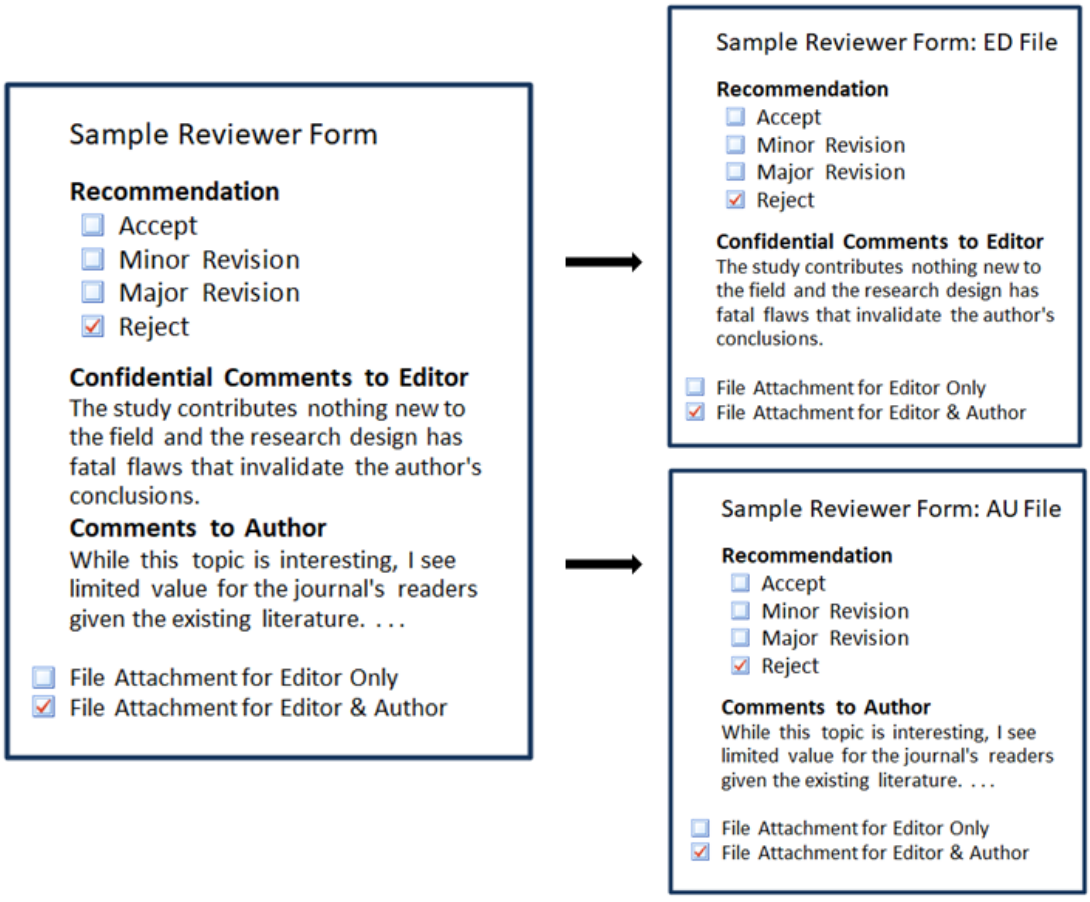


Figure 3.2: Audience-segmented reviewer files.

For analysis purposes, I created blank placeholder files when the reviewer provided no comments; some reviewers provided comments to the editor but not to the author or vice versa, while other reviewers provided comments to both. Duplicate content was flagged; for example, if a reviewer uploaded a file with 500 words of comments and pasted those same 500 words into both the "confidential comments to the editor" section and the "comments to the author" section, then each coder coded those 500 words once only. (All recorded word counts represent the number of words in the blinded version of the texts.)

From the NVivo files, I created NVivo cases. NVivo uses cases to represent units of observation, meaning the "specific item measured at an individual level" (Thayer et al., 2007, p. 270). In this study, the units of observation were reviewer reports, segmented by audience (editor or author). Metadata were entered for each case in the form of NVivo case attributes (e.g., journal, comment audience, review word count, presence of file attachment, manuscript version, reviewer number, reviewer publication recommendation, editor decision, final disposition of manuscript). These metadata were also recorded in an Excel spreadsheet for separate statistical analyses that are described in more detail in the Statistical Analysis section.

Textual Analysis

After preparing the text for analysis, I developed the coding scheme, codebook, coding form, and custom dictionary. Then I established procedures for pilot testing, coding, and analysis. These coding tools and research procedures are discussed in the next sections.

Coding schemes.

The CA method requires mutually exclusive, predetermined coding categories, and best practices dictate that researchers should use existing coding schemes when available (Thayer et al., 2007). I could not find a relevant coding scheme from the TC literature. Rather than devise a custom

coding scheme from scratch, I looked to other fields. I modified a coding scheme created by Bornmann, Nast, and Daniel (2008), who analyzed 46 studies on manuscript evaluation criteria for fields ranging from the social sciences to chemistry. From the literature, Bornmann et al. identified 542 unique criteria and reasons for accepting or rejecting manuscripts; they sorted these assessment criteria into nine categories:

1. Relevance of contribution
2. Writing/presentation
3. Design/conception
4. Method/statistics
5. Discussion of results
6. Reference to the literature and documentation
7. Theory
8. Author's reputation/institutional affiliation
9. Ethics.

These manuscript assessment criteria are applicable to research articles in most, if not all, fields, including technical and professional communication. The categories are similar to ones used by Bakanic et al. (1989): topic, theory, review of the literature, design, data, sample, measurement, analysis, results, style, ad hominem, and general. More importantly, the categories fit within the framework of genre theory, and, with minor modifications, the categories paralleled my research questions and hypotheses.

I modified Bornmann, Nast, and Daniel's (2008) criteria by omitting category 8 (author's reputation/institutional affiliation) because it was not relevant to this study, which examines publication-process artifacts from journals that use double-blind peer review—reviewers would not have access to this type of information. To better align the categories with my research questions and hypotheses, I divided category 2 into two writing/presentation categories: higher level concerns and lower level concerns. I added an "Other" category not only to ensure that all data could be categorized but also to allow for unexpected results.

The revised categories follow:

1. Relevance of contribution
2. Writing/presentation (higher order)
3. Writing/presentation (lower order)
4. Design/conception
5. Method/statistics
6. Discussion of results
7. Reference to the literature and documentation
8. Theory
9. Ethics
10. Other.

Gosden (2003) used complete sentences as the unit of analysis—the "general idea or phenomenon being studied" (Thayer et al., 2007, p. 270); however, knowing that the literature on feedback comments recommends sandwiching criticism between praise, I anticipated encountering compound and complex sentences that required multiple codes—a coding practice criticized in the technical communication field (Boettger & Palmer, 2010). Therefore, for the CA, the unit of analysis ranged from a word to several sentences.

Coder training.

After agreeing to the nondisclosure terms, two research assistants with backgrounds in writing and editing were trained to code the data using the codebook (see Appendix C). The male research assistant had taken a rhetoric class, and the female research assistant had taken a technical writing class; since neither assistant works in the technical communication (TC) field, core TC concepts and theories were explained to both assistants.

As part of their training, the research assistants read portions of two articles (i.e., Bornmann, Nast, & Daniel, 2008; and Bornmann, Weymuth, & Daniel, 2010) that detail the original coding scheme. The variables of interest in this dissertation study were discussed, but to minimize bias and "demand characteristic" (coding to please the researcher), the research assistants were not provided with the study's research questions or hypotheses (Neuendorf, 2017, p. 158).

Codebook & coding form.

Research assistants were provided with a codebook (see Appendix C) that is based on a coding scheme developed by Bornmann, Nast, and Daniel (2008). The codebook includes directions for coding and using the two-part coding form as well as definitions and examples of each code. The examples come primarily from excerpts of reviews published in Lay's (2004) *Technical Communication Quarterly* article.

For analysis purposes, the coding form was split into two parts. Part one of the coding form (see Appendix C) provides space to tally each code identified in the publication-process artifacts. The elements coded are shown in Figure 3.3. Part two of the coding form integrates elements of Swales' (1990) and Gosden's (2001, 2003) genre analysis models, which were discussed in the Structural Elements section of this chapter. The second part of the form also features elements of van Rooyen, Black, and Godlee's (1999) review quality instrument (RQI), which is discussed in the next section. While these instruments had been previously validated for other uses, to the best of my knowledge, neither had been used in studies of technical communication. As described in the next section, the RQI was modified to fit the research study.

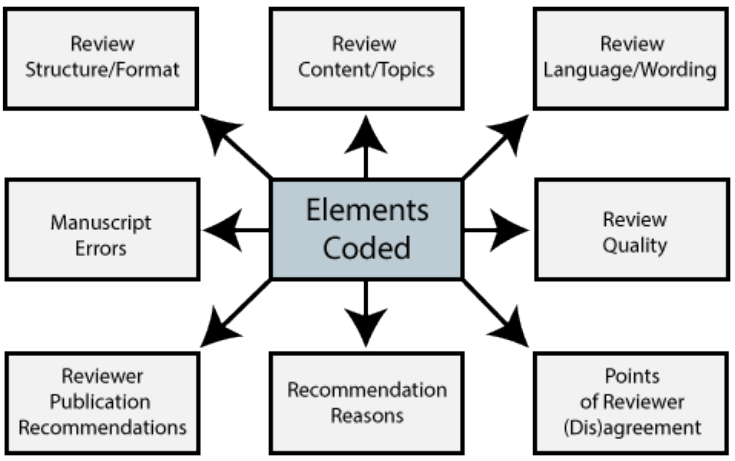


Figure 3.3: Elements coded.

Review quality instrument.

The coding form integrates a modified version of van Rooyen, Black, and Godlee's (1999) review quality instrument (RQI), an instrument that measures "the extent to which a peer reviewer has considered key aspects of a manuscript" (p. 628). I modified the RQI to align with Bornmann, Nast, and Daniel's (2008) categories, which are discussed in the Textual Analysis section of this chapter. Specifically, I split question 3 of the RQI into two questions to differentiate the research study design from the methods, and I split question 4 of the RQI into one question about high-level writing concerns and one about low-level writing concerns. The original wording of question 4 was problematic in that it rated four elements (writing, organization, tables, and figures) and did not allow for the possibility of elements being rated differently—for example, one element being rated "not at all" and one as "extensive." Minor changes in the wording of questions and scale labels were also made to fit the context of this study.

Pilot testing.

The codebook was pilot tested on a representative sample of reviewers' reports and refined as needed to resolve disagreements in applying codes. While it is common to use 5% to 10% of a data set for pilot testing—a time-intensive process for large data sets—Thayer et al. (2007) found that 5 pages of text were adequate for calculating intercoder reliability in their study. For this study's intercoder reliability sample, I used a similar data set from a previous unpublished study (N = 28 manuscripts [54 reviewer reports]). The two research assistants pilot tested the predetermined codes on a stratified sample (n = 4 [reviewer reports]) drawn from that unpublished study.[8] The sample was stratified by publication recommendation (e.g., accept,

[8] UMCIRB 17-001261.

major revision) to ensure that the coders were familiar with the full range of reviewer feedback. Pilot testing concluded when the Cohen's kappa showed an interrater reliability rate of at least 70% (Boettger & Palmer, 2010). The intercoder agreements ratings are shown in Table 3.5.

Table 3.5

Pilot Testing Intercoder Agreement Levels

Agreement Ratings	PI & Coder 3	PI & Coder 13	Coder 3 & Coder 13
Percent agreement	90%	80%	90%
Cohen's kappa	.89	.78	.89
Confidence interval	95% (90 ± 13.5)	95% (80 ± 18.6)	95% (90 ± 13.5)

Coding procedures.

Due to limited resources, I, as the principal investigator (PI), coded all of the data. Following training and pilot testing, the research assistants independently coded 20% of the reviewers' reports connected to the 61 manuscripts collected for this study. The PI and the research assistants used separate coding forms for each document analyzed (see Appendix C). Reviewer comments to authors (and any accompanying file attachments) were coded first so that the PI and the research assistants were initially interpreting the text as presented to the authors—disciplinary knowledge-making from the perspective of the authors. Afterwards, the comments to the editors were coded, providing a more complete perspective on disciplinary knowledge-making.

As part of the coding process, the PI and the research assistants counted the number of unique problems identified in each review. For example, if the author misspelled the same word multiple times, the misspelling was counted as one problem; if the author misspelled five different words, the misspellings were counted as five problems (see part one of Appendix C).

The final intercoder reliability from the training session was used as the baseline interrater reliability (IRR); see Table 3.5. The IRR was reassessed after both research assistants had coded 10 reviews so that coder recalibration or further refinements to the codebook could be made if needed (see Table 3.6). The research assistants, who completed the first part of the coding form only, recorded notes about reviewers' comments that they found difficult to code; these notes were used to help resolve coding disagreements and to refine the codebook. Calculating the IRR for each variable coded also helped pinpoint areas of disagreement between coders.

Table 3.6

First Check of IRR

	PI & Coder 3 (*n* = 10)		**PI & Coder 13 (*n* = 14)**	
Category	**PA_O**	**kappa**	**PA_O**	**kappa**
Contribution	90%	.89	85.7%	.8517
Writing/higher	70%	.6774	92.8%	.9249
Writing/lower	70%	.6939	85.7%	.8573
Design/conception	100%	1.0	71.4%	.711
Methods/statistics	70%	.6842	100%	1.0
Results	80%	.7826	57.1%	.5643
Lit/documentation	90%	.89	92.8%	.9249
Theory	80%	.7872	92.8%	.9273
Ethics	40%	.4	85.7%	.8555
Other	90%	.9	64.3%	.6411
Overall IRR	78%	.7788	82.9%	.8286
Confidence interval	95% (78.0 ± 9.29)		95% (82.83 ± 8.15)	

The IRR was calculated again after the coders had completed coding a subset of the sample (i.e. at least 20% of the reviews for the initial manuscripts); Table 3.7 shows the IRR for those coded reviews. Coder 3 did not complete the assigned reviews in time to include those IRR calculations; they will be included in follow-up studies.

Due to limited resources, I completed the second part of the coding forms (see part two of Appendix C). The second part of the form was used to record details about (A) reviewer recommendations, (B) review structure, (C) alignment with manuscript evaluation instruments (i.e., reviewers' guidelines/rubrics), and (D) review quality. Most of these elements required evaluation of surface features of the text or required evaluation of textual elements that could be triangulated through computer assisted coding or with other data. For example, during the publication process, the editors had rated the quality and timeliness of many reviews. The research assistants coded textual elements that required some interpretation.

Table 3.7

Second Check of IRR

	PI & Coder 13 (*n* = 34)	
Category	**PA_O**	**kappa**
Contribution	76.5%	.7706
Writing/higher	79.4%	.7988
Writing/lower	79.4%	.8003
Design/conception	79.4%	.7968
Methods/statistics	79.4%	.7989
Results	79.4%	.8016
Lit/documentation	79.4%	.8002
Theory	79.4%	.7975
Ethics	79.4%	.8009
Other	79.4%	.8016
Overall IRR	78.8%	.7879
Confidence interval	95% (79.11 ± 34.44)	

Statistical Analysis

To be clear, I did not evaluate or analyze the merits of the manuscripts; rather, I analyzed the reviewers' evaluations of those manuscripts and how those evaluative comments influence

content development and disciplinary knowledge-making. Specifically, I analyzed the reviewers' publication recommendations (e.g., accept, reject) to determine the extent to which reviewers (dis)agree. I looked for patterns in the reviewers' publication recommendation decisions using quantitative content analyses, comparative content analyses, computer-aided text analysis, and descriptive and inferential statistical analysis. All of the analyses were based on my coding of the data.

I collected several types of data that were recorded in an Excel spreadsheet and analyzed in JMP Pro 14.2, SPSS 25, and NVivo 12.4 Pro, all of which are commonly used data analysis programs. JMP is easier to use than SPSS and tends to be better for creating charts, but SPSS provides a wider selection of statistical tests. NVivo is limited in its graphical rendering of large data sets, but it generally works well for frequency counts, word clouds, and correlation analysis; however, it does not support custom dictionaries. For additional computer-aided text analysis (CATA), I used Yoshikoder 0.6.5.0, XML-based freeware recommended by Neuendorf (2017) that does support custom dictionaries but has limited features (Lowe, 2019).

Descriptive statistics.

I ran descriptive statistics (e.g., mean, mode, standard deviation, variance, range), checked the data for normalcy, and visually inspected the data to detect any problems in the data set that would affect inferential analysis. The descriptive statistics indicated that the most of the data was not normally distributed; therefore, nonparametric tests would be needed for the inferential statistical analyses.

Inferential statistics.

Inferential statistics (e.g., nonparametric tests such as the Wilcoxon signed rank test, Wilcoxon each pair test, and Tukey-Kramer HSD), and other analyses[9] were performed as warranted by the data. The statistics included Cohen's kappa and rank order correlation. The Cohen's kappa is used to measure agreement between categorical variables (e.g., accept, reject) and can account for chance agreement between reviewers; the rank order correlation (Wilcoxon/Kruskal-Wallis rank sums) was used to compare the reviewers' publication recommendations to the editor's publication decision. The Tukey-Kramer test is equivalent to a paired t test and is designed to protect the overall error rate; the test can help identify the significant pairs of variables (e.g., Reviewer 1 versus Reviewer 2's publication recommendation of a specific manuscript).

Inferential tests used a *p* value of .05 or less as an indicator of statistically significant results. The data were coded for statistical analysis (e.g., Reviewer decision: 1 = Reject; 2 = Major revision; 3 = Minor revision; 4 = Accept[10]). In some instances, the data were binned for analysis based on the data distribution; for example, the reviewers' other-than-reject publication recommendations (i.e., major revision, minor revision, and accept) were combined into one group and the audience-segmented reviewer report files (i.e., the author and editor parts of the reviewer form) were analyzed as one file.

Mean priority scores.

The reviewers' publication recommendations were analyzed using a hybrid version of Sposato Ovbiagele, Johnston, Fisher, and Saposnik's (2014) mean priority score and Eberley and

[9] Other studies (e.g., Bornmann, Mutz, & Daniel, 2010; Jackson et al., 2011; Kravitz et al., 2010) have used intraclass correlation (ICC) and Pearson produce-moment correlation (r) to determine interrater reliability (IRR) and/or have used Fisher Z-transformed correlation to correct for scale issues.

[10] This numerical coding aligns with Sposato, Ovbiagele, Johnston, Fisher, and Saposnik's (2014) reviewer decisions rankings, which are used to calculate manuscripts' mean priority scores and predict editors' final publication decisions.

Warner's (1990) agreement on recommendation score. Eberley and Warner's (1990) scale is used to calculate perfect agreement and perfect disagreement between reviewers; however, their numeric system (i.e., 1 = Accept and 4 = Reject) is opposite to Sposato, Ovbiagele, Johnston, Fisher, and Saposnik's (2014) system of representing manuscript quality (i.e., 1 = Reject and 4 = Accept). I realigned Eberley and Warner's scale with Sposato et al.'s numeric system and adjusted the calculations as needed (i.e., I used absolute values). For both metrics, the cumulative scores are divided by the number of reviews to determine the score; although problematic mathematically, these scales are useful for representing the data patterns visually.

Structural Analysis

Based on data collected from part two of coding form, I analyzed the structure of the reviewers' reports in terms of genre characteristics, rhetorical moves (Swales, 1990), response patterns (Gosden, 2001, 2003), document design elements, and organizational patterns. I compared those textual structures to the structure of the journals' reviewers' guidelines and rubrics.

Content Analysis

I conducted a content analysis (CA) of the evaluative comments in reviewers' reports to determine (1) what aspects of the manuscripts peer reviewers are evaluating, (2) what points and to what extent peer reviewers (dis)agree, (3) the number, types, and severity of problems mentioned in reviewers' reports, and (4) how peer review affects content development.

Comparative content analysis.

Following the quantitative content analyses of the reviewer reports for each manuscript, I performed a comparative content analysis (CCA), a term Bornmann, Mutz, and Daniel (2010) introduced in the future research section of their article. By CCA, I simply mean that I have

compared the results of two quantitative content analyses; for each manuscript, I compared the content analysis of one reviewer report with that of the other reviewer report (that is, I compared the codes associated with reviewer one with the codes associated with reviewer two). This method is similar to correspondence analysis in that it compares sets of categorical variables; however, correspondence analysis is not appropriate for hypothesis testing (Lam, 2016).

Specifically, the CCA examines

1. the types and quantities of manuscript problems that both reviewers discussed (i.e., the points of agreement)
2. the types and quantities of reviewers' contradictory comments (i.e., the points of disagreements)
3. the elements of the manuscript or the research study that were discussed, namely the
    - importance of the research question
    - relevance of the contribution
    - originality of the paper
    - higher order writing/presentation concerns (e.g., organization)
    - lower order writing/presentation concerns (e.g., grammar)
    - strengths/weaknesses of the study design
    - strengths/weaknesses of the methods
    - author's discussion/interpretation of results
    - literature/documentation
    - theory
    - ethics

Data for the first two aspects of the CCA were obtained from part one of the coding form. Data for the remaining aspects of CCA were obtained from part two of the coding form.

Computer-aided text analysis.

Each review was compared to reviewers' guidelines and rubrics using a combination of manual coding and computer-aided text analysis (CATA). Although manual coding involves a degree of subjectivity, human coders are superior to computers in quantifying tone, style, and nuanced

meaning (Thayer et al., 2007). Computers are effective tools for identifying and quantifying the presence of specified terms.

Custom dictionary & stop words.

The CATA required the use of several custom dictionaries (see Appendix D). I created one custom dictionary using terms compiled from technical and professional communication journals' reviewers' guidelines and rubrics, publishers' guides (e.g., Sage Journals, n.d.; Taylor & Francis, 2019), and from ethical guidelines published by the Council on Publication Ethics (COPE Council, 2013; 2017, September).

The 761-term dictionary represents a refinement of the 1,000 most frequent stemmed words (e.g., alert, alerted, alerting) of at least three letters that appeared in the compiled text; stop words (e.g., conjunctions, helping verbs, proper nouns, URLs, and numbers) were excluded from the dictionary. Many of the dictionary terms include wildcards (*); for example, the term *abide** would match *abide*, *abides*, and *abided*.

Because wildcards can yield unexpected term matches, the dictionary was tested on reviewer reports from this study. The dictionary was refined multiple times (based on concordances) to ensure that the computer assisted textual analysis counted as many terms as possible that are relevant to the study while minimizing the number of irrelevant terms counted. The dictionary was designed to err on the side of matching too many terms, which, upon closer analysis, could be excluded later; the relevance of terms in a given data set may or may not be apparent until patterns emerge through the CATA.

I triangulated the manual coding with a computer-aided text analysis of word frequencies, using another custom dictionary based on the peer review literature (Appendix D) and the stop words listed in Appendix E. Stop words are frequently used words such as articles, conjunctions,

linking verbs, and other terms that, within the context of the study, function like noise in the data.

Each pair of reviews was analyzed with Yoshikoder; the CATA program counted the number of words in each review and the number of dictionary terms that appeared in each review. The program also calculated the proportion of matching terms in each review, the percentage change between review 1 and review 2, and the relative risk ratio. The relative risk ratio indicates "the relative probability of seeing each [dictionary term] in each document, controlling for their document lengths"; the computations reflect a 95% confidence interval (Yoshikoder, 2015).

In the next chapter, I report the results of the textual analyses.

Chapter 4

Results

In this chapter, I report the results of my study using a combination of narrative description, descriptive statistics, and inferential statistics. I begin by describing the characteristics of the data set and then present the results that relate to my research questions and hypotheses.

Characteristics of the Data Set

The data set includes copies of 61 initial manuscripts (and the corresponding publication-process artifacts) submitted to two peer reviewed technical communication journals; by *initial,* I mean the manuscripts had been submitted to the respective journals for the first time. The term *publication-process artifact* refers to documents such as reviewers' guidelines and reviewers' reports. Among the publication-process artifacts are 16 revised manuscripts, which were developed from 16 of the initial 61 manuscript submissions, as well as numerous editors' decision letters, authors' response letters, and other artifacts that served as background information; those artifacts will be analyzed in follow-up studies.

The current study focuses primarily on the reviewers' guidelines and the pairs of peer review reports (also called reviews) that accompanied each manuscript in the data set. The breakdown of the peer review report data set follows:

Journal 1

- 60 peer review reports from 30 initial manuscripts
- 2 peer review reports from 1 first-round revised manuscript

Journal 2

- 62 peer review reports from 31 initial manuscripts
- 26 peer review reports from 13 first-round revised manuscripts
- 4 peer review reports from 2 second-round revised manuscripts.

Reviews for revised manuscripts were included in some high-level analyses; however, I prioritized the analysis of initial manuscripts because all but one of the revised manuscripts was submitted to Journal 2. Further analysis of the revised manuscripts and the associated publication-process artifacts will be completed in follow-up studies.

Reviewer Reports

Peer review report forms are typically designed for two distinct audiences (authors and editors), so each of the 154 peer review reports was separated into two audience-segmented files for coding: an author (AU) file and an editor (ED) file (see Figure 3.2). As a result, the number of reviewer report files doubled to 308 audience-segmented report files.

The audience-segmented files simplified some analyses yet complicated others because 15% (n = 46) of the reviewer report files (n = 308) were incomplete—either the AU section of the report form or the ED section had been left blank. Of those 46 files, 39 corresponded with initial manuscript submissions and 7 with revised manuscripts. (Manuscripts with both sections of the form blank were excluded from the sample.) Data from the AU and ED section of each report were combined for most statistical analyses.

When reviewer feedback was provided as both comments on the form and as a file attachment, the form and file text were combined to ensure that the coders had ample context for coding the text. File attachments were analyzed with the author files except for one file attachment that was designated for the editor.

Reviewer report word counts.

The length of the reviewer reports varied by audience segment (i.e., author vs. editor), manuscript status (i.e., initial manuscript vs. revised manuscript), reviewer, manuscript topic, manuscript genre (e.g., research report vs. rhetorical analysis), and journal. The entire sample of reviews comprised 102,820 words, which works out to an average word count of 1,335 words per manuscript reviewed and an average word count of 668 words per reviewer report. Across the sample, the word count for an individual review ranged from zero words to more than 2,500 words. The low end of the range reflects a blank author section of the reviewer report form (i.e., the section of the reviewer report form intended primarily for the author) or a blank editor section of the reviewer report form (i.e., the confidential section of the reviewer report form intended for the editor only); see Figure 3.2. Of those reports with a blank section, about 80% of them had blank editor sections. A journal-by-journal breakdown of the review lengths follows:

Journal 1.

- Total word count of initial manuscript reviewer reports: 44,540 words
- Average word count per manuscript reviewed: 1485 words
- Average word count per reviewer report: 742 words
- Word count range: 0 words to more than 2,500 words

Journal 2.

- Total word count of initial manuscript reviewer reports: 45,981
- Average word count per manuscript reviewed: 1483 words
- Average word count per reviewer report: 742 words
- Word count range: 0 words to more than 2,000 words

Combined Journal 1 and Journal 2 revised manuscripts.

- Total word count of revised manuscript reviewer reports: 12,299 words
- Average word count per revised manuscript reviewed: 769 words
- Average word count per reviewer report: 384 words
- Word count range: 0 words to more than 1,500 words

Structural Analysis

The reviewer forms that were analyzed in this study provided leeway in how the reviewers structured their report. The next five tables summarize the various approaches reviewers used to present their evaluative comments; the data were compiled from questions 8–12 of the coding form. The tallies of the reviewers' use and placement of summarizing judgments are listed in Table 4.1. Nearly half of the Journal 1 reviewers included a summarizing remark in their opening remarks (40% to authors and 48.3% to editors). By comparison, more than two-thirds of Journal 2 reviewers included a summarizing remark in their opening remarks (72.6% to authors and 67.7% to editors).

Table 4.1

Reviewers' Summarizing Judgments

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Summarizing Judgments	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Opening Remarks	24	40	29	48.3	45	72.6	42	67.7
Closing Remarks	3	5	1	1.7	6	9.7	3	4.8
Opening & Closing Remarks	5	8.3	1	1.7	4	6.5	0	0
In File Attachment	2	3.3	0	0	3	4.8	1	1.6
With File Attachment	5	8.3	2	3.3	3	4.8	1	1.6
None	12	20	2	3.3	1	1.6	3	4.8
Other	8	13.3	13	21.7	0	0	8	12.9
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	108.2	60	115	62	103.2	62	112.8

In Table 4.1 the total percentage exceeds the number of review files because multiple options could be selected on the coding form. The "other" category was typically selected when the review consisted of one paragraph and the summarizing judgment was embedded in the opening

or closing remarks of the single paragraph. Single-paragraph comments tended to be found in the editor sections of reviews. Although revised manuscripts are not included in this table, single-paragraph comments were common in second- or third-round reviews (i.e., reviews of revised manuscripts) that generally confirmed the author(s) had successfully completed the reviewers' requested revisions.

As indicated in Table 4.2, less than 15% of the reviewers for either journal provided an outline of the article being reviewed. Journal 2 reviewers (12.9%) were nearly twice as likely to include an article outline in the author section of the comments as Journal 1 reviewers (6.7%). In this sample, none of the reviewers included an article outline in the editor section of the review form.

Table 4.2

Reviewers' Outline of Article Reviewed

	Journal 1				**Journal 2**			
	Author		**Editor**		**Author**		**Editor**	
Inclusion of Outline	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Included Outline of Article	4	6.7	0	0	8	12.9	0	0
No Outline of Article Included	50	83.3	39	65	52	83.9	50	80.6
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

The use of a conclusion paragraph was less consistent. As shown in Table 4.3, Journal 1 and Journal 2 reviewers were nearly evenly split on using them in the author section of the review form; about half of Journal 1 reviewers didn't use them while just over half of Journal 2 reviewers used them. The "other" category muddies the results for the editor sections. Other was usually selected when the review consisted of one paragraph and the review conclusion was

embedded in the single paragraph. Some total percentages exceeded 100 because more than one choice could be selected on the coding form.

Table 4.3

Review Conclusion Paragraph

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Review Conclusion	#	%	#	%	#	%	#	%
Included Conclusion Paragraph	19	31.7	15	25	33	53.2	8	12.9
No Conclusion Paragraph	31	51.7	13	21.7	27	43.5	33	53.2
Other	4	6.7	11	18.3	0	0	9	14.5
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100.1	60	100	62	100	62	100

Reviewers' most common approaches to presenting their evaluative comments were an unnumbered point-by-point format (i.e., paragraphs or bulleted lists) and a combination of point-by-point and location-based comments. See Table 4.4 for details; the total percentage may exceed 100 because some combinations were counted in the individual categories too. The category "other" was used when reviews consisted of a single paragraph, when reviewers asked series of questions, and when reviewers quoted sections of the authors' manuscripts. Although revised manuscripts are not included in this table, the "other" category was used for revisions when reviewers responded to authors' statements.

For both journals, reviewers' most commonly used feedback approach was direct criticism (e.g., *X* is irrelevant to your argument), with praise-criticism pairs the second most commonly used approach (see Table 4.5). For instances of praise and criticism to be counted as a praise-criticism pair, the praise and criticism had to be contained in the same sentence (e.g., *X* is effective; however, *Y* needs work); otherwise, the instances of praise (e.g., direct praise: *X* is

effective) and criticism (e.g., direct criticism: *Y* needs work) were counted separately as either direct or hedged forms of praise (e.g., hedged praise: *X seems* to support your argument) or criticism (e.g., hedged criticism: *Y* is confusing *but* that could be my reading of the text).

Table 4.4

Review Comment Presentation Format

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Review Comment Format	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Numbered Point-by-Point	3	5	2	3.3	5	8.1	2	3.2
Unnumbered Point-by-Point	16	26.7	20	33.3	21	33.9	33	53.2
Combination Numbered/Unnumbered	0	0	0	0	2	3.2	1	1.6
Page-by-Page (location based)	7	11.7	0	0	3	4.8	2	3.2
Section-by-Section (location based)	4	6.7	0	0	2	3.2	0	0
Combination of Location Based	5	8.3	0	0	4	6.5	1	1.6
Combination of Point/Location Based	12	20	5	8.3	23	37.1	2	3.2
Other	7	11.7	14	23.3	0	0	15	24.2
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100.1	60	103.2	62	100	62	109.6

Data for the revised manuscripts are shown in Table 4.6 to show how the nature of the feedback changes with the status of the manuscript (i.e., initial submission versus revision). Although the amount of data from Journal 1 is inadequate for comparison, the data from Journal 2 indicate that praise is more common when evaluating revised manuscripts. Criticism, however, remains a frequently used feedback approach in comments intended for authors and editors.

Table 4.5

Reviewers' Feedback Approaches: Original Manuscripts

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Feedback Approach	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Praise/Criticism Pairs	38	63.3	17	28.3	30	48.4	22	35.5
Hedged Praise	7	11.7	4	6.7	2	3.2	1	1.6
Hedged Criticism	12	20	2	3.3	0	0	2	3.2
Praise/Journal Criteria	0	0	0	0	0	0	0	0
Criticism/Journal Criteria	1	1.7	2	3.3	3	4.8	2	3.2
Direct Praise	23	38.3	9	15	48	77.4	17	27.4
Direct Criticism	50	83.3	28	46.7	60	96.8	42	67.7
Other	1	1.7	3	5	1	1.6	2	3.2
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	230	60	143.3	62	235.4	62	161.2

Table 4.6

Reviewers' Feedback Approaches: Revised Manuscripts

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Feedback Approach	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Praise/Criticism Pairs	0	0	2	100	7	23.3	7	23.3
Hedged Praise	0	0	0	0	3	10	0	0
Hedged Criticism	0	0	0	0	1	3.3	1	3.3
Praise/Journal Criteria	0	0	0	0	0	0	0	0
Criticism/Journal Criteria	0	0	0	0	0	0	2	6.7
Direct Praise	0	0	0	0	29	96.7	19	63.3
Direct Criticism	0	0	0	0	19	63.3	13	43.3
Other	1	50	0	0	0	0	2	6.7
Blank Form	1	50	0	0	0	0	6	20
Total Review Files	2	100	2	100	30	196.6	30	166.6

These data have provided insight on how reviewers present their evaluative comments to authors and editors. The next data look beyond the presentation of the comments to the alignment of the reviewers' comments with the manuscript evaluation instrument—more commonly known as the reviewer guidelines or reviewer assessment rubric.

Manuscript Evaluation Instrument

The next three tables report the Likert rating data from question 13 of the coding form (see Appendix C). Data for revised manuscripts are not shown because those reviews tended to be shorter and less aligned with the reviewer guidelines and rubrics; instead, the reviews of revised manuscripts often mirrored the format of the authors' response letters or the original review.

At best, the review data for the original manuscripts seem to show moderate alignment with the reviewers' guidelines or assessment rubrics for any of the categories evaluated. The Likert ratings in Table 4.7 reveal a slightly stronger alignment of the review content/topics for Journal 2, while the ratings in Table 4.8 point to a somewhat stronger alignment with the structure/format/order for Journal 1 in the author section of the form; however, the number of blank editor sections of the form muddies those results. Based on the data in Table 4.9, Journal 2 appears to align more strongly with the language/wording of its reviewer guidelines/rubric than Journal 1 does with its respective guidelines/rubric.

Table 4.7

Alignment of Review with Content/Topics of Manuscript Evaluation Instruments

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Alignment Level	#	%	#	%	#	%	#	%
Strongly Disagree	1	1.7	0	0	0	0	0	0
Disagree	0	0	2	3.3	0	0	1	1.6
Somewhat Disagree	5	8.3	4	6.7	2	3.2	0	0
Somewhat Agree	39	65	30	50	28	45.2	37	6
Agree	9	15	3	5	24	38.7	10	16.1
Strongly Agree	0	0	0	0	6	9.7	2	3.2
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.8

Alignment of Review with Structure/Format/Order of Manuscript Evaluation Instruments

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Alignment Level	#	%	#	%	#	%	#	%
Strongly Disagree	0	0	0	0	0	0	0	0
Disagree	1	1.7	4	6.7	0	0	2	3.2
Somewhat Disagree	11	18.3	11	18.3	7	11.3	10	16.1
Somewhat Agree	34	56.7	22	36.7	34	54.8	31	50
Agree	8	13.3	2	3.3	13	21	5	8.1
Strongly Agree	0	0	0	0	6	9.7	2	3.2
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.9

Alignment of Review with Language/Wording of Manuscript Evaluation Instruments

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Alignment Level	#	%	#	%	#	%	#	%
Strongly Disagree	0	0	0	0	0	0	0	0
Disagree	0	0	3	5	0	0	1	1.6
Somewhat Disagree	5	8.3	2	3.3	2	3.2	0	0
Somewhat Agree	41	68.3	32	53.3	28	45.16	36	48.4
Agree	8	13.3	2	3.3	24	38.7	11	17.7
Strongly Agree	0	0	0	0	6	9.7	2	3.2
Blank Form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

To triangulate the subjective Likert ratings reported in Tables 4.7–4.9, I performed a computer-aided text analysis (CATA). The results of the CATA follow.

Computer-Aided Text Analysis (CATA) Results

The computer-aided text analysis (CATA) compared the text of the reviewer reports to terms in custom dictionaries and calculated the number of matches, proportion of matches, direction of the matches (i.e., whether Reviewer 1 or Reviewer 2 matched more dictionary terms), and the risk ratios (the expected number of matches based on the review word counts). The custom dictionary terms were compiled from technical and professional communication journals' reviewers' guidelines and rubrics, publishers' guides (e.g., Sage Journals, n.d.; Taylor & Francis, 2019), and from ethical guidelines published by the Council on Publication Ethics (COPE Council, 2013, 2017).

The CATA results (Table 4.10) show minimal alignment with the journals' respective guidelines/rubrics. At most, in the editor section of the form, Journal 1 matched 16.7% of the

Table 4.10

CATA—Comparison of Reviewer Reports to Reviewer Rubrics and Guidelines

	Journal 1		Journal 2	
Overall	**Comment Audience**		**Comment Audience**	
Reviewer Rubric/Guidelines Source	Author	Editor	Author	Editor
Journal 1	0%–7.8%	0%–16.7%	n/a	n/a
Journal 2	n/a	n/a	0%–13.5%	0%–12.8%
TC Journals, Sage, T&F, & COPE	0%–40.0%	11.1%–46.9%	18.1%–36.7%	3.8%–41.3%
Initial Submissions	Journal 1		Journal 2	
Reviewer Rubric/Guidelines Source	Author	Editor	Author	Editor
Journal 1	0%–7.8%	0%–16.7%	n/a	n/a
Journal 2	n/a	n/a	0.06%–13.5%	0%–11.7%
TC Journals, Sage, T&F, & COPE	19.1%–40.0%	11.1%–46.9%	18.1%–34.3%	3.8%–41.3%
First Revisions	Journal 1		Journal 2	
Reviewer Rubric/Guidelines Source	Author	Editor	Author	Editor
Journal 1 (1 manuscript in sample)	0%	2.4%–4.8%	n/a	n/a
Journal 2 (13 manuscripts in sample)	n/a	n/a	0%–12.8%	3.8%–12.8%
TC Journals, Sage, T&F, & COPE	0%	22.2%–28.6%	21.3%–36.7%	20.7%–38.7%
Second Revisions	Journal 1		Journal 2	
Reviewer Rubric/Guidelines Source	Author	Editor	Author	Editor
Journal 1 (no manuscripts in sample)	n/a	n/a	n/a	n/a
Journal 2 (2 manuscripts in sample)	n/a	n/a	0%–13.0%	4.6%–7.6%
TC Journals, Sage, T&F, & COPE	n/a	n/a	18.7%–30.4%	19.7–32.1%
R1/R2 Combined All Versions	Journal 1		Journal 2	
Reviewer Rubric/Guidelines Source	Author	Editor	Author	Editor
Journal 1	9.4%	4.1%	n/a	n/a
Journal 2	n/a	n/a	4.8%	6.5%
TC Journals, Sage, T&F, & COPE	27.10%	29.3%	26.4%	28.1%
R1/R2 & AU/ED Combined	Journal 1		Journal 2	
Reviewer Rubric/Guidelines Source	Author & Editor		Author & Editor	
Journal 1	9.9%		n/a	
Journal 2	n/a		5.1%	
TC Journals, Sage, T&F, & COPE	27.7%		26.8%	

terms, and, in the author section of the form, Journal 2 matched 13.5% of the dictionary terms. A comparison of the reviewer reports with a broader dictionary comprising terms from additional resources (e.g., additional technical communication journals, Sage, Taylor & Francis, COPE) yielded better results. The content/topics and language/wording of the reviewer reports aligned with 46.9% of the terms in the editor section of Journal 1's reviews and 41.3% of the editor section of Journal 2's reviews.

Using the respective journal dictionaries, the Journal 1 comments to the editor matched more dictionary terms than did the Journal 1 comments to the authors, while the Journal 2 comments to the author matched the same or slightly more dictionary terms than did the Journal 2 comments to the editors. Though not a fair comparison given unequal sample sizes, for Journal 1, the first revision comments to the editor matched more terms, and the opposite was true for the Journal 2 first revisions. For the Journal 2 second revisions, the comments to the author matched nearly twice as many dictionary terms compared to the comments to the editor; however, this data set consisted of two manuscripts, which isn't a large enough data set for meaningful analysis. For both journals overall, the comments to the editor matched more dictionary terms (TC Journals, Sage, Taylor & Francis, COPE) than did the comments to the authors.

The CATA showed significant results for several manuscripts. When comparing the Journal 1 dictionary terms to the Journal 1 author-segmented (AU) reviews, six manuscripts had significant risk ratios; of those, four Reviewer 2s matched significantly more terms than the Reviewer 1s and two Reviewer 1s matched significantly more terms than the Reviewer 2s. (Risk ratios are descriptive statistics and as such do not indicate statistical significance; in other words, the significant results cannot used to predict outcomes or generalize results.)

When comparing the broader dictionary terms to the Journal 1 AU reviews, eight manuscripts had significant risk ratios; in those, the Reviewer 1s and Reviewer 2s were evenly split. None of the Journal 1 editor-segmented (ED) reviews showed significant results.

In contrast, the CATA showed significant results for nine initial AU manuscript reviews and two first-revision AU reviews when comparing the Journal 2 dictionary terms to the Journal 2 AU reviews. For the initial manuscript reviews, 6 Reviewer 2s matched significantly more terms than the Reviewer 1s and three Reviewer 1s matched significantly more terms than the corresponding Reviewer 2s.

When comparing the broader dictionary terms to the Journal 2 AU reviews, 13 manuscripts had significant risk ratios; of those, two were revised manuscripts where the Reviewer 1s and Reviewer 2s were evenly split over the most number of matches. For the 11 initial manuscripts, 7 Reviewer 1s matched more terms than the Reviewer 2s and 4 Reviewer 2s matched more terms than the Reviewer 1s.

Journal 2 also showed significant results for two manuscripts when comparing the ED reviews with the Journal 2 dictionary; in both cases the Reviewer 2s matched significantly more terms than the Reviewer 1s. With the broader dictionary, one of the same manuscripts again showed significant results but to a lesser degree, and another manuscript showed significant results with Reviewer 1 matching significantly more terms than Reviewer 2.

When the data were analyzed with the Reviewer 1 (R1) and Reviewer 2 (R2) reports combined (i.e., all the R1 and R2 author files were combined and all the R1 and R2 editor files were combined), for each journal across all manuscript submission versions, the results were significant for both journals for both dictionaries. Likewise, when the data were analyzed with all the reviews combined by journal the results were significant when compared against the broader

dictionary. The audience-segmented analyses showed negative percentage changes, whereas the combined journal-by-journal analysis showed a positive percentage change. Based on the results of these various analyses, hypothesis 1 was rejected. Hypothesis 1 posited that no significant difference exists between the content of the reviewers' evaluative comments and the content of the journal's reviewers' guidelines or scoring rubrics.

The data reported in this section addressed the alignment of the reviewer reports with the journals' manuscript evaluation instruments. The data in the next section relate to the quality of the reviewer reports.

Review Quality

After analyzing how well the reviews aligned with the journals' reviewers' guidelines and rubrics, I used a modified version of van Rooyen, Black, and Godlee's (1999) review quality instrument (RQI) to determine which elements of the manuscripts had been evaluated by the reviewers. The RQI was integrated into questions 14–18 of the coding form; those results are presented in Tables 4.11–4.22.

In this data set, the majority of the reviews for both journals devoted little attention to the importance of the research question (Table 4.11). In particular, 65% to 85% of the reviews for Journal 1 were rated as discussing the research question as "none at all" or "a little." Similarly, 80% to 97% of the reviews for Journal 2 were rated in the same two categories.

Reviewers spent even less time discussing the originality of the manuscripts (Table 4.12). Between 53.3% and 71.7% of the Journal 1 reviews did not discuss the topic at all, whereas between 62.9% and 67.7% of the Journal 2 reviews did not discuss the topic.

Reviewers for both journals also spent little time discussing the strengths of the study designs (Table 4.13); however, the reviewers discussed the weaknesses of the study designs about three

times as much (Table 4.14). The reviewers' criticisms appealed more in the author (AU) section of the form. About 40% of the Journal 1 AU comments were rated "a little" or "a moderate amount" versus about 45% of the Journal 2 AU comments.

Table 4.11

Importance of Research Question

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	27	45	16	26.7	22	35.5	37	59.7
A little	24	40	23	38.3	37	59.7	13	21
A moderate amount	3	5	0	0	1	1.6	0	0
A lot	0	0	0	0	0	0	0	0
A great deal	0	0	0	0	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.12

Originality of Paper

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	43	71.7	32	53.3	42	67.7	39	62.9
A little	9	15	7	11.7	17	27.4	11	17.7
A moderate amount	1	1.7	0	0	1	1.6	0	0
A lot	1	1.7	0	0	0	0	0	0
A great deal	0	0	0	0	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.13

Strengths of Study Design

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	48	80	37	61.7	52	83.9	48	77.4
A little	6	10	2	3.3	7	11.3	2	3.2
A moderate amount	0	0	0	0	1	1.6	0	0
A lot	0	0	0	0	0	0	0	0
A great deal	0	0	0	0	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

The results for reviewers' discussion of the strengths of the methods (Table 4.15) were similar to those relating to the strengths of the study design. Nearly all of the reports were rated as discussing the topic "none at all" or "a little" with about two-thirds to three-quarter of the responses falling in the "none" category for both journals.

Table 4.14

Weaknesses of Study Design

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	26	43.3	27	45	31	50	40	64.5
A little	17	28.3	6	10	19	30.6	8	12.9
A moderate amount	7	11.7	5	8.3	9	14.5	2	3.2
A lot	4	6.7	1	1.7	1	1.6	0	0
A great deal	0	0	0	0	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.15

Strengths of Methods

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	43	71.7	37	61.7	51	82.3	47	75.8
A little	11	18.3	2	3.3	8	12.9	3	4.8
A moderate amount	0	0	0	0	1	1.6	0	0
A lot	0	0	0	0	0	0	0	0
A great deal	0	0	0	0	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Again, the reviewers for both journals discussed weaknesses more than strengths (Table 4.16). Journal 1 reviewers discussed weaknesses of methods more than Journal 2 reviewers; at least twelve of the Journal 1 reviewers discussed them "a moderate amount" to "a great deal." About two-thirds of the Journal 2 reviewers did not discuss the topic at all compared to about half of the Journal 1 reviewers.

Table 4.16

Weaknesses of Methods

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	31	51.7	26	43.3	39	62.9	42	67.7
A little	11	18.3	9	15	13	21	7	11.3
A moderate amount	10	16.7	3	5	7	11.3	1	1.6
A lot	1	1.7	0	0	1	1.6	0	0
A great deal	1	1.7	1	1.7	0	0	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

About 30% of the Journal 1 and Journal 2 reviewers were rated as discussing the author's interpretations of results "a moderate amount" or "a lot." About 40% of the reviewers discussed the topic "a little" (see Table 4.17).

Table 4.17

Author's Interpretations of Results

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	12	20	21	35	9	14.5	23	37.1
A little	24	40	14	23.3	21	33.9	24	38.7
A moderate amount	15	25	3	5	24	38.7	3	4.8
A lot	3	5	1	1.7	5	8.1	0	0
A great deal	0	0	0	0	1	1.6	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Combined, about 75% of the Journal 1 AU reports discussed higher order writing concerns "a little," "a moderate amount," or "a lot" (Table 4.18). By comparison, nearly 90% of the Journal 2 AU reports clustered in those same Likert scale scores.

The results for the extent reviewers discussed lower order writing concerns trended toward the lower ends of the Likert scale (Table 4.19). Approximately 40% of Journal 1 reviewers did not discuss the topic at all, while more than 50% of Journal 2 reviewers did not discuss lower order writing concerns.

For both journals, approximately 70% of the reviewer's comments were rated as "moderately" or "very" constructive (Table 4.20). The majority of reviews for both journals were rated as having substantiated "most comments" (Table 4.21). Within that rating category,

reviewers had substantiated comments three times more often in the AU section of the form than in the ED section of the form.

Table 4.18

Higher Order Writing Concerns

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	9	15	16	26.7	6	9.7	19	30.6
A little	25	41.7	17	28.3	15	24.2	25	40.3
A moderate amount	18	30	5	8.3	28	45.2	4	6.5
A lot	2	3.3	1	1.7	10	16.1	2	3.2
A great deal	0	0	0	0	1	1.6	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.19

Lower Order Writing Concerns

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Extent Discussed	#	%	#	%	#	%	#	%
None at all	23	38.3	26	43.3	33	53.2	35	56.5
A little	21	35	9	15	19	30.6	10	16.1
A moderate amount	6	10	4	6.7	4	6.5	5	8.1
A lot	2	3.3	0	0	3	4.8	0	0
A great deal	2	3.3	0	0	1	1.6	0	0
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.20

Constructiveness of Reviewer's Comments

	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
Constructiveness	**#**	**%**	**#**	**%**	**#**	**%**	**#**	**%**
Not at all	1	1.7	1	1.7	0	0	3	4.8
Slightly	6	10	12	20	3	4.8	16	25.8
Moderately	22	36.7	20	33.3	23	37.1	22	35.5
Very	20	33.3	4	6.7	29	46.8	8	12.9
Extremely	5	8.3	2	3.3	5	8.1	1	1.6
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

For both journals, the overall review quality is rated as "average" to "excellent" for about 80% of the reviews (Table 4.22). These rankings are similar to—if not somewhat lower than—the ones that accompany about half of the reviews in this sample. The journals' reviewer report forms allow the editor to rank the reviewers' timeliness and review quality on a three-point scale with one being lowest and three highest. Although the scale used on the coding form is not identical to the one used by the journals in this sample, for comparison purposes, "poor" and "fair" was mapped to an editor rating of one; "average" was mapped to two, and "good" and "excellent" were mapped to three. Of the reports in this sample rated by editors for quality, one report was rated one, three reports were rated two, and the rest were rated three.

Table 4.21

Substantiated Reviewer Comments

Substantiated Comments	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
	#	%	#	%	#	%	#	%
No comments	3	5	7	11.7	0	0	6	9.7
Few comments	3	5	4	6.7	1	1.6	8	12.9
Some comments	9	15	14	23.3	6	9.7	16	25.8
Most comments	36	60	12	20	49	79	16	25.8
All comments	3	5	2	3.3	4	6.5	4	6.5
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

Table 4.22

Overall Review Quality

Review Quality	Journal 1				Journal 2			
	Author		Editor		Author		Editor	
	#	%	#	%	#	%	#	%
Poor	2	3.3	2	3.3	0	0	2	3.2
Fair	8	13.3	10	16.7	3	4.8	16	25.8
Average	14	23.3	16	26.7	7	11.3	13	21
Good	25	41.7	9	15	45	72.6	18	29
Excellent	5	8.3	2	3.3	5	8.1	1	1.6
Blank form	6	10	21	35	2	3.2	12	19.4
Total Review Files	60	100	60	100	62	100	62	100

The data reported in this section addressed the quality of the reviewer reports. The next section covers in more detail what the reviewers said in their evaluative comments.

Quantitative Content Analysis of Reviewers' Evaluative Comments

The reviewers' reports were coded using a modified version of Bornmann, Nast, and Daniel's (2008) coding scheme. The aggregate coding counts for the reviewers' evaluation of initial manuscript submissions are displayed in Table 4.23, which shows the data for Journal 1, and Table 4.24, which shows the data for Journal 2; this information comes from coding form questions 4 and 5.

Descriptive statistics confirmed that the data are not distributed normally, thus various nonparametric tests were used for inferential statistical analyses.

Elements Discussed by Reviewers

For each journal, the aggregate Reviewer 1 and Reviewer 2 coding counts appear similar. The percentages of each element discussed by the reviewers are shown in Figure 4.1, which shows the results for Journal 1, and Figure 4.2, which shows the results for Journal 2. The gap between the Journal 1 reviewers' discussion of Methods is about 5 percentage points larger than gap between the Journal 2 reviewers' discussion of Contribution.

Table 4.23

Comparison of Journal 1 Reviewers' Evaluative Comments on Initial Manuscript Submissions

Journal 1	**Reviewer 1 to Author**					**Reviewer 2 to Author**					**Reviewer 1 vs. Reviewer 2**	
	Element Discussed by Reviewer		**Elements Discussed** (positive, negative, or neutral)	**Problems Reviewer Discussed** (negative elements)		**Element Discussed by Reviewer**		**Elements Discussed** (positive, negative, or neutral)	**Problems Reviewer Discussed** (negative elements)		**Problems Both Reviewers Discussed**	**Contradictory Comments**
Comment Category	Yes	No	Times Coded	Times Discussed	# Unique Problems	Yes	No	Times Coded	Times Discussed	# Unique Problems	# Times Agreed	# Times Disagreed
Contribution	20	5	56	44	31	25	4	112	88	56	9	1
Writing Higher	21	4	122	112	84	23	6	114	105	86	16	2
Writing Lower	12	13	141	134	125	18	11	82	79	71	2	0
Design	14	11	57	50	31	18	11	58	51	37	2	1
Methods	10	15	67	51	42	18	11	75	58	48	7	1
Results	17	8	90	82	68	24	5	97	80	63	6	0
Literature	17	8	52	41	31	22	7	78	63	49	5	2
Theory	7	18	13	12	10	11	18	26	19	16	1	0
Ethics	2	23	2	2	2	1	28	1	1	1	0	0
Other	1	24	1	0	0	0	29	0	0	0	0	0
Total	121	129	601	528	424	160	130	643	544	427	48	7
Blank Form			5 of 30					1 of 30			6 of 60	
	Reviewer 1 to Editor					**Reviewer 2 to Editor**					**Reviewer 1 vs. Reviewer 2**	
Contribution	18	4	36	29	22	12	5	29	24	18	4	3
Writing Higher	13	9	52	50	33	11	6	22	18	17	1	2
Writing Lower	8	14	20	15	13	5	12	11	10	9	0	1
Design	8	14	27	22	14	7	10	20	18	16	1	2
Methods	7	15	37	30	17	7	10	18	14	10	0	0
Results	12	10	38	34	25	7	10	16	14	9	1	0
Literature	12	10	23	20	19	8	9	13	12	10	1	0
Theory	2	20	5	5	2	1	16	1	1	1	0	0
Ethics	1	21	2	2	1	1	16	1	1	1	0	0
Other	4	18	4	1	1	6	11	6	5	5	1	0
Total	85	135	244	208	147	65	105	137	117	96	9	8
Blank Form			8 of 30					13 of 30			21 of 60	

Table 4.24

Comparison of Journal 2 Reviewers' Evaluative Comments on Initial Manuscript Submissions

Journal 2	**Reviewer 1 to Author**					**Reviewer 2 to Author**					**Reviewer 1 vs. Reviewer 2**	
	Element Discussed by Reviewer		**Elements Discussed** (positive, negative, or neutral)	**Problems Reviewer Discussed** (negative elements)		**Element Discussed by Reviewer**		**Elements Discussed** (positive, negative, or neutral)	**Problems Reviewer Discussed** (negative elements)		**Problems Both Reviewers Discussed**	**Contradictory Comments**
Comment Category	Yes	No	Times Coded	Times Discussed	# Unique Problems	Yes	No	Times Coded	Times Discussed	# Unique Problems	# Times Agreed	# Times Disagreed
Contribution	24	5	54	39	28	22	9	44	26	18	6	0
Writing Higher	26	3	139	126	107	28	3	146	131	103	10	0
Writing Lower	12	17	39	35	33	15	16	47	45	43	2	0
Design	19	10	54	43	30	18	13	35	29	23	5	0
Methods	16	13	40	34	29	14	17	48	33	29	3	0
Results	26	3	104	88	79	25	6	83	72	68	4	0
Literature	24	5	82	69	56	28	3	61	45	42	8	1
Theory	14	15	31	27	15	13	18	31	23	20	4	0
Ethics	0	29	0	0	0	0	31	0	0	0	0	0
Other	3	26	3	2	2	4	27	5	3	3	0	0
Total	164	126	546	463	379	167	143	500	407	349	42	1
Blank Form			2 of 31					0 of 31			2 of 62	
	Reviewer 1 to Editor					**Reviewer 2 to Editor**					**Reviewer 1 vs. Reviewer 2**	
Contribution	19	7	32	21	15	14	10	18	12	11	1	0
Writing Higher	19	7	45	43	30	13	11	37	36	28	2	0
Writing Lower	9	17	30	26	24	6	18	13	13	12	1	1
Design	6	20	19	18	11	6	18	8	8	6	1	0
Methods	8	18	19	15	7	6	18	10	7	6	1	0
Results	12	14	22	20	20	15	9	23	22	21	1	0
Literature	8	18	21	13	10	5	19	9	7	5	2	0
Theory	4	22	6	4	3	3	21	5	5	5	0	0
Ethics	2	24	7	7	4	0	24	0	0	0	0	0
Other	4	22	5	4	3	5	19	5	4	4	0	0
Total	91	169	206	171	127	73	167	128	114	98	9	1
Blank Form			5 of 31					7 of 31			12 of 62	

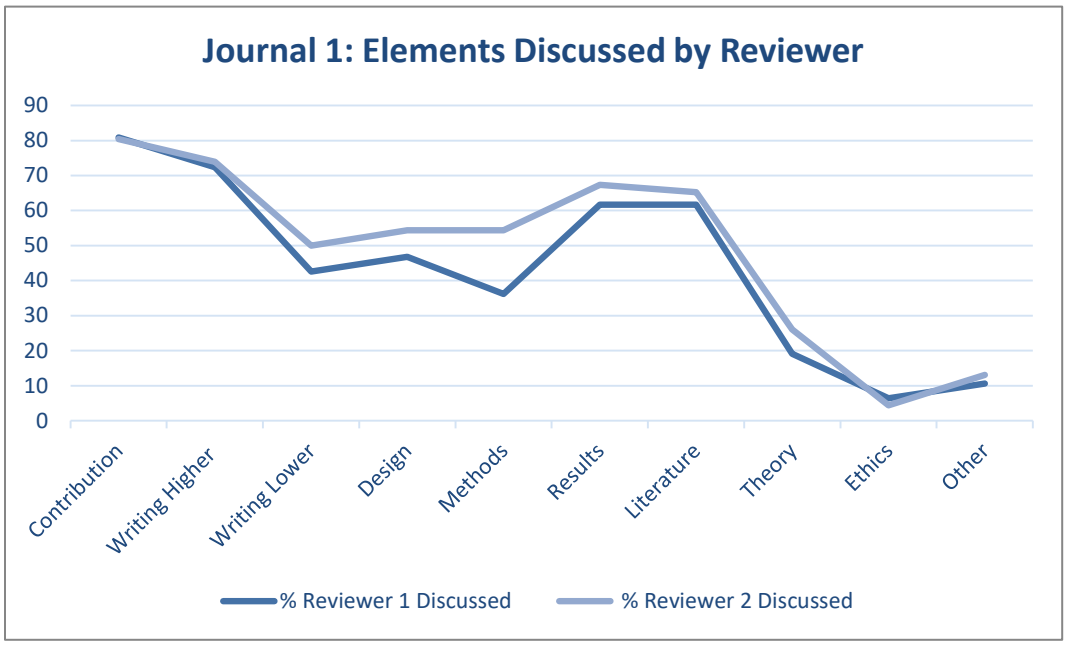


Figure 4.1.

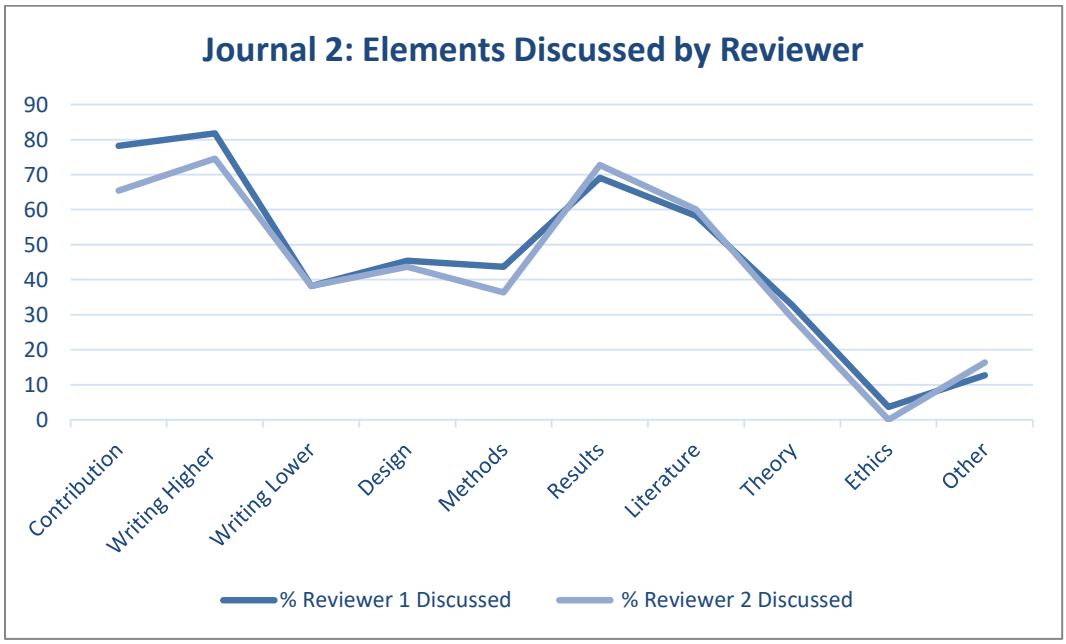


Figure 4.2.

An examination of the reviews by audience segment shows more variation between data points (Figures 4.3 and 4.4) than in the comparison between Reviewer 1 and Reviewer 2 (Figures 4.1 and 4.2). For both journals, the graph lines differ by at least 6 percentage points for each category except Ethics and Other (Figures 4.3 and 4.4). Notably, the reviewers discussed ethics slightly more frequently when their comments were intended for editors than authors. Similarly, the data shows that topics categorized as Other tended to be discussed more when the comments were intended for editors.

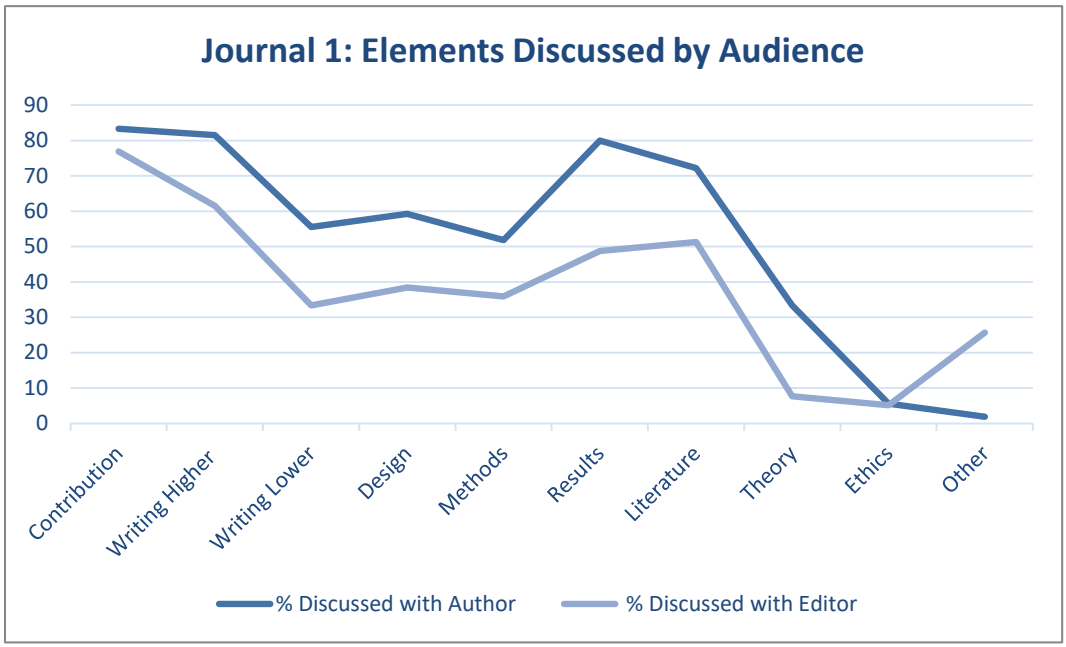


Figure 4.3.

When the reviewer data for Journal 1 and Journal 2 are combined (Figure 4.5), the by-audience discussion patterns change little. The graph peaks flatten in places and the spread between lines widens by a few percentage points in places (e.g., Contribution) yet narrows in some places (e.g., Other).

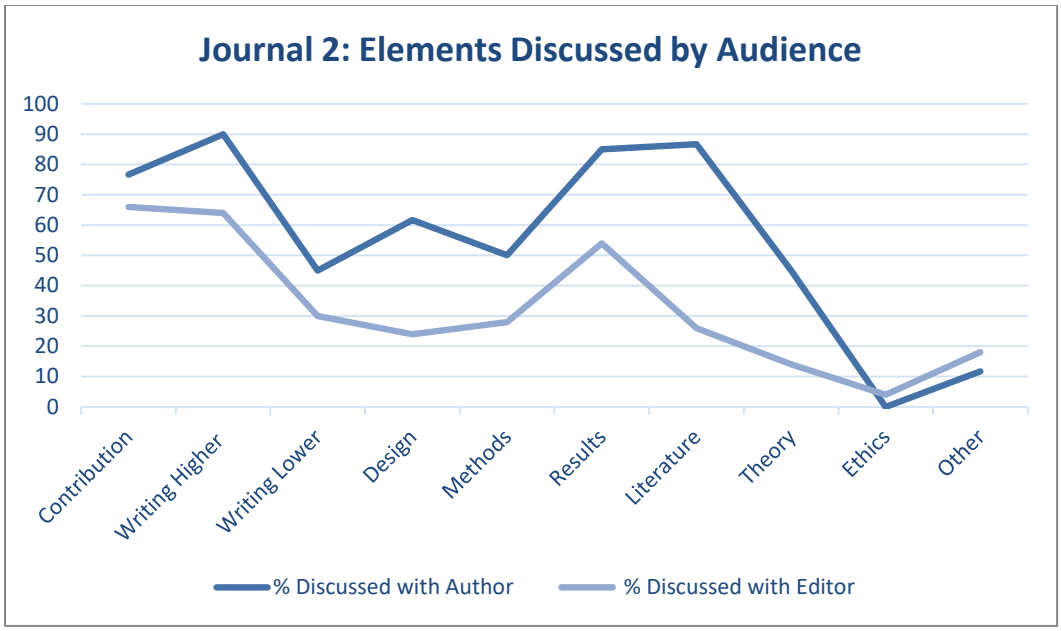


Figure 4.4.

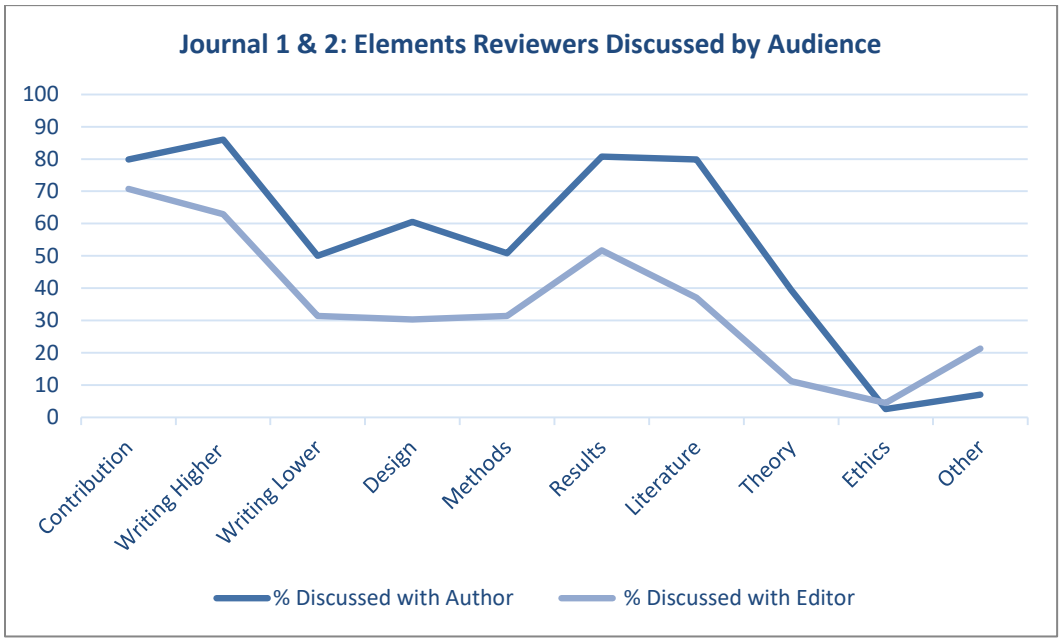


Figure 4.5.

Elements Discussed by Topic and Severity Level

The next analyses compare aspects of the data that relate directly to the study's research questions and hypotheses; these are topic-by-topic comparisons within groups (i.e., Journal 1 and Journal 2).

High-level writing comments versus low-level writing comments.

The number of comments (positive, negative, or neutral) each reviewer made in the categories of High-Level Writing/Presentation and Low-Level Writing/Presentation were compared. For Journal 1, the statistical analysis of the number of comments related to high-level writing versus low-level writing produced statistically significant results for the sign test (p = .0307) and the Wilcoxon signed rank test was nearly significant (p = .0501), which suggests the two reviewers evaluated those areas of the manuscripts to different extents (see Figure 4.6).

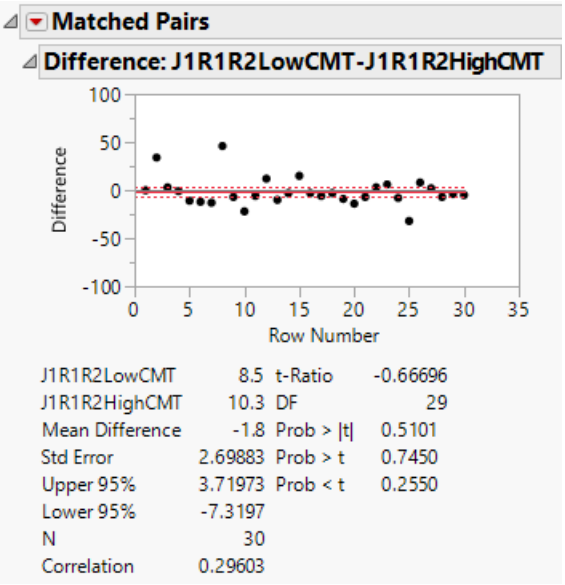


Figure 4.6: Journal 1: Low vs high comments.

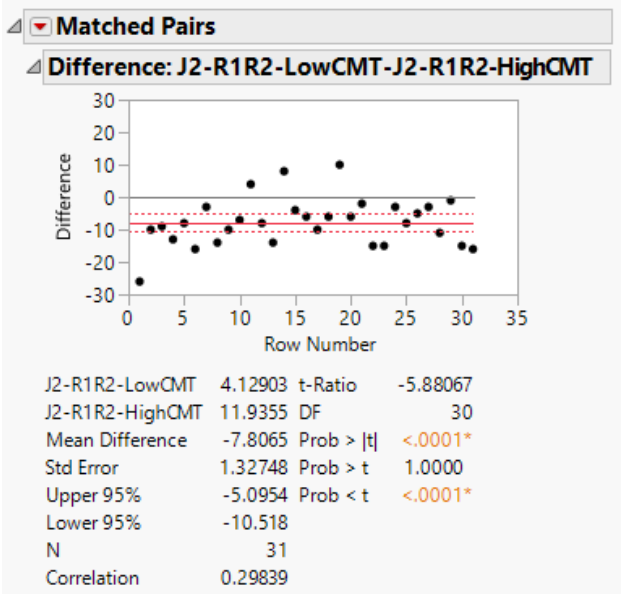


Figure 4.7: Journal 2: Low vs high comments.

The Journal 2 data distribution (Figure. 4.7) differs markedly from Journal 1 (Figure 4.6) and the Wilcoxon signed rank and sign test results are highly significant (p < .0001) for the comparison

of the number of comments each reviewer made in the categories of high-level writing concerns vs. low-level writing concerns.

High-level problems versus low-level problems.

The next analyses compare the combined number of reviewers comments related to all high-level problems (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) versus comments related to low-level problems (i.e., low-level writing and presentation problems). The high-level problem data includes the data from all coding categories except Low-Level Writing/Presentation and Other; comments coded as Other were excluded because they could have been high- or low-level problems. The low-level problem data includes only the data from comments coded as Low-Level Writing/Presentation.

For both journals, the Wilcoxon signed rank and sign test results were highly significant ($p < .0001$); therefore, hypothesis 2, which posited that no significant difference exists between the number of reviewer comments associated with higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) and the number of reviewer comments associated with lower level concerns (e.g., grammar, mechanics, style, citations), was rejected. See Figures 4.8 and 4.9.

Category-by-category unique problems.

An examination of the distribution and number of unique problems identified shows (Figure 4.10) that the top three issues that the Journal 1 reviewers discussed were high-level writing concerns (20.11%), low-level writing concerns (19.93%), and the author's discussion of the study results (15.08%). For Journal 2, the top three issues were high-level writing concerns (28.12%), the author's discussion of the study results (19.72%), and literature and documentation (11.86%), followed closely by low-level writing concerns (11.75%).

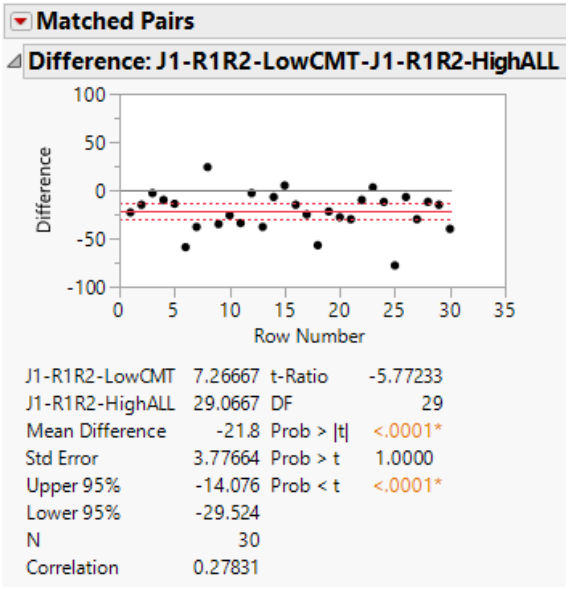


Figure 4.8: Journal 1: Low vs all comments.

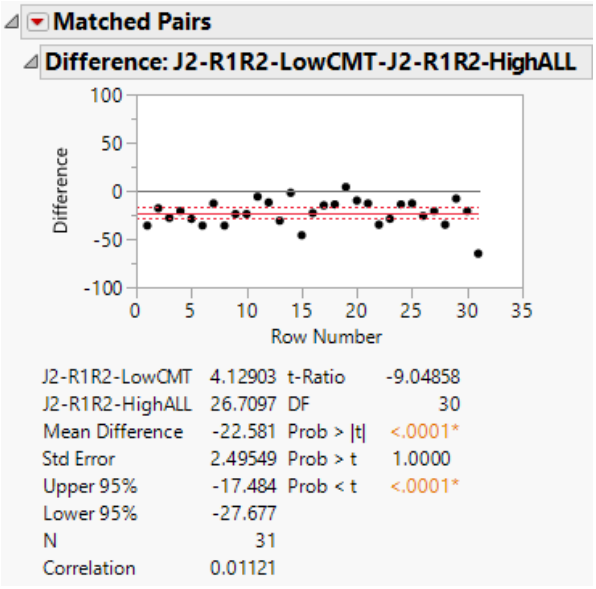


Figure 4.9: Journal 2: Low vs all comments.

Category-by-category analyses of the unique problems each reviewer identified yielded several statistically significant results for each journal. The data were first analyzed by yes–no binaries (i.e., the reviewers did/did not identify problems in the category).

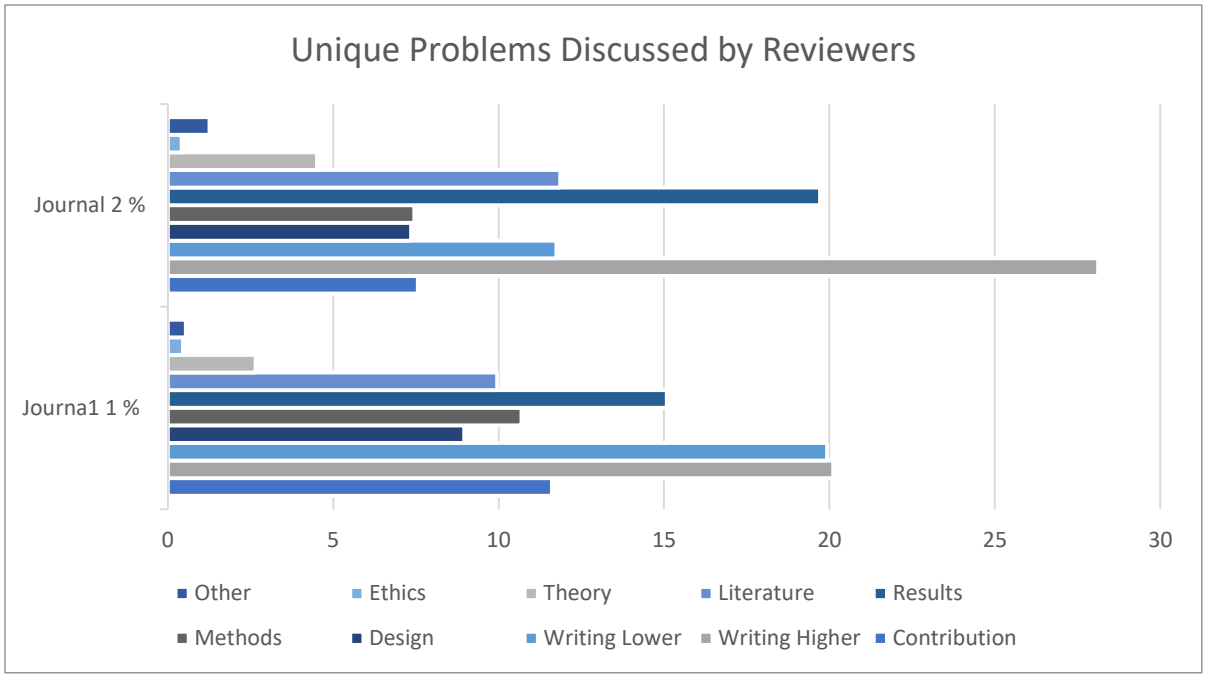


Figure 4.10.

Journal 1.

The kappa scores for interreviewer agreement level ranged from −.34043 (Design) to .385965 (tie: Results and Theory). The Pearson correlation coefficients were significant in these categories:

- Design (r = .0371)
- Results (r = .0284)
- Theory (r = .0284).

Journal 2.

The kappa scores for interreviewer agreement level ranged from −.10714 (Results) to .640371 (Theory). The Pearson correlation coefficients were significant in these categories:

- Methods (r = .0484)
- Theory (r = .0004).

Given a significant Pearson r for both journals for the category Theory and the fact that each journal had at least two categories with significant Pearson correlation coefficients, hypothesis 4 was rejected. Hypothesis 4 posited that no significant difference exists between the types of problems each reviewer identifies.

Next, the data were analyzed by problem counts (i.e., the number of problems Reviewer 1 identified versus the number of problems that Reviewer 2 identified. The results follow.

Journal 1.

The kappa scores for interreviewer agreement level ranged from −.14286 (Design) to .212598 (Theory). The Pearson correlation coefficients were significant in these categories:

- Writing Low-Level Concerns (r = .0029)
- Results (r = .0422).

Journal 2.

The kappa scores for interreviewer agreement level ranged from −.12198 (Literature/Documentation) to .276265 (Theory). The Pearson correlation coefficients were significant in these categories:

- Design (r = .0014)
- Method (r = .0053)
- Theory (r = .0001).

Elements Discussed by Reviewers

The next analyses compare the types of problems that reviewers identified and the degree of agreement. These are primarily topic-by-topic comparisons between groups (i.e., Reviewer 1 and Reviewer 2).

Reviewer 1 versus Reviewer 2 identification of unique problems.

A Wilcoxon signed rank test did not show statistically significant results for either journal when comparing the unique problems identified by Reviewer 1 and Reviewer 2 for the respective journals; therefore, hypothesis 5 was not rejected. Hypothesis 5 posited that no significant difference exists between the numbers of manuscript problems each reviewer identifies. (See Figures 4.11 and 4.12.)

Comparative Content Analysis

The results of the content analysis for each reviewer were compared. An analysis of specific problems that both reviewers identified shows very few points of agreement. Journal 1 reviewers agreed on 5.21% of the identified problems and Journal 2 reviewers agreed on 5.35% of the identified problems.

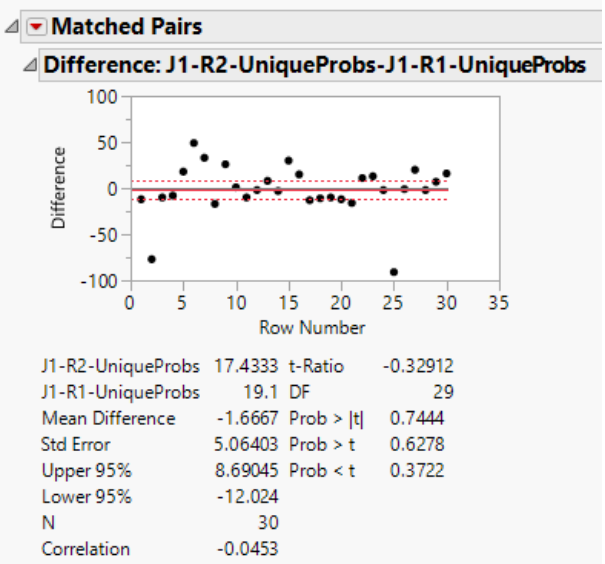


Figure 4.11: Journal 1—Reviewer 1 vs Reviewer 2 unique problems.

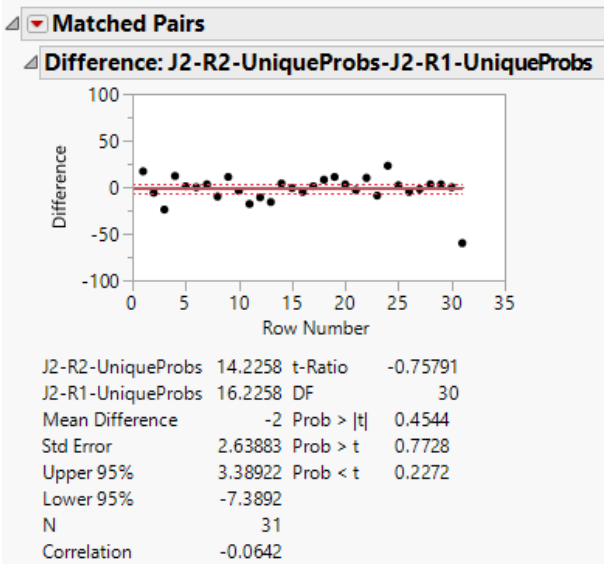


Figure 4.12: Journal 2—Reviewer 1 vs Reviewer 2 unique problems.

There were even fewer points of disagreement: Journal 1 reviewers disagreed on 1.37% of elements discussed and Journal 2 reviewers disagreed on 0.21% of the elements discussed. Figure 4.13 shows the category-by-category breakdown of the points of agreement and disagreement. The remaining percentages (93.42% and 94.44%, respectively) can be attributed to specific elements of the manuscripts that the reviewers neither agreed nor disagreed upon.

For the most part, the reviewers were discussing the same general topics; however, upon closer examination, the data show few points of direct overlap or direct disagreement in the reviewers' evaluative comments.

For Journal 1, Wilcoxon signed rank and sign tests did not show significant results when comparing Reviewer 1 and Reviewer 2's combined comments on low-level writing problems with Reviewer 1 and Reviewer 2's combined comments on the high-level writing problems. However, the results for Journal 2 were highly significant ($p < .0001$); see Figures 4.14 and 4.15.

The data reported in this section addressed the content of reviewers' evaluative comments. The next section discusses the relationship between the review content and editorial decisions.

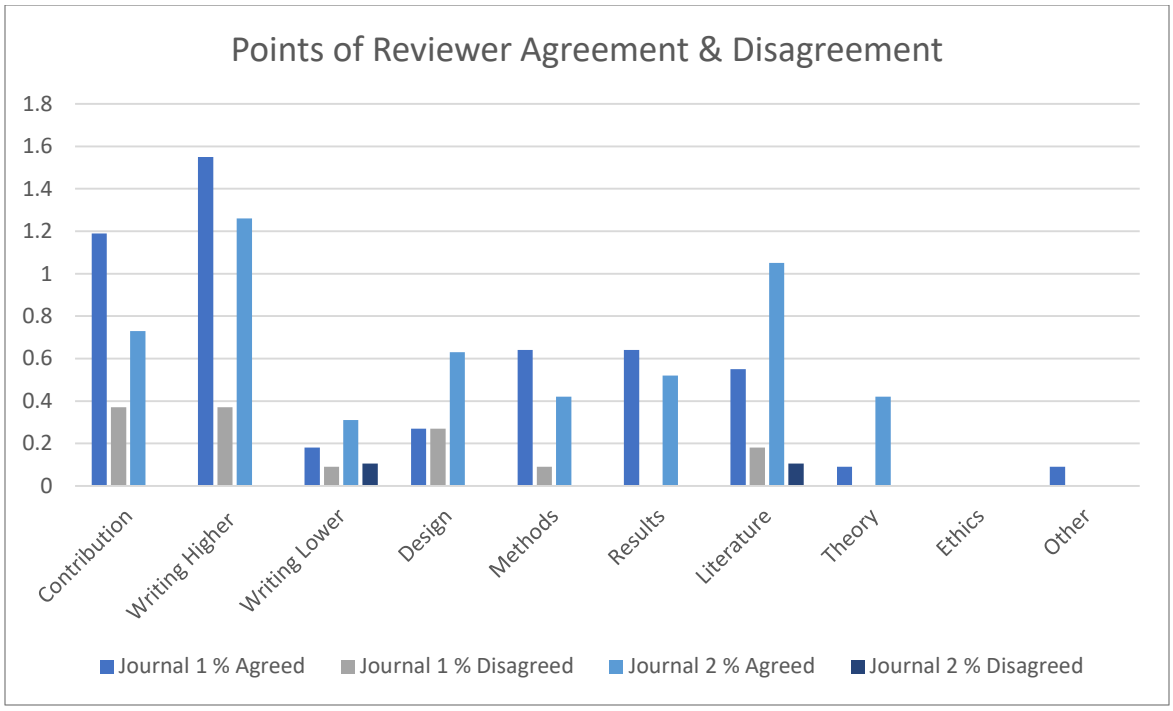


Figure 4.13.

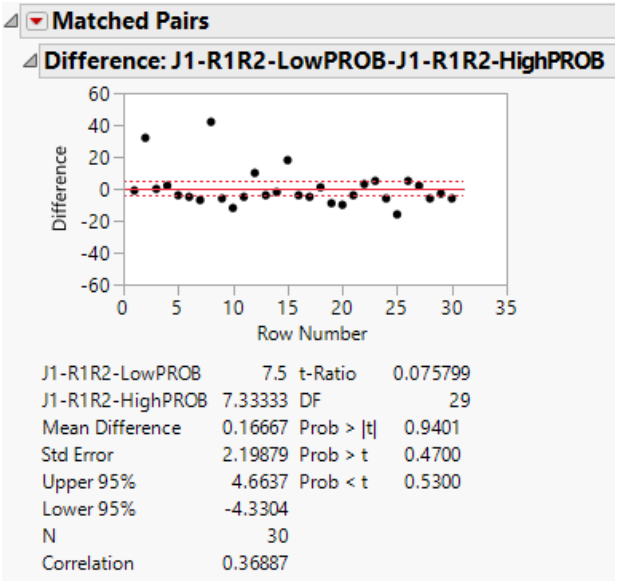


Figure 4.14: Journal 1—Low problems vs high problems.

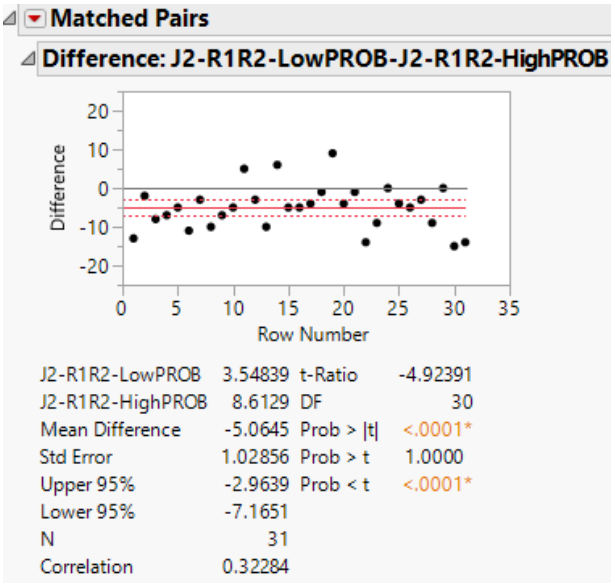


Figure 4.15: Journal 2—Low problems vs high problems.

Reviewer Publication Recommendations

Although most peer review report forms include a form field where the reviewer is supposed to indicate their publication recommendation (e.g., accept, minor revision, major revision, reject)—a form field that is not typically displayed to authors—the data in this study show that reviewers sometimes indicate their publication recommendation within other sections of the report forms, including the author-segmented (AU) section, which contains comments are that are usually shared with the authors. Based on data compiled from questions 6 and 7 of the coding form, 53% of time, the reviewers in this sample explicitly stated their publication recommendation (n = 82) in the AU or editor-segmented (ED) section of the reviewer form (see Table 4.25). Reviewers expressed uncertainty in their decision in 22% of those instances (n = 18).

The data set did not include any manuscripts that both reviewers recommended acceptance upon initial submission. All of the manuscripts in the sample were either rejected or required at least one round of revision prior to being accepted for publication. Figure 4.16 shows the extent that the reviewers agreed or disagreed with one another in their publication recommendations.

Perfect agreement, which means that the reviewers' recommendations were identical (e.g., reject/reject), occurred 5 times among Journal 1 reviewers and 23 times among Journal 2 reviewers. At the opposite end of the spectrum is perfect disagreement (i.e., accept/reject), which occurred once in this sample among Journal 2 reviewers for a revised manuscript. Between the extremes were 17 instances of one-category disagreement (e.g., major/minor) among both Journal 1 and Journal 2 reviewers (34 instance total) and a combined 13 instances of two-category disagreement (e.g., minor/reject) among Journal 1 and Journal 2 reviewers.

Table 4.25

Publication Recommendations Stated within Reviewer Report Comments

	Manuscript Types	# reviews	Publication Recommendation Stated		Recommendation Uncertainty Expressed	
			Author	**Editor**	**Author**	**Editor**
Journal 1	initial submission	60	11	16	3	6
	first revision	2	0	1	0	1
	second revision	0	0	0	0	0
	total all versions	62	12	17	3	7
Journal 2	initial submission	62	18	22	1	6
	first revision	26	4	7	1	0
	second revision	4	1	1	0	0
	total all versions	92	23	30	2	6
	Combined Totals	**154**	**35**	**47**	**5**	**13**

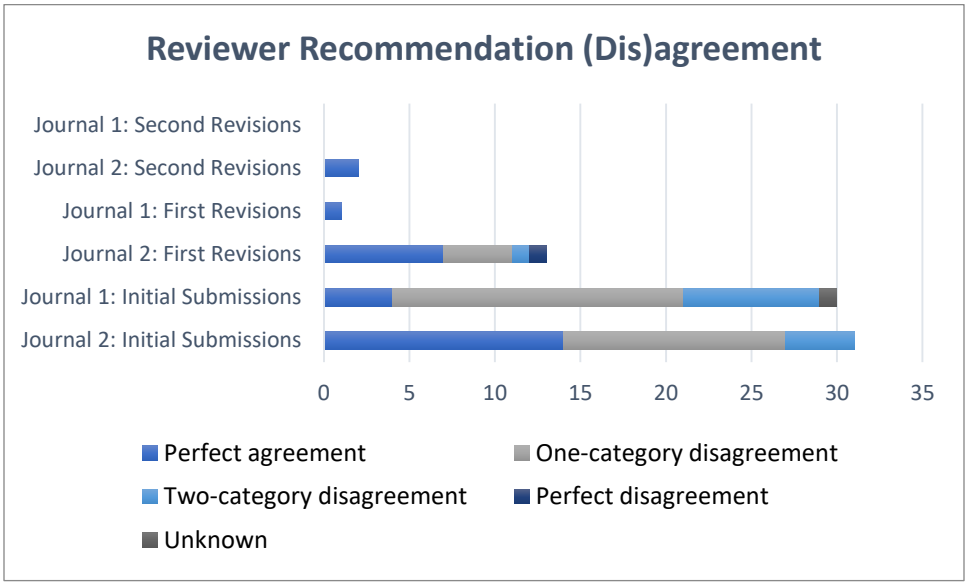


Figure 4.16: A comparison of reviewers' publication recommendations by manuscript submission round.

Problems Identified and Publication Recommendations

The number of unique problems reviewers identified was compared to the reviewers' publication recommendations. No statistically significant relationship between these variables was found for

either journal; however, for Journal 1, the results for Reviewer 1 (number of unique problems vs. publication recommendation) were nearly significant (p = .0561) for the Wilcoxon/Kruskal-Wallis rank sums test and (p = .0519) the Wilcoxon each pair method. Given those results, hypothesis 6 was not rejected. Hypothesis 6 posited that no significant relationship exists between the reviewer's publication recommendation and the number of manuscript problems the reviewer identified.

Editorial Decisions

Reviewer recommendations inform editorial decisions. Following peer review, at least 25 of the manuscripts in the data set were evaluated by an associate editor (AE) prior to the editor-in-chief (EIC) making a publication decision. With those manuscripts, the associate editor's publication recommendation matched the editor-in-chief's publication decision 84% of the time. In the 16% of cases where the AE and EIC disagreed, the editors disagreed by one degree (e.g., reject/major revision, minor revision/major revision, accept/minor revision). Agreement levels were slightly higher (87.5%) for revised manuscripts.

Journal 1 & Journal 2

The Journal 1 reviewers' publication recommendations are fairly evenly distributed compared to the editor-in-chief's publication decisions (Figure 4.17). Reviewers' publication recommendations were not available for one manuscript in this data set; only the review content and the EIC's publication decision were available. For Journal 2, the distribution of the publication recommendations reflects the number of revised manuscripts in the sample (Figure 4.18). For a better comparison, the Journal 1 distributions are shown in Figure 4.19 without the one revised manuscript and one manuscript with unknown reviewer recommendations.

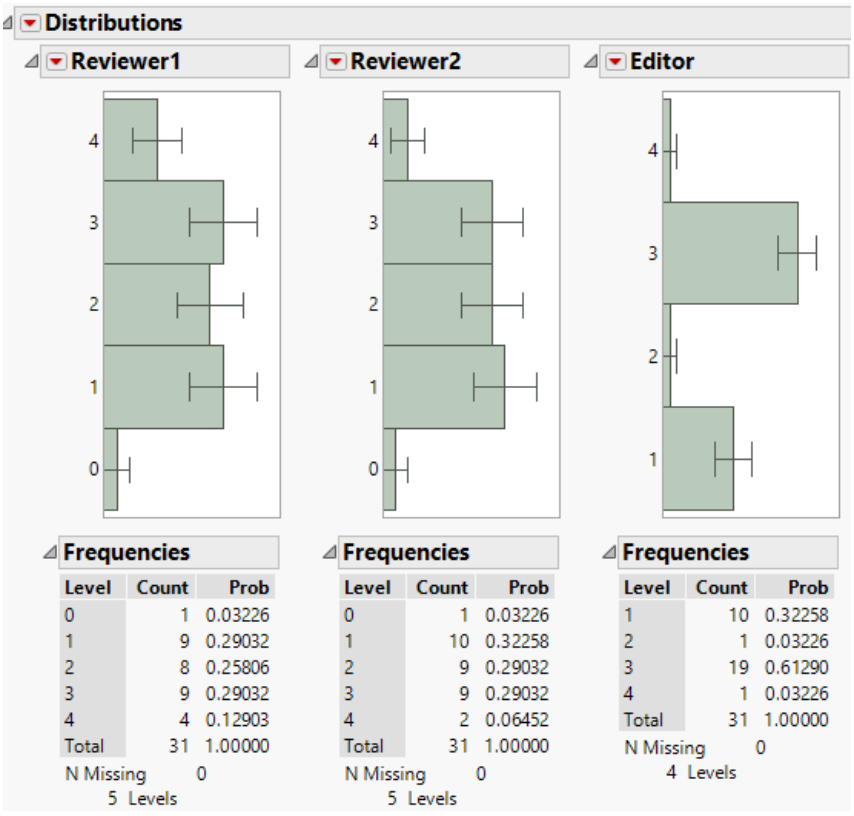


Figure 4.17: Journal 1—Reviewer publication recommendations vs editor decisions.

The Journal 2 distributions are also broken out by initial submissions (Figure 4.20) and revised submissions (Figure 4.21). Visualizing the data in these different ways is helpful for interpreting the statistical results.

An analysis of Journal 1 reviewers' publication recommendations versus the editor-in-chief's publication decision showed poor levels of agreement for the 31 manuscripts reviewed.

- Reviewer 1 vs the EIC (kappa = .324; SE = .110)
- Reviewer 2 vs the EIC (kappa = .224; SE = .104)
- Reviewer 1 vs Reviewer 2 (kappa = −0.093; SE = .113)

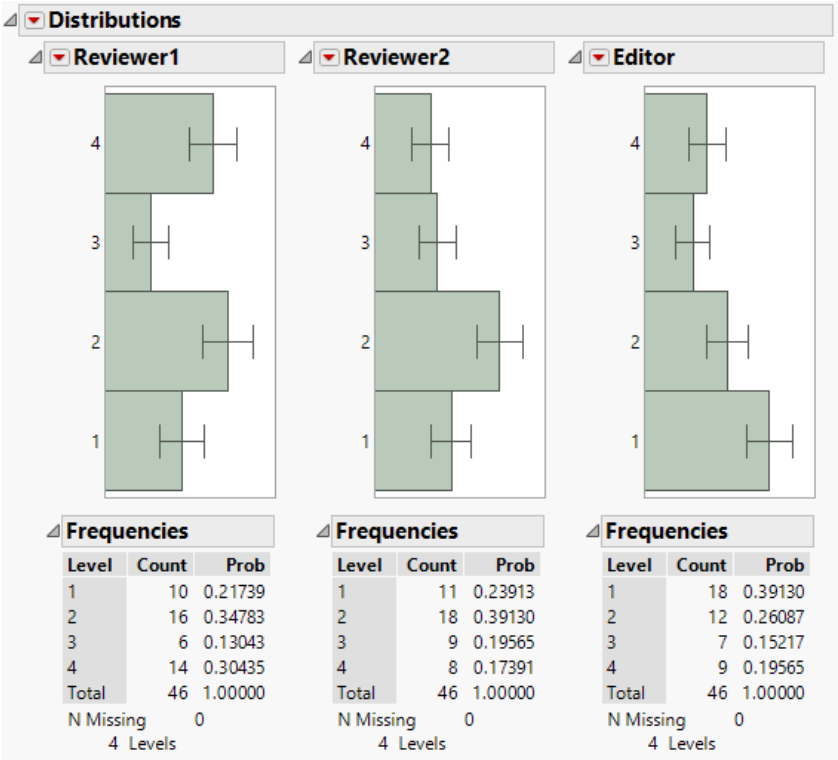


Figure 4.18: Journal 2—Reviewer publication recommendations vs editor decisions.

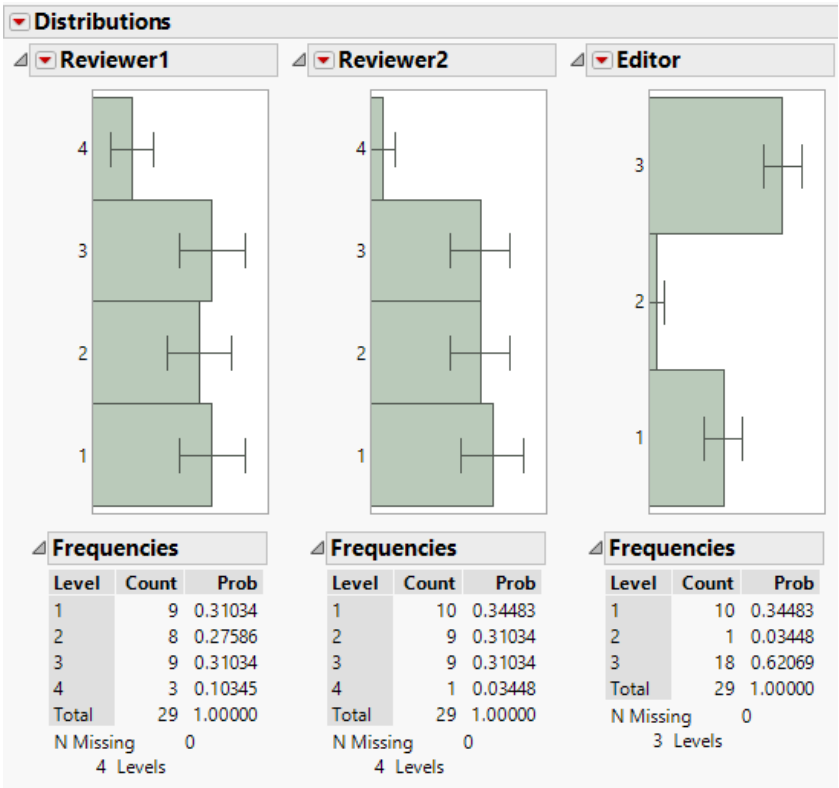


Figure 4.19: Journal 1—Reviewer publication recommendations vs editor decisions for initial submissions only.

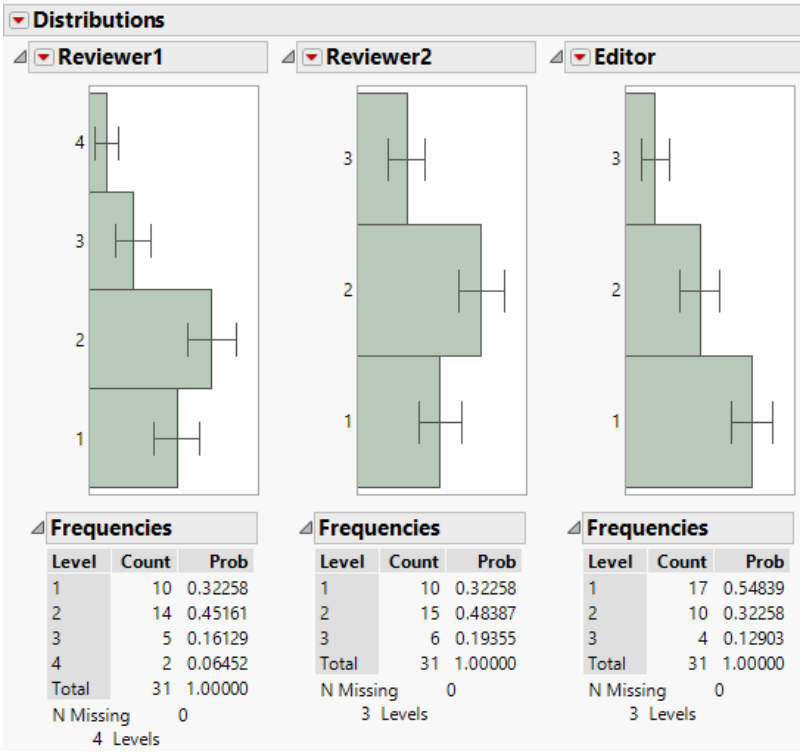


Figure 4.20: Journal 2—Reviewer publication recommendations vs editor decisions for initial submissions only.

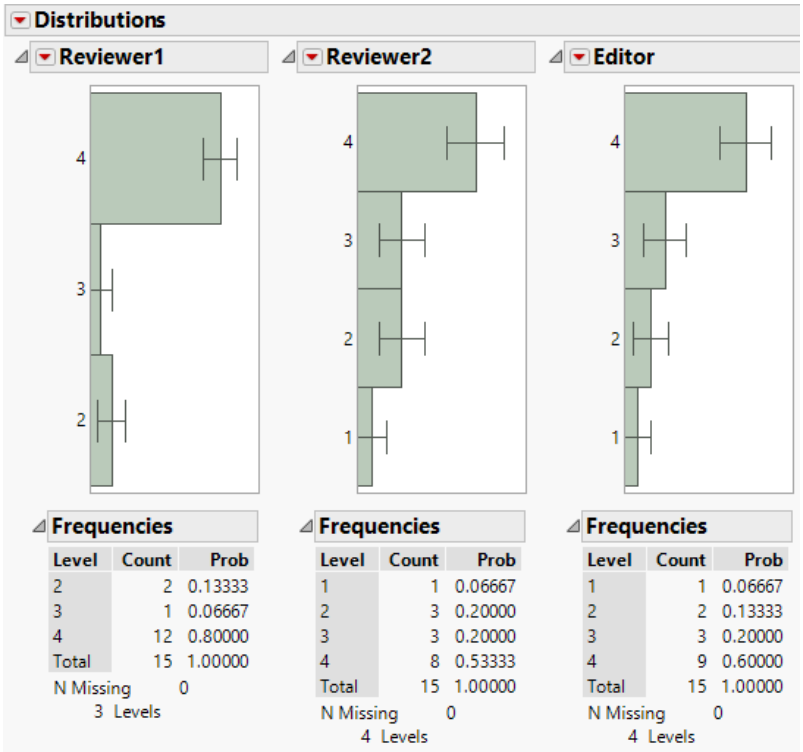


Figure 4.21: Journal 2—Reviewer publication recommendations vs editor decisions for revised submissions only.

An analysis of Journal 2 reviewers' publication recommendations versus the editor-in-chief's publication decision for the 46 manuscripts reviewed showed a higher degree of agreement than the Journal 1 data.

- Reviewer 1 vs the EIC (kappa = .387; SE = .092)
- Reviewer 2 vs the EIC (kappa = .589; SE = .092)
- Reviewer 1 vs Reviewer 2 (kappa = .318; SE = .097)

For Journal 1, Pearson tests showed significant *r* values for the comparisons of Reviewer 1 vs the EIC (*r* = .0433) and for Reviewer 2 vs the EIC (*r* = .0333) and highly significant values for Reviewer 1 vs Reviewer 2 (*r* = .0002); however, the chi squares are suspect due to the distribution of the values and the sample size.

A different picture emerges when only the initial manuscripts (*n* = 29) are compared.

- Reviewer 1 vs the EIC (kappa = .301; SE = .113)
- Reviewer 2 vs the EIC (kappa = .186; SE = .104)
- Reviewer 1 vs Reviewer 2 (kappa = −0.21849; SE = .090)

The kappa values shift slightly and the values of the Pearson correlation coefficients are no longer significant between Reviewer 2 and the EIC (*r* = .3447) or between Reviewer 1 and Reviewer 2 (*r* = .2610); the results for Reviewer 1 vs the EIC remain significant (*r* = .0340) and the chi squares remain suspect for all comparisons.

For Journal 2, Pearson tests showed *r* values less than .0001 for comparisons of Reviewer 1 vs the EIC and Reviewer 2 vs the EIC, a number that, without the suspect chi square warning, would indicate that the editor almost never agreed with the reviewers, which the raw data show is not true. A comparison of Reviewer 1 to Reviewer 2 showed significant results (*r* = .0044) with the same suspect chi square caveat.

When the initial and revised manuscripts from Journal 2 are analyzed separately, the chi squares remain problematic; however, the Pearson correlation coefficients point to some degree

of agreement between Reviewer 1 and the EIC (r = .0056, initial manuscripts; r = .0830, revised manuscripts); the results are not significantly different for the revised manuscripts. The results for Reviewer 2 vs the EIC are significantly different for the initial manuscripts (r = .0054) and the revised manuscripts (r = .0004). In contrast, the comparisons of Reviewer 1 vs Reviewer 2 did not show significantly different results for either the initial manuscripts (r = .3050) or the revised manuscripts (r = .2500).

Despite the suspect chi squares, the data as a whole (e.g., Figure 4.16) seem to support the rejection of hypothesis 3. Hypothesis 3 posited that no significant difference exists between reviewers' publication recommendations and editors' publication decisions.

Editor Publication Decisions

Regardless of the level of agreement between reviewers, editors must decide whether manuscripts warrant acceptance, rejection, or—most likely—major to minor revision. One tool that can help editors-in-chief weigh reviewers' and associate editors' publication recommendations is the mean priority score (MPS). The MPS is essentially an average of the reviewer scores. Authors who receive a revise and resubmit decision letter can use the scores to predict the final disposition of their manuscript; the data shown in Table 4.26 suggests that the higher the MPS score, the more likely a subsequent revision will result in publication. Most initial manuscripts with an MPS of less than 2 were rejected regardless of the journal; those with an MPS of 2 were harder to predict. Manuscripts with scores of 2.5 of higher usually resulted in publication following one or two rounds of revision.

In the Journal 1 sample (N = 31), the agreement between individual reviewers' publication recommendations and the editor-in-chief's publication decision is considered fair to poor depending on the literature consulted. For Journal 1, the kappa value for Reviewer 1 and the EIC

is .324; the kappa for Reviewer 2 and the EIC is .224. The maximum value of kappa is one, which indicates perfect agreement; given that the statistic accounts for chance agreement, the value of kappa can be negative.

In the Journal 2 sample (N = 46), the agreement between individual reviewers' publication recommendations and the editor-in-chief's publication decision is considered fair to good. The kappa value for Reviewer 1 and the EIC is .387, and the kappa for Reviewer 2 and the EIC is .589. Both samples include revised manuscripts (Journal 1: one manuscript; Journal 2, 15 manuscripts).

Of the 30 initial manuscripts in the Journal 1 sample, 10 were rejected and 20 were accepted; of the 31 initial manuscripts in the Journal 2 sample, 17 were rejected, 12 were accepted, and 2 were major revisions that the authors opted not to revise. (Although Table 4.26 shows 9 accepted, 12 were accepted; the table legend explains the discrepancy.) Acceptance rates cannot be calculated from the figures because of the way the sample was drawn.

In this chapter, I summarized the characteristics of the data set, described the structure of reviewer reports, and reported the results of the quantitative content analysis, comparative content analysis, and computer-assisted textual analysis. In the next chapter, I unpack the results and connect the findings to the literature.

Table 4.26

Analysis of Mean Priority Scores & Editors' Publication Decisions

	Manuscript Types	# mss	Mean Priority Scores‡							Editors' Publication Decisions			
			1	**1.5**	**2**	**2.5**	**3**	**3.5**	**4**	Reject	Major	Minor	Accept
Journal 1	initial submission	30*	2	8	8	6	2	3	0	10	1	19	0
	first revision	20**	0	0	0	0	0	0	1	0	0	1	19
	second revision	0	0	0	0	0	0	0	0	0	0	0	1
Journal 2	initial submission	31	6	6	9	7	3	0	0	17	10	4	0
	first revision	13†	0	0	2	1	1	4	5	1	2	3	7
	second revision	2	0	0	0	0	0	0	2	0	0	0	2

* The reviewers' publication recommendations for one manuscript are unknown.

** One manuscript was sent out for review following the first revision; the other 19 manuscripts were evaluated by the editor.

† The number of first and second revisions differs from the number of initial submissions because some manuscripts were not sent out for review again, some authors opted not to revise, and one author resubmitted a rejected manuscript.

‡ Mean priority score scale: 1 = reject; 4 = accept

Chapter 5

Discussion

In this chapter, I discuss the study results and connect the findings to the research questions and hypotheses. Guided by the theoretical framework of genre ecology, I discuss the significance of the findings and link them to existing scholarship.

The research questions asked about reviewer reports and review guidelines, reviewer roles, and reviewers' impact on editorial decisions and content development. The associated hypotheses tested relationships between variables such as (1) reviewers' comments and journals' reviewers' guidelines, (2) higher level and lower level writing concerns, (3) publication recommendations and decisions, (4) types and number of manuscript problems; and (5) publication recommendations and manuscript problems. The chapter is organized in clusters of research questions and the associated hypotheses. The first cluster deals with reviewer reports review guidelines.

Reviewer Reports & Review Guidelines

RQ_1: How well do the content and structure of the reviewers' evaluative comments align with the journal's reviewers' guidelines or scoring rubrics?

H_1: No significant difference exists between the content of the reviewers' evaluative comments and the content of the journal's reviewers' guidelines or scoring rubrics.

The null hypothesis associated with the first research question was rejected (H_1). The reason will be explained following the discussion of the research question.

To determine how well the content and structure of the reviewers' evaluative comments aligned with the reviewers' guidelines or scoring rubrics from the respective journals, I focused on the genre of peer review reports and analyzed the reports from multiple angles: review content and structure, reviewer response patterns, and review wording.

Review Content and Structure

Van Rooyen, Black, and Godlee's (1999) review quality instrument (RQI) was modified and used to evaluate not only the quality of the reviews but also to examine the content and structure of the reviews. By means of Likert scales, I rated the degree of alignment of various review elements (e.g., discussion of research questions, originality, strengths and weakness of the research design, methods, and writing; see Appendix C, coding form—part 2, Review Quality Instrument section). The RQI results are reported in Tables 4.11–4.22 and summarized immediately below.

Interestingly, the element that drives studies—the research question—was discussed "a little" (21%–40%) or "none at all" (27%–60%) by the majority of reviewers (Table 4.11). Similarly, the majority of reviewers (53%–72%) did not discuss the originality of the paper (Table 4.12), the strengths of the study design (62%–84%; Table 4.13), or the strengths of the methods (62%–83%; Table 4.15). Granted, a lack of discussion does not necessarily mean that reviewers did not evaluate that aspect of the manuscript. To the contrary, the results could be read as indicating that a particular aspect of a manuscript was not problematic or that other issues were more pressing. For instance, the Likert ratings move toward the middle of the scale for weaknesses in the study design (Table 4.14), and weaknesses in the methods (Table 4.16), and authors' interpretations of results (Table 4.17) with several ratings falling in the categories "a moderate amount," "a lot," or "a great deal." An absence of discussion could also indicate that the topic

was not emphasized on the reviewers' guidelines, rubric, or scoring form. A comparison of these results with the reviewers' publication recommendations and the editor's publication decisions did not reveal any notable patterns.

On the whole, reviewers devoted more discussion to higher order writing concerns (7%–45% "a moderate amount") than lower order writing concerns (7%–10% "a moderate amount"; Tables 4.18 & 4.19). Higher order writing concerns include purpose, thesis, tone, definition of terms, organization, clarity, completeness, development of ideas, and argumentation; lower order concerns include grammar, mechanics, word choice, style, citation formatting, and document design. It is not surprising that reviewers focused more on higher order concerns—most reviewers' guidelines and rubrics emphasize higher order concerns, and many reviewers' guidelines explicitly discourage reviewers from copyediting or proofreading, both of which are editorial tasks that typically align with lower order writing concerns.

In this sample, the reviewers' higher order concerns chiefly related to organization; argumentation, especially concerning assumptions and fallacies; underdeveloped ideas; and perhaps the most frequent concern: definition of terms—from missing or unclear definitions to definitions that contradicted customary or disciplinary usage (primarily the latter, and a marker that the manuscript did not belong in a technical communication journal). While definitions are key elements in any academic argument, definitions are critical rhetorical moves in the technical communication discipline—moves that help define the technical communication genre itself, as both form and social action (Miller, 1984).

Review Quality

Certainly, quality reviews are desirable; however, perhaps more consideration should be given to how the information in the reviews will be used and how easy it is to use the information. In

other words, what makes a review useful? What makes a review usable? How, if at all, do those review characteristics differ by audience? by journal? by discipline? What will authors do with the information? What will editors do with the information?

The COPE Ethical Guidelines for Peer Reviewers instruct reviewers to

> Be *objective* and **constructive** in your review, providing feedback that will help the AUTHORS to improve their manuscript. For example, be **specific** in your critique, and provide **supporting evidence** with appropriate references to **substantiate general statements**, to help EDITORS in their evaluation. (COPE, 2013, "Conducting a review")

The italicized term not only invokes Merton's norms of universalism (Zuckerman and Merton, 1971) and Popper's (2002) philosophy of empirical verification (e.g., falsification) but also the positivistic notion that peer review operates as a tool for "obtaining reliable knowledge about objective reality" (Howard, 2012, p. 326)—that peer review functions as an objective standard in the knowledge-making process.

The terms in bold face align with questions in the RQI and thus criteria of a quality review. The terms in all capital letters denote the audiences who are expected to act on the contents of the review. The underlined terms indicate the general actions the audiences are expected to perform (the editors are expected to evaluate the manuscript).

The overall review quality—the most subjective data in the study—was rated as "average" to "excellent" for about 80% of the reviews (Table 4.22). Where possible, these ratings were triangulated with the editors' ratings of review quality. Approximately half the reviews in the sample had been rated by the editors; of those, all but four received the highest rating. Even accounting for differences in scales (3 points vs 5 points), my ratings of the reviews tended to be slightly lower than the editors' ratings of the reviews, particularly for the editor portions of the reviews. Although I had read the author-segmented portions of the reviews prior to reading the editor-segmented portions, I was rating each audience-segmented section of the review

separately; the editor portions of the review were usually shorter and focused on a few key points, whereas the author portions were typically longer and covered a broader range of topics that fell within the purview of the RQI criteria. Had both sections of the review been evaluated together as one comprehensive review, my ratings for overall review quality probably would have been higher.

Irrespective of the ratings, the average length of the reviews (742 words) indicates that reviewers spent consider time evaluating the manuscripts—a few reviews were less than 100 words and a few exceeded 2,000 words. In this sample, the reviewers usually provided adequate information for the editors to perform their expected action—evaluate the manuscript (a task, that arguable, the editors could perform without reviewers' assistance, though not as expertly, efficiently, objectively, etc.). However, the reviewers did not always provide adequate information for the authors to perform their expected action—improve their manuscript.

Many of the reviews in the sample provided ample details on what was wrong with manuscripts, and the majority of the reviews (78%–98%) substantiated "some" to "all" comments (Table 4.21). Yet, many reviews required authors to read between the lines to figure out what to do with the information. The "how-to-improve" part was implied. From the author's perspective, a review that explains how to correct manuscript flaws is more useful and usable than one that leaves the author pondering how to proceed. Editors often provide guidance to authors when reviewers offer contradictory feedback or when reviewers offer suggestions that can be disregarded. However, the task of revision falls to the author. In short, the author must figure out how to fix the problems.

The majority of the reviews in the sample (70%–95%) were rated as "moderately" to "extremely" constructive (Table 4.20). The extremely constructive reviews were the ones that

laid out, step-by-step, how to fix specific weaknesses in the manuscript. This approach to peer review incorporates developmental editing strategies and is much like the approach writing instructors use when providing feedback on student papers.

The reviews in this study show wide variations in how reviews are structured (perhaps signifying an unawareness of—or resistance to—peer review genre conventions). The varied approaches to peer review and the tensions between the needs of authors, reviewers, and editors point to a disconnect between the form and function of the occluded peer review genre—particularly from the author's perspective. Among other things, the author needs to know how to fix the manuscript problems, the reviewer needs acknowledgement for the knowledge-making contribution, and the editor needs quality research to fill journal pages.

The peer review genre assumes knowledge of scholarly conventions, yet at the same time operates to enculturate—and discipline—those in a particular field, such technical communication. This assumption applies to the reviewer who must write the review, to the editor who must extract usable information from the review and relay it to the author, and to the author who must digest the criticism within the review and apply it to another genre (e.g., research article) in order to produce a publishable manuscript.

The disconnect between the form and function of some reviewer reports cannot easily be corrected without increasing labor demands; nevertheless, a few minor changes could help. For example, before submitting manuscripts, authors should run spell check and proofread the entire document so that reviewers and editors are not distracted by minor errors. When writing their evaluative comments, reviewers should consider who will be acting on the information and how the information will be used (i.e., genre as social action). In short, try to anticipate both the editor's and the author's needs when responding to a manuscript. Before selecting reviewers,

editors should consider how the pairing of reviewers will aid with editorial decision-making (e.g., what roles do the reviewers need to play when evaluating a particular manuscript?).

Reviewer Response Patterns

The reviewer reports were also coded to determine how well the reviewers' evaluative comments aligned with Gosden's (2001, 2003) response pattern norms (e.g., formulaic opening remarks, point-by-point replies, and closing remarks; see Appendix C, coding form—part 2, Review Structure section). These results, which are reported in Tables 4.1–4.6, correspond to the genre of the peer review report as enacted by the reviewers in this sample. A summary of the results follows.

The majority of the reviewers (40%–73%) structured their reports with summarizing judgments as opening remarks (Table 4.1), and about half of the reviewers included conclusion paragraphs (Table 4.3). Review comments were most commonly presented in unnumbered point-by-point formats (27%–53%), with the higher percentages occurring in the editor sections of reviews (Table 4.4). Most of the reviews contained direct criticism (47%–97%) with the higher percentages found in the author sections of reviews; direct praise (27%–77%) and praise/criticism pairs (28%–63%) moderated the negative feedback (Table 4.5). These findings are consistent with those of similar studies from other disciplines (e.g., Bakanic et al., 1989; Gosden, 2001, 2003).

Few reviewers (7%–13%) included outlines of the article being reviewed (Table 4.2); the eight instances of outlines were located in author sections of the review forms. Three reviewers reverse outlined sections of manuscripts and suggested alternate organizational schemes (i.e., a modified outline) to the authors. The other five reviewers provided high-level overviews in outline format that summarized the reviewers' understanding of the author's work as presented

in the manuscript. Unlike the summarizing judgments, the outlines were objective synopses of the manuscript content, much like an abstract. For example,

- The author hypothesized *A.*
- The author did *B.*
- The author found *C.*
- The author concluded *D.*

Though this approach was not commonly used, I would argue that more reviewers should include similar outlines because they communicate to the editor and the author that the reviewer read the entire article, and the outline may help identify shortcomings in organization, logic, topic coverage, etc. In addition to their own reading of the manuscript, the editor may find a reviewer's outline of the author's article useful in the publication decision-making process in that they can quickly assess whether the author actually did what they said they did in the abstract. It is not uncommon for abstracts to differ from the actual manuscript content (at least one review in this study notes as much)—the abstract may instead reflect the author's original intentions or a previous iteration of the manuscript.

Either way, the presence of an outline builds ethos in an anonymous peer review relationship in which, by design, the reviewers' reputation is unknown and the editor mediates the indirect conversations between the author and the reviewers. In that author–reviewer relationship, the power differential favors the reviewers (Fortanet, 2008) even though the author–reviewer conversations are mediated. Despite the power imbalance, reviewers need to cultivate trust and establish credibility. Assuming one of the primary purposes of peer review is to ensure that the journal publishes quality scholarship, then each reviewer needs to convince the author that their feedback is valid, especially when juxtaposed with feedback from the other reviewer(s) and perhaps the editor. In many cases, the reviewer must persuade the author to make the requested

revisions, though undoubtedly, some authors will feel obligated to comply to meet the publish-or-perish demands of academia.

Journal publication criteria.

Somewhat surprisingly, reviewers rarely (7%) attributed their praise (e.g., *X* meets the journal's guidelines regarding *Y*) or criticism (e.g., *X* does not follow the journal's guidelines regarding *Y*) to journal criteria outlined in the journals' authors' guidelines; in those instances when they did, the comments appeared equally in the author and editor sections of the report for initial manuscripts (Table 4.5) and only in the editor sections for revised manuscripts (Table 4.6). Examples of criticisms linked to journal publication criteria include the following:

- journal scope (e.g., manuscript topic was not relevant)
- discipline specific (i.e., manuscripts were too generic, not enough TC characteristics)
- reader's needs (e.g., manuscripts offered nothing new or nothing for TC practitioners)
- genre components (e.g., manuscripts lacked takeaways or best practices)
- empirical research (i.e., manuscripts that were misrepresented as "research" articles).

This finding could be troubling if interpreted to mean that reviewers are not basing their evaluations on the journal's publication criteria. Alternately, this finding could be interpreted as the reviewers' assumption of shared knowledge (e.g., scope is obvious to those familiar with a particular journal), or merely that the reviewers are not explicitly connecting their evaluative comments to the criteria.

Evidently, the reviewers' guidelines are not being used as a genre model in terms of structuring the review report; however, to varying degrees, they do appear to influence the content of the reports. The variations in response patterns suggest that the peer review process is largely functioning in terms of social actions (e.g., the ways in which the feedback will be used by editors and authors for various purposes) rather than in terms of genre form (e.g., formulaic structure and document design).

After examining the reviewer response patterns, I analyzed the reviews at a more granular level: at the word level.

Review Wording: Manual Analysis

In order to determine how well the reviews matched the journal's reviewer guidelines or scoring rubrics (also called manuscript evaluation instruments), I focused on (1) what was said—that is, the content and topics discussed, (2) how the information was presented—that is, the structure, format, and order of the reviewer's comments, and (3) how the information was communicated—that is, the language or wording repeated from the reviewer guidelines or rubrics. The results are reported in Tables 4.7–4.9 and summarized in this section (see also Appendix C, coding form—part 2, Review Manuscript Evaluation Instrument section).

The first two areas of interest intersect with aspects of the RQI, but the ratings are not comparable. In each of these three areas, my ratings pointed toward moderate alignment with the respective manuscript evaluation instruments. Six Journal 2 reviews (less than 10%) aligned well in all three areas. With exception of those six reviews, I could not successfully sort the blinded reviews by journal based on the review content, structure, or language alone. In other words, in blinded form, most of the Journal 1 reviews read like Journal 2 reviews and vice versa. In many respects, the reviews read like generic reviews of technical communication manuscripts—or reviews of manuscripts purporting to fit the technical communication discipline.

This homogenization of the peer review process as enacted by the reviewers for these two technical communication journals speaks to cross-pollination within the discipline itself. The technical communication field is small and scholars typically serve as reviewers for multiple journals within the field. Each journal has a distinct personality (e.g., scope, aims, professional affiliations, readership demographics), and though the data indicate that the reviewers are aware

of the types of scholarship associated with each journal as well as those readers' needs, to some degree, the cumulative interactions—or cross-pollinations—with multiple journals may impact each reviewer's approach to manuscript evaluation. For instance, if Journal 1 asks reviewers to evaluate aspect *A* of a manuscript; Journal 2 asks reviewers to evaluate aspect *B*; Journal 3, aspect *C*, etc., then the reviewer may find it more efficient to always evaluate aspects *A*, *B*, and *C* of all manuscripts regardless of the journal. Similarly, reviewers may find it more efficient to use a generic peer review template of their own devising than to customize reviewer reports for each journal.

Some journals provide comprehensive review forms that disrupt the disciplinary homogenization of peer review, both in structure and content. (For data anonymity purposes, I have intentionally not disclosed whether either journal in this sample provides such forms.) Since many aspects of the reviewers' guidelines are embedded in these forms, structurally, the forms enforce genre conventions and constrain the reviewers' evaluation. In the process, the forms essentially homogenize peer review within the journal, assuming reviewers complete the forms as intended. In the end, what reviewers say is more important than how reviewers present the information.

Review Wording: Computer-Aided Textual Analysis

To counter the subjective comparisons of the reviews with the manuscript evaluation instruments, the data were triangulated through a computer-aided text analysis (CATA), the results of which appear in Table 4.10. A mixture of manual analysis and computer analysis (Lauer, Brumberger, & Beveridge, 2018) is helpful in contexts where the researcher must balance the complexities of language nuances and large amounts of data. In this study, the reviewers' evaluative comments total nearly 103,000 words. The average length of a review is

742 words for initial manuscript submissions and 384 words for revised manuscripts; the combined evaluative comments from Reviewer 1 and Reviewer 2 average 1484 words per initial manuscript and 769 words for revised manuscripts.

The CATA compared terms in each review to terms in custom dictionaries (two journal-specific dictionaries and one broader technical communication-based dictionary; see Chapter 3 for details). Each review was compared to its corresponding journal dictionary and the broader dictionary. (Of course, the match percentages were expected to—and did—increase when the reviews were compared to a discipline-specific dictionary that contained more terms.)

While the analysis revealed significant differences between several pairs of reviews (i.e., one reviewer matched significantly more terms than the other reviewer when accounting for the review word counts), the analysis showed limited correspondence with the journal-specific dictionaries (0%–17%) and marginally improved correspondence with the broader dictionary (0%–47%). One could argue that the 30% difference between dictionary matches provides weak evidence of reviewers' knowledge of peer review practices and disciplinary expectations; more likely than not, the difference can be attributed to a knowledgeable outlier in the data set or the nature of a particular manuscript.

When the Reviewer 1 and Reviewer 2 reviews were combined by author or editor audience segments across all the entire sample of initial and revised manuscripts, the match with journal-specific dictionaries was less than 10% for both journals, and the match with the broader dictionary was less than 30%; the percentages were slightly higher for Journal 1 (Journal 1 dictionary: mean 3.7%, median 3.4%, mode 3.8%; broader dictionary: mean 27.3%, median 26.8%, mode 32.3%). When all the journal reviews were combined (i.e., combining reviewers and audiences for each journal), the results changed little. The match with the respective journal

dictionaries remained less than 10% and the match to the broader dictionary was less than 28% of the terms.

I could not find any previous research to use as a basis for interpreting these percentages, so I cannot determine whether they are higher or lower than what would be expected; an analysis of a large corpus of reviews from technical communication journals is needed to determine the normal distributions. (Other disciplines have studied various aspects of peer review, but to the best of my knowledge, none of the studies compared reviewer guideline text to review text.) Whatever the norm, for several manuscripts and the corresponding pairs of reviews, the data indicate a significant difference in the language being used by the pairs of reviewers in their evaluations.

Moreover, the eight Journal 2 reviews that were rated "strongly agree" for alignment with the manuscript evaluation instrument (Tables 4.7–4.9), matched, on average, nearly twice as many Journal 2 dictionary terms than the other reviews in the Journal 2 sample (Journal 2: subsample mean 10.2%; full sample mean 5.2%). However, that difference is less noticeable when those "strongly agree" reviews are compared to the broader dictionary; the mean (28.4%) of that subsample is only slightly higher than the mean of the full Journal 2 sample (27.9%). These eight reviews contain phrases taken directly from the reviewers' guidelines; this finding indicates that the reviewers consulted the guidelines and intentionally engaged with journal-approved peer review discourse when writing their reviews.

Five of those eight reviews had significant CATA results, which means that not only are those individual reviewers matching more dictionary terms than most of the other reviewers in the sample, but they are also matching significantly more terms than the reviewer with whom they are paired. Among those five reviewers, one reviewer matched significantly more terms in

both dictionaries. This result might be attributable to a novice reviewer who was carefully following the review criteria.

Chance, wildcards, and stop words must be weighed in the interpretation of the results. Some of the matches can be attributed to chance. Given the context of manuscript review, it is highly likely that a reviewer would use words such as manuscript (or a similar term), author (or a similar term), and terms related to the topic of the manuscript. The latter were not included in the dictionaries to ensure the data were not identifiable. Another caveat is that common words would be expected to match more frequently than uncommon words. For example, within the broader dictionary, the term *idea* is more likely to produce a match with the reviewer's comments than the term *libel* unless the manuscript topic relates to libel.

The dictionary contained numerous wildcards so that different forms of similar words were not missed in the analysis (e.g., wildcard* would match wildcard and wildcards but not wild card). The use of wildcards could have inflated the number of matches, potentially erroneously. For instance, if I used wild* instead of wildcard*, the results might include terms such as wilderness that are not relevant to this research study. Although I did not notice any spurious wildcard matches—a known risk with custom dictionaries (Garten et al., 2018)—some may have occurred. Realistically, the low percentage of matches suggests that neither chance nor wildcards impacted the results significantly. (Excluding stop words, the Journal 1 dictionary consisted of 49 terms, of which 34 were in wildcard form; the Journal 2 dictionary consisted of 41 terms, of which 32 were in wildcard form, and the broad dictionary consisted of 761 terms, of which 545 were in wildcard form. While the proportions of wildcards are high in each dictionary, most of the matches to wildcards were to the base term (e.g., the term editor* usually matched editor rather than editors).

At the same time, stop words prevented matches to common terms such as a, an, the, and other words that provided no useful information in respect to the research study. The number of matches would have increased significantly had irrelevant terms not been excluded. Presumably many of the unmatched words in the reviews relate to the topic of the manuscript.

Together, these analyses were used to test the hypothesis that no significant difference existed between the review content and the review guidelines or rubrics (H_1). The results, especially those from the CATA, indicate that hypothesis 1 should be rejected. While a few individual audience-segmented reviews matched 47% of the terms in the broader custom dictionary, the individual reviews matched less than 17% of the terms in the narrower journal-specific dictionaries. Subsequent analyses will help answer what the remaining 53%–83% of the review terms contribute to the manuscript evaluations and what those terms reveal about reviewer roles.

Taken together, these results seem to indicate that the reviewers are relying primarily on existing knowledge of peer review processes rather than on the reviewers' guidelines and rubrics. As a publication-process genre, the guidelines appear to be functioning more so in terms of social action than form. In other words, the reviewers in this sample seldom used the reviewers' guidelines as a structural model for their reviewer reports; the presentation and discourse of the reports varied considerably. However, the reviewers' guidelines—or prior knowledge—enabled the reviewers to perform the assigned task: evaluate the manuscript.

Reviewer Roles

RQ_2: What role(s) do journal reviewers in the technical communication field play?

H_2: No significant difference exists between the number of reviewer comments associated with higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) and the number of reviewer comments associated with lower level concerns (e.g., grammar, mechanics, style, citations).

H_4: For each manuscript, no significant difference exists between the types of problems each reviewer identifies.

H_5: For each manuscript, no significant difference exists between the number of manuscript problems each reviewer identifies.

Two of the null hypotheses associated with the second research question were rejected (H_2 and H_4) and one hypothesis was not rejected (H_5). Each hypothesis will be covered in further detail as the discussion of this research question unfolds.

The roles that journal reviewers in the technical communication field play have been inferred from my reading of the reviews during the coding process and my analysis of the types of problems the reviewers discussed. In follow-up studies, I will approach this research question through interviews with editors and reviewers and through close readings of a stratified sample of reviewer reports, manuscripts, editor's decision letters, and author's response letters.

The manuscript problems that bother reviewers tell us something about the roles the reviewer plays. First of all, the reviewer must identify the problems, by which I mean the reviewer must perceive and name the problem (Boettger, 2014). The reviewer's level of experience in the technical communication field will likely affect the types of problems the reviewer identifies. For example, a newly minted assistant professor who is reviewing for the first time (a novice reviewer) is unlikely to have the breadth and depth of knowledge or the nuanced understanding of a full professor (a veteran reviewer)—who may or may not have industry experience—or the specialized knowledge, experience, and perspective of a subject matter expert who works as a

practitioner in the technical communication field (a practitioner reviewer). Some TC journal editors intentionally pair academic reviewers with practitioner reviewers to obtain wider perspectives. Regardless of experience level or workplace background, all peer reviewers are cast in the role of expert—an expert who identifies manuscript problems and evaluates manuscript merits.

Reviewer Expertise

While a novice reviewer is capable of competently evaluating manuscripts within their areas of expertise, a novice reviewer is unlikely to identify some of the problems that veteran and practitioner reviewers would notice. For that matter, even if all three of these reviewers were to identify the same problems, they would be unlikely to be bothered by or to prioritize the problems the same. The same could hold true for any three reviewers within the same experience/background level; as it is, academics and practitioners are bothered by different types of errors (Boettger & Moore, 2018), which means an academic reviewer will probably discuss different types of manuscript problems than a practitioner reviewer. That is not to say that the types of problems that a novice reviewer identifies are necessarily more or less important than the types of problems that other reviewers identify. However, higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) generally take precedence over lower level concerns (e.g., grammar, mechanics, style, citations). The latter are easily remedied, while the former may render a manuscript unpublishable (e.g., a fatally flawed research design). Regardless of their expertise, the reviewers in this study appeared to prioritize manuscript problems based on their perception of a reviewer's editorial role.

Editorial Roles

Editorial roles frequently shift and overlap in the realm of academic publishing. For example, an editor may take on various editor-specific roles (e.g. copyeditor, managing editor, editor-in-chief); at the same time, the editor may assume the role of author (e.g. writing editorials or submitting articles to other journals), or the editor may serve as a reviewer for another journal. Similarly, a reviewer may take on roles usually associated with editors, such as that of a copyeditor or a developmental editor. Or, reviewers may view their role broadly as that of a problem spotter. If the goal of peer review is to facilitate the "publication of high-quality research" (Fischer, 2011, p. 227), then reviewers must prioritize identifying critical manuscript problems.

Reviewers as problem spotters.

Analyses of the reviews in this sample showed that reviewers discussed higher level writing concerns significantly more often than lower level writing concerns—an indicator that critical problems are likely being spotted. Based on the results of Wilcoxon signed rank tests ($p < .0001$ for Journal 1 and Journal 2), the hypothesis (H_2) is rejected for both journals. Hypothesis 2 posits that no significant difference exists between the number of reviewer comments associated with higher level concerns and the number of reviewer comments associated with lower level concerns. For this analysis, the comments could be positive, negative, or neutral; see Figures 4.8–4.9.

If the comparison is limited to the comments dealing with higher level problems and lower level problems (negative comments only), the results are not significant for Journal 1, but the results for Journal 2 remain highly significant ($p < .0001$); see Figures 4.14 and 4.15. If the comparison is limited to strictly writing/presentation issues (that is, without higher level issues

such as theoretical framework, data analysis, and conclusions), the Wilcoxon signed rank test is nearly significant for Journal 1 ($p = .0501$) and the results of the sign test are significant ($p = .0307$); the results remain unchanged for Journal 2 ($p < .0001$); see Figures 4.6 and 4.7.

Reviewers as pseudo-developmental editors.

Though reviewers play various roles from expert to gatekeeper, depending on the perceived purpose of peer review, the data in this study indicate that the majority of these reviewers (about 60%) prioritized higher level writing concerns (e.g., big picture problems related to organization and the development of ideas). The reviewers not only commented on the big picture concerns significantly more often than on the lower level concerns but also identified significantly more problems that required substantive or developmental editing.

This type of editing involves major restructuring or changes to content (Norton, 2009)—changes that, if approved by the editor-in-chief and enacted by the author, would have considerable impact on the manuscript content. Developmental editing can shape how arguments are framed and executed, among other things.

As a whole, the data suggest that peer review is operating as a form of social action in which the reviewers are functioning as pseudo-developmental editors in the field's disciplinary knowledge-making processes. The data do not reveal whether this is an appropriate role for reviewers. Are editors casting reviewers in this role as part of their editorial strategy? Are reviewers overstepping editorial boundaries? Is this role critical to the development of quality scholarship?

I concede that, by definition, reviewers are stepping into the role of developmental editor merely by commenting on higher level writing issues. However, the argument is not entirely circular. The extent to which reviewers in this study focused on the evaluation of the writing of

the manuscript rather than the evaluation of the research itself leads me to this conclusion. Had the majority of the reviewers' comments related to Design or Methods, I would have reached a different conclusion. Likewise, given different manuscripts, the reviewers may have focused on entirely different aspects of the manuscripts. Granted, many reviewers' guidelines emphasize higher level concerns, but those concerns extend beyond higher level *writing* concerns to research design, methods, etc.

Reviewers as copyeditors.

About 40% of the reviewers' comments (Figures 4.1 and 4.2) dealt with lower level writing concerns, which indicates that reviewers are acting in the role of copyeditor—a role that most reviewers' guidelines discourage. Granted, the percentage is misleading because some reviewers made few comments about lower level writing concerns while others made dozens. The reviewers who made a single comment stating that the manuscript needs copyediting are not acting as copyeditors. The reviewers who summarized types of recurring errors (e.g., check for faulty parallelism and punctuation errors) or who flagged one instance of each type of recurring error have begun to cross into copyediting territory. The reviewers who marked each error are playing the role of copyeditor. This finding is problematic unless the journal editor wants reviewers to serve as copyeditors.

From a pragmatic standpoint, it can be inefficient for reviewers to adopt the role of copyeditor. If the editor decides the manuscript needs a major revision, then the reviewer's copyediting work is potentially wasted—the error-ridden text may be deleted as part of the major revision (or even as part of some minor revisions). Moreover, the journal's copyeditors will correct these types of problems and apply house style, which may contradict many of the reviewer's well-intentioned suggestions.

The percentage of comments related to lower level writing concerns raises several questions.

- Did the manuscript lack higher level problems?
- Did the reviewer comment on both higher and lower level problems?
- Did the reviewer read the reviewers' guidelines or rubric?
    - If so, did the guidelines instruct the reviewer to copyedit?
    - If the guidelines did not specify copyediting, why is the reviewer copyediting?
- Were the copyediting errors so bothersome—fingers-scraping-the-chalkboard irritating—that the reviewer felt compelled to comment on them?
- Did the copyediting errors function as noise that obscured the higher level problems?
- Did the reviewer go for the low-hanging fruit (the lower level problems) due to time pressures?
- Did the reviewer feel qualified to comment on the higher level problems?
- Has the reviewer received reviews that consisted primary of copyediting comments?
- Has the reviewer been taught how to peer review manuscripts for academic journals?
- What other types of manuscript problems did the reviewer discuss?
- Are reviewers so accustomed to writing student commentary that they feel obliged to copyedit, even when they know they should not?

Types of Manuscript Problems

Based on categories used in previous studies (Bornmann, Nast, & Daniel, 2008), I coded each review for the presence of manuscript problems related to the following areas:

1. relevance of contribution
2. writing/presentation (higher order elements)
3. writing/presentation (lower order elements)
4. design/conception
5. methods/statistics
6. discussion of results
7. reference to the literature and documentation
8. theory
9. ethics
10. other.

Reviewers identified manuscript problems that fit all ten of the coding categories. The analysis of the types of problem each reviewer identified yielded significant results in some categories (Figures 4.10–4.12); however, the results are more difficult to interpret.

For Journal 1, the Pearson correlation coefficients were significant in three categories: Design (r = .0371), Results (r = .0284), and Theory (r = .0284). For Journal 2, the results were significant in two categories: Methods (r = .0484) and Theory (r = .0004). The results for Theory were significant for both journals.

Given that the results for Theory were significant for both journals, and that both journals had at least two categories with significant results, the hypothesis (H_4) that no significant difference exists between the types of problems each reviewer identifies is rejected.

These findings may point to genre differences of the manuscripts in the sample (i.e., between traditional research articles, theoretical articles, and rhetorical analyses). The findings may also suggest that some reviewers are more attuned to research design, methods, results, or theory than others. Future work could investigate across fields the roles that the reviewers take; some reviewers appear to be playing the roles of empiricists, rhetoricians, or theorists. That is, the reviewers who embrace the role of empiricist focus on the rigor of the research design and the research itself and prioritize those elements when they evaluate manuscripts. The reviewers who take on the role of rhetorician prioritize aspects of the manuscript such as argument development, logical fallacies, and rhetorical appeals. And the theorists prioritize theory building, theoretical frameworks, application and interpretation of existing theories. Of course, reviewers often take on multiple roles, including that of reader.

Another notable finding is the low percentage of problems related to ethics. Of course, an absence of identified ethical problems is a good thing, unless ethical problems indeed exist but

were not identified or were not discussed. Few reviewers appear to be playing the role of ethicist. Nonetheless, if one of the purposes of peer review is to validate scholarship, then perhaps more attention should be devoted to this aspect of manuscripts (e.g., if nothing else, explicitly confirm the absence of ethical problems).

The remaining category with low percentages of identified problems was Other. The low percentages in this category seem to indicate that the coding scheme was adequate for describing the majority of the review content.

Quantities of Manuscript Problems

Although the category-by-category comparison of problems revealed significant differences for a few categories, overall, the number of problems each reviewer identified was not significantly different. Based on the results of the Wilcoxon signed rank test results, the hypothesis (H_5), which posited that no significant difference exists between the numbers of manuscript problems each reviewer identifies, was not rejected.

Nevertheless, the data provide insight on reviewers' identification of manuscript problems. Higher level writing problems were identified by reviewers in the largest quantities. On average, each Journal 1 reviewer identified 18 problems in each manuscript, of which about 5 were related to higher level writing concerns. Similarly, each Journal 2 reviewer identified, on average, 15 problems per manuscript, of which about 3 were related to higher level writing concerns. Journal 1 reviewers identified approximately 8.6% more problems than Journal 2 reviewers; however, this difference does not mean that Journal 1 reviewers were slightly better at identifying problems; the difference merely indicates that more problems were identified in a particular group of manuscripts (see Figures 4.11 and 4.12). These findings could be interpreted

as reviewers acting as editors, teachers, or mentors; in each of these capacities reviewers are expected to identify most of the critical problems.

The number of problems related to the author's discussion of the results perhaps point to reviewers acting as quality control managers. The reviewers hold the authors accountable for the scope of the discussion and the interpretation of the data to the correctness of the conclusions and the practicality of the takeaways. The comments about these types of problems tend to be framed in terms of audience needs and takeaways.

In contrast, the number of literature-related problems and the nature of the reviewer's comments position the reviewers in matchmaker roles—intertextual matchmakers that is. Reviewers commonly provide specific examples of literature that connect to the author's topic. The reviewers commented on the literature approximately 65% of the time for Journal 1 and 60% of the time for Journal 2. Typically, those comments deal with literature gaps.

Gaps in manuscript literature reviews.

This observation about literature gaps begs the question, why are so many manuscripts missing key literature? Several possible reasons follow:

- findability issues (i.e. keyword- and algorithm-related problems)
- resource limitations (e.g., lack access to certain databases; interlibrary loans with short borrowing periods)
- unfamiliarity with TC journals (e.g., authors submitted blindly to the journals without reading any issues)
- disciplinary boundaries (e.g. viewing technical communication too broadly/narrowly)
- interdisciplinary issues (e.g., limited the scope to technical communication literature)
- Dunning-Kruger effect (i.e., you don't know what you don't know)
- bare minimum research (e.g., cherry picked literature to make barebones argument)

- timeliness (e.g., did not locate current literature)
- thoroughness (e.g., did not locate seminal works)
- amount of literature (e.g., overwhelmed by the amount of literature)

Whatever the reason for the author's omission, the reviewers usually devise corrective steps for the author; the presentation of those steps varies according to the reviewer's perception of his or her role in the publication process. Some reviewers chastise the author, while other reviewers give the author the benefit of the doubt. In the end, the role(s) the reviewer enacted—as well as their publication recommendation—will impact editorial decisions and perhaps content development.

Reviewers' Impact on Editorial Decisions and Content Development

RQ_3: In what ways do reviewers' publication recommendations and evaluative comments shape editorial decisions and content development?

H_3: No significant difference exists between reviewers' publication recommendations and editors' publication decisions.

H_6: For each manuscript, no significant relationship exists between the reviewer's publication recommendation and the number of manuscript problems the reviewer identified.

One null hypotheses associated with the third research question was rejected (H_3) and the other was not (H_6). These decisions will be explained as the discussion of this research question progresses.

To begin to understand the ways in which reviewers' publication recommendations and evaluative comments shape editorial decisions and content development requires comparing the reviewers' publication recommendations to each other as well as comparing each reviewer's recommendation against the editor's publication decision. Partly due to sampling procedures, the number of manuscripts with which both reviewers agreed in their recommendations (i.e., perfect

agreement; see Figure 4.16) was much higher than expected, particularly for Journal 2; at the same time, the number of manuscripts with which both reviewers perfectly disagreed (that is, one recommended accept; the other, reject) was lower than expected compared to the data in my pilot study sample. Even so, for both journals, comparisons of the reviewers' publication recommendations yielded Pearson correlation coefficient values that were significant when the initial and revised manuscripts were combined (Journal 1: $r = .0002$; Journal 2: $r = .0044$). When the recommendations were analyzed by manuscript status (i.e., initial manuscript submission or revised manuscript), no significant difference was found between the reviewers' publication recommendations for either journal (Journal 1: $r = .2610$; Journal 2: $r = .3050$).

Predictably, most of the reviewer-by-editor comparisons showed significantly different results. If Reviewer 1 and Reviewer 2 are not in agreement, then the editor must disagree with at least one of the reviewers.

For both journals, comparisons of each reviewer's publication recommendation to the editor's publication decision yielded Pearson correlation coefficient values that were significant when the initial and revised manuscripts were combined (Journal 1: $r = .0433$, $r = .0333$; Journal 2: $r < .0001$, $r < .0001$). However, the correlation coefficients for Journal 2 were suspect; they indicated that the editor almost never agreed with the reviewers' recommendations, which was not true. These extremely low r values can probably be attributed to a quirk of random fluctuation within the sample.

When the recommendations were analyzed by manuscript status, the Journal 1 initial manuscript results were mixed ($r = .0340$; $r = .3447$); the difference between the group of Reviewer 1s and the editor was significant, but the difference between the group of Reviewer 2s and the editor was not. In contrast, the Journal 2 initial manuscript results showed significant

differences between each reviewer group and the editor (r = .0056; r = .0054); the Journal 2 revised manuscript results were also mixed (r = .0830; r = .0004); the difference between the group of Reviewer 1s and the editor was not significant, but the difference between the group of Reviewer 2s and the editor was.

Although the chi squares were suspect for all reviewer–reviewer and reviewer–editor comparisons, the hypothesis (H_3) was rejected. Hypothesis 3 posited that no significant difference exists between reviewers' publication recommendations and editors' publication decisions.

Numerous factors complicate the interpretation of these results. The first consideration is the editor's selection of reviewers. The editor may have intentionally paired an academic reviewer with a practitioner reviewer. For manuscripts with interdisciplinary topics, the editor may have selected reviewers with different areas of expertise. In both cases, the pairs of reviewers are likely to disagree in their publication recommendations, and the overlap in their evaluative comments would likely be minimal. In turn, the editor will have to weigh the evidence and reconcile any differences in publication recommendations.

From a decision-making perspective, the editor's job is presumably easier when the reviewers are in agreement; assuming the editor agrees with the reviewers' evaluations of the manuscript, then those reviews lend credence to the editor's publication decision. But, decision-making is rarely a straightforward process; the editor must consider myriad factors. For example, why did each reviewer recommend *X*.? What aspects of the manuscript did each reviewer evaluate? What types of problems did each reviewer identify? How many problems did each reviewer identify? What aspects of the manuscript did neither reviewer evaluate?

On the whole, a lack of agreement between reviewers or between reviewers and editors is not necessarily evidence that peer review does not work. The (dis)agreement must be viewed within the full context (e.g., purpose, situation).

Reviewer Influence

Perceptions of power dynamics and editorial influence vary (Devitt et al., 2003). Some view editors as most influential (e.g., Fortanet, 2008) and some view reviewers as most influential (e.g, Enos quoted in Gale, 1998). Others contend the power is distributed between authors, reviewers, and editors (e.g., Burbules, 2014). A comparison of reviewers' publication recommendations and editor's publication decisions provides insight on the power dynamics and the shaping of disciplinary knowledge.

When the reviewers' publication recommendations differed, the majority (34 pairs) differed by only one category (Figure 4.16). The influence of reviewers' publication recommendations on editors' publication decisions could be interpreted in opposing ways; for instance, perfect agreement between reviewers could be viewed as having the most or the least influence on editors' decisions. Some might argue that, given the reviewers' consensus, the editor would have likely reached the same decision independently; therefore, the reviewers had little influence.

On the other hand, I would argue that, since editors tend to agree with reviewers' consensus decisions, the reviewers exert the most influence when their recommendations concur with one another. Assuming all things are equal (e.g., the reviewer's credentials and the editor's consideration of the issues raised by reviewers), as the reviewers' recommendations diverge, at least one reviewer's influence on the editor's decision weakens. That is, if one reviewer recommends minor revision and the other reviewer recommends major revision, then the editor is likely to agree with one of those recommendations unless the editor's reading of the

manuscript differs considerably or other variables intervene; hence, the influence of one or more reviewers weakens.

As the degree of reviewer disagreement increases, the influence of at least one reviewer may further weaken. For example, if the reviewers disagree by two categories (e.g., minor revision versus reject) and the editor decides to split the difference with a major revision—essentially averaging the recommendations in a similar manner to how Sposato, Ovbiagele, Johnston, Fisher, and Saposnik's (2014) mean priority score is calculated—then each reviewer's influence weakens to the same degree. Were the editor to instead decide to accept the manuscript, then the influence of one reviewer weakens by one decision category and the other by two categories.

The data from this study are insufficient to adequately assess reviewer influence. To go beyond conjecture and better understand how reviewers impact editorial decisions requires further research (e.g., interviewing editors). Only the editors can explain (1) what purpose(s) peer review serves at a specific journal, (2) how the editor perceives reviewers' roles, (3), what strategies the editor uses for selecting reviewers, (4) how much influence reviewers have on editorial decisions, and (5) how the editor resolves differences between reviewers' recommendations. (Anecdotally, two manuscripts were excluded from this sample because third reviewers were consulted to resolve contradictory reviews.)

Paradoxically—and still conjecturally—as the reviewers' influence weakens, the influence of the author—by proxy of the manuscript and its merits—increases, in that the editor must reconcile the opposing reviewers' recommendations by weighing his or her reading of the manuscript against the reviewers' evaluative comments. Of course, the reviewers' evaluative comments go hand-in-hand with their publication recommendations; even when the reviewers

are in agreement, the editor presumably looks beyond the publication recommendations to the evaluative comments and to the manuscript itself.

In some respects, the influence of the reviewer depends on the role(s) the editor expected the reviewer to play, the reviewer's perception of their role, and the role(s) the reviewer enacted. To a certain degree, the reviewers' guidelines/rubrics dictate the reviewer's role and the types of problems they should be identifying (assuming the problems exist in the manuscript); granted, the results of this study indicated that the reviewers' adherence to the guidelines was minimal. Even so, the reviewer's perception of their role did seem to affect the types of problems they identified and indirectly their publication recommendation.

Above all, the reviewer's perception of their role seemed to determine how they responded to the manuscript problems. For example, depending on the reviewer's perception of their role, they might respond to the author via the evaluative comments in the review in a gatekeeping manner (e.g., this is not technical communication) or they might take a mentoring approach (e.g., this is how to write a literature review). Regardless of the reviewer's approach or the final disposition of the manuscript, the reviewer can influence how manuscript content develops (e.g., authors of rejected manuscripts can use the reviewers' evaluative comments to revise their manuscript for submission elsewhere or for resubmission to the same journal).

Statements of Publication Recommendation

Complicating the decision-making process, 35 of the reviewers in this sample expressed their publication recommendation in the author section of the reviewer form and 47 did so in the editor section of the form (Table 4.25). With the journals in this study, it was not necessary to state the publication recommendation within the evaluative comments because a form field was provided for publication recommendations.

Although stating the publication recommendation in the author section of the form provides transparency, I question the usefulness of this practice, especially when the reviewer expresses uncertainty about their publication recommendation or disagrees with the other reviewer—something not known until later. How does this kind of information help the author? Does knowing that one reviewer recommended outcome *A* and the other recommended outcome *B*, help the author revise their manuscript or make the author feel better about a rejected manuscript? Does the author really need to know the reviewers' publication recommendations or merely the editor's publication decision? Knowing the reviewers' publication recommendations potentially pits one reviewer against the other or against the editor, neither of which is helpful.

The editor, on the other hand, needs to understand the reviewers' reasoning (e.g., I recommended outcome *A* for reason *X*) and be aware of any uncertainties (e.g., I can't decide between outcome *A* and outcome *B*) in order to make well-informed and appropriate publication decisions. Of the 47 reviewers who stated their publication recommendation in the editor section of the form, 13 of them discussed their decision-making dilemmas.

A spot check of 10 decision letters from each journal indicates that, in their decision letters, these editors usually included reviewers' verbatim comments. In at least one case, harsh reviewer's comments had been omitted. In contrast, no editorial changes had been made to the five reviews in which reviewers expressed uncertainty over their publication recommendation. Table 5.1 shows the recommendation options the reviewers were considering, the recommendation they made, and the editor's publication decision. In four of five instances, the editor's publication decision matched one of the publication recommendations under consideration by the reviewer. In two instances, the findings are perplexing. One reviewer

considered categories that were not adjacent (i.e., accept and major revision) and another reviewer's recommendation did not match the options mentioned in their evaluative comments.

A closer look at the 13 reviews in which reviewers expressed their uncertainty to the editor revealed other puzzling findings. First, two reviews are not represented in Table 5.2 because it was not clear which categories the reviewers were debating between. As before, the data show another reviewer considered nonadjacent categories (i.e., accept and major revision) and one reviewer's recommendation did not match the options mentioned in their evaluative comments. In this instance, the reviewer felt one part of the manuscript merited publication but another part required extensive revision to be publishable.

Somewhat predictably given the literature and the (dis)agreement findings in this study, one editor's publication recommendation matched neither the options the reviewer mentioned in their evaluative comments nor the reviewer's publication recommendation. The editor's publication decisions were listed in Tables 5.1 and 5.2 primarily to provide context for the reviewer's uncertainty. Several decision-making dilemmas related to journal fit and technical communication disciplinary boundaries—in their comments, the reviewers provided either the names of journals that would be better fits for the manuscript under review or a list of disciplines that would be more receptive to the topic and/or methodology. These findings indicate that gatekeeping is occurring in respect to knowledge-making boundaries. The reviewers are not necessarily saying the knowledge or methodologies have no merit, rather they are saying that they do not belong in the respective journals and/or the field of technical communication.

Table 5.1

Recommendation Uncertainty Expressed to Authors (Journal 1 & Journal 2)

Recommendations Reviewer Considered	Reviewer's Recommendation	Editor's Publication Decision
accept & minor revision	accept	minor revision
accept & major revision**	major revision	major revision
minor revision & major revision	minor revision	minor revision
minor revision & major revision	reject*	reject†
major revision & reject	major revision	major revision

* This recommendation does not match the options considered by the reviewer in the evaluative comments.
** This reviewer may have interpreted accept to mean minor revision.
† This decision does not match either option the reviewer considered but does match recommendation submitted.

Table 5.2

Recommendation Uncertainty Expressed to Editors (Journal 1 & Journal 2)

# Instances	Recommendations Reviewer Considered	Reviewer's Recommendation	Editor's Publication Decision
2	accept & minor revision	minor revision	minor revision
1	accept & minor revision	reject*	minor revision
1	accept & major revision**	major revision	major revision
1	major revision & reject	major revision	minor revision‡
2	major revision & reject	major revision	reject
4	major revision & reject	reject	reject

* This recommendation does not match the options considered by the reviewer in the evaluative comments.
** This reviewer may have interpreted accept to mean minor revision.
‡ This decision does not match either recommendation option the reviewer considered.

I can only speculate on the reasons for the anomalies where the reviewers recommended something other than the publication outcome options they had stated in their evaluative comments. The simplest explanation is human error; perhaps the reviewers checked the wrong box on the form. Perhaps the reviewers changed their minds at the last minute. Perhaps the reviewers wanted to save face by presenting themselves as "nice"—the least likely explanation,

though not impossible, particularly for niche subspecialties where authors could guess the identities of the reviewers. Or perhaps the reviewer wanted to appear polite (Johnson, 1992).

The data also showed that reviewers who seemed to be certain in their publication recommendations often interpreted the scope of minor and major revision very differently. Some reviewers noted that the manuscript needed only minor revisions and then proceeded to outline extensive revisions, such as organizational changes and implementation of different theoretical frameworks. (Possibly, those long lists of extensive revisions were minor compared to those that the same reviewer would have requested for a major revision.)

Together, these findings suggest that the reviewers' guidelines and rubrics could use refinement. The criteria for each publication recommendation should be clearly delineated so that the distinctions between accept/minor revision, minor revision/major revision, and major revision/reject are evident. The addition of microcopy (i.e., explanatory text beneath each selection choice) to the reviewer forms could help reviewers discern the difference between recommendation options.

Reviewer calibration also seems warranted in respect to terminology. Peer review cannot function in a useful or usable manner unless key terms are understood to mean the same thing to each person (Locke, 1700). Editors, reviewers, and authors need shared understandings of the terms *accept*, *minor revision*, *major revision*, and *reject* (or the individual journal's equivalent terms for manuscript dispositions—for consistency, this study uses these four terms regardless of the terms used by the individual journal). The characteristics of each manuscript disposition should be delineated in documents such as reviewers' guidelines not only to help calibrate reviewers but also to minimize editorial indecision.

Editorial indecision parallels concerns with Type I errors—accepting a piece that should have been rejected; a false positive—and Type II errors—rejecting a piece that's later published elsewhere; a false negative (Eden, 2008). Some would argue that Type I and Type II errors do not apply here because, in their original context of statistical hypothesis testing, these errors relate to probability and prediction. However, I am not using these terms in the literal statistical sense; I am using the terms figuratively to refer to publication decision-making errors (i.e., perceived errors in judgment). Conceivably, if both the editor and the reviewers were sensitized to these types of errors, the resultant hesitance to make the wrong decision could have a compounding effect unless the reviewers communicate their concerns to the editor. Arguably, these error constructs oversimplify decision-making in publishing contexts. Myriad factors impact what is and is not published at any given time (e.g., editorial strategies, editorial philosophies, available space, available content, editorial calendar, content strategies).

Editorial Decision-Making

In their decision-making process editors might consider the numbers of problems each reviewer identifies in the manuscript. The numbers, types, and severity of the problems are potentially useful in differentiating between a minor or major revision, for example; however, this study found no significant relationship between the reviewer's publication recommendation and the number of manuscript problems the reviewer identified. (Hypothesis 6 was not rejected based on the results of Wilcoxon/Kruskal-Wallis rank sums tests; see Tables 4.23–4.24.)

Beyond reviewer agreement, editors should consider factors such as reviewer toughness, expertise, and experience (Marsh & Ball, 1989). At times, editors receive additional decision-making input from associate editors. In those instances, the associate editor's publication recommendation initially functions as a dependent variable in the decision-making process; it

then becomes a mediating variable in the editor-in-chief's decision-making process. In this study, the associate editors and editors usually agreed.

Evaluative comments: Points of intersection.

Not only do reviewers infrequently agree in their publication recommendations, they also infrequently identify the same problems in a manuscript. The comparative content analysis (Figure 4.13) revealed that reviewers discussed the same broad topics, but the discussions rarely intersected. Journal 1 reviewers agreed on 5.21% (57 problems) of the identified problems and Journal 2 reviewers agreed on 5.35% (51 problems) of the identified problems. Between both journals, the reviewers agreed on 108 problems.

Most of the intersecting problems occurred in the categories of Higher Level Writing (29 problems), Contribution (20 problems), and Literature (16 problems). A missing or inappropriate term definition is a representative example of higher level writing problems that reviewers agreed upon. The act of defining terms is rhetorical move characteristic of the technical communication genre and the field itself; the definition (genre as form) and the act of defining (genre as social action) are markers that one belongs to the discourse community (Gee, 2014; Miller, 1984). Other in-common problems dealt with argumentation, faulty assumptions, and weak arguments.

Problems associated with the Contribution category usually related to novelty or relevance to the technical communication field; reviewers readily agreed when manuscripts contributed nothing new or when topics were not germane to the journals and their readers. In the Literature category, reviewers often noted the absence of technical communication literature, seminal literature, misrepresented literature, or disproportional literature reviews (i.e., too much lit on

one topic and not enough on another). These findings indicate that some reviewers are enforcing disciplinary knowledge-making conventions and policing the technical communication poseurs.

Evaluative comments: Points of disagreement.

At the same time, when the reviewers' evaluative comments intersected, the points of explicit apples-to-apples disagreement were rare; for example, Reviewer 1 said the apples were sweet, yet Reviewer 2 said the apples were sour. Across the entire sample from both journals, the reviewers provided contradictory feedback only 17 times. Journal 1 reviewers disagreed on 1.37% (15) of the elements discussed and Journal 2 reviewers disagreed on 0.21% (2) of the elements discussed. These results are consistent with those from Fiske and Fogg's (1990) study.

The majority of the contradictory feedback happened in the following categories:

- Contribution (4 instances)
- Higher Level Writing (4 instances)
- Design (3 instances)
- Literature (3 instances).

No instances of contradictory feedback were reported in four categories (Results, Theory, Ethics, and Other). This data cannot be analyzed statistically given the few instances of contradictory feedback and the distribution of the contradictions among categories. I can only make tentative inferences based on the content and nature of the contradictory comments.

In content and nature, the comments hinted at disciplinary tensions and knowledge-making conventions. For example, in the Contribution category, pairs of reviewers said certain manuscripts (1) did/did not contribute knowledge to the field, (2) were/were not useful to readers, or (3) were/were not interesting to readers. Contradictions in the Higher Level Writing category included disputes over terminology usage and whether a revision did/did not improve the manuscript.

Similarly puzzling were diametrically opposed views on the merits of research designs and literature reviews. Disagreements over research designs could indicate that one reviewer was less experienced than the other reviewer; however, experience level cannot be determined from anonymous reviews. Disagreements over research designs could reflect differences in reviewers' training, qualitative versus quantitative biases, or other factors that cannot be determined from anonymous reviews. Disagreements over the literature reviews could speak to the reviewers' perceptions of the technical communication field (i.e., a broad view or a narrow view) or could be attributed to one or more of the postulated reasons for disagreements over research designs, among other things.

Publication recommendations: Reviewer (dis)agreement.

The literature from other fields discusses reviewer disagreement extensively (e.g., Fischer, 2011; Gebhardt et al., 1995) as well as shortcomings of statistical calculations of interrater reliability (e.g., Perreault & Leigh, 1989; Stemler, 2001). Meta-analyses of studies on peer review reported low Cohen's kappa values that averaged .17, and the highest scores generally corresponded to small samples (Bornmann, Mutz, et al., 2010); for comparison, here are the figures for this study: Journal 1 (kappa = −0.093, SE = .113) and Journal 2 (kappa = .318, SE = .097). These results are consistent with studies that found reviewer agreement to be little better than chance (e.g., Kravitz et al., 2010; Rothwell & Martyn, 2000).

Counterintuitively, I would argue that these findings are evidence that peer review is working efficiently, assuming the purpose of peer review is to facilitate the publication of quality research. These findings confirm that the reviewers are focused on different aspects of the same manuscript; together, the reviews provide a more comprehensive evaluation than had both reviewers examined the exact same aspects of the manuscript. From a quality standpoint, it does

not matter that the reviewers identify few of the same problems. What matters is that the reviewers identify as many problems as possible, particularly the critical problems. Much like usability testers, the reviewers, through their combined yet independent efforts, should detect most of the problems in the manuscript.

Reviewers frequently disagree in their publication recommendation partly because they are essentially comparing apples and oranges. There is minimal overlap in the review content. Were neither reviewer to evaluate the apples, that incomplete evaluation or omission could be far more problematic than disagreement over publication recommendations.

In terms of practical significance, does it really matter whether reviewers agree or disagree, much less whether the (dis)agreement is statistically significant? Possibly. It depends on the purpose of peer review as perceived by the editor (or the journal's board, the discipline, the tenure-conferring department, academia, etc.).

While we do not know whether editors expect peer review to provide valid, reliable data to support decision-making, if editors do rely on reviews in this way, then the interrater reliability (IRR) numbers may matter. However, unless the editor selects pairs of reviewers who are identical in most every respect (Hirschauer, 2010), who diligently follow standardized review procedures (Gosden, 2003; Hirst & Altman, 2012) that are specified in minute detail, and who evaluate the same aspects of the manuscript (Fiske & Fogg, 1990), then the editor will likely fail to achieve desirable IRR numbers.

If the editor conceives of peer review as a gatekeeping mechanism, then the numbers probably do not matter. Reasons to reject a manuscript are plentiful—the IRR is not relevant if the goal is to find a reason to reject.

If the editor views peer review as a publication-process genre that operates as both a form and a social action, then the numbers do not matter. The editor is more concerned with the publishing process and producing a quality final product.

If the editor views peer review from the perspective of usability, then the IRR numbers are irrelevant. The editor is more concerned with the users' needs.

If the editor views peer review through the lens of content strategy, the numbers still do not matter. The editor is focused on developing quality content that meets the stakeholders' needs.

Unless the purpose of peer review is defined in positivist terms, then what reviewers say about the manuscript is more important than whether the reviewers agree with one another.

In this chapter, I discussed the results that directly relate to my research questions and hypotheses as well as the significant findings. In the next chapter, I outline the conclusions, limitations, future research, and implications of this study.

Chapter 6

Conclusions

In this chapter, I (1) draw conclusions based on my interpretation of the results, (2) note the limitations of my findings, (3) outline recommended actions, and (4) suggest future lines of inquiry. I conclude with a recap of the most important results; discuss the implications for technical communication, technical communication pedagogy, and practitioners; and explain how these findings advance our understanding of how peer review shapes technical communication scholarship.

Genre theory provided a flexible framework for analyzing peer review and its subgenres as both form and social action (Miller, 1984); functioning as a genre ecology or ecosystem (Spinuzzi & Zachry, 2000), the peer review subgenres work together to support the editor-mediated production of technical communication scholarship. The hypothesized rhetorical context (Figure 6.1) in which the genres operate has been revised to reflect my current understanding of the complex interaction of variables that shape content and, in turn, knowledge (Figure 6.2). A shaded hexagon has been added to the diagram to represent editorial strategies (ES) because ES appear to link the elements within the peer review genre ecology. Editorial strategies may involve

- editorial roles (e.g., primary roles of author, reviewer, editor)
- editorial tasks (e.g., writing, reviewing, editing)
- journal policies (e.g., double-blind peer review)

- standards (e.g., writers' guidelines, reviewers' rubrics)
- publication goals (e.g., quality scholarship)
- audience needs (e.g., academic vs. practitioner)
- content priorities (e.g., empirical research)
- organization (e.g., genre conventions; Halvorson & Rach, 2012).

Further research is needed to fully understand key aspects of the ES, particularly editors' perceptions of reviewers' editorial roles and reviewers' editorial tasks.

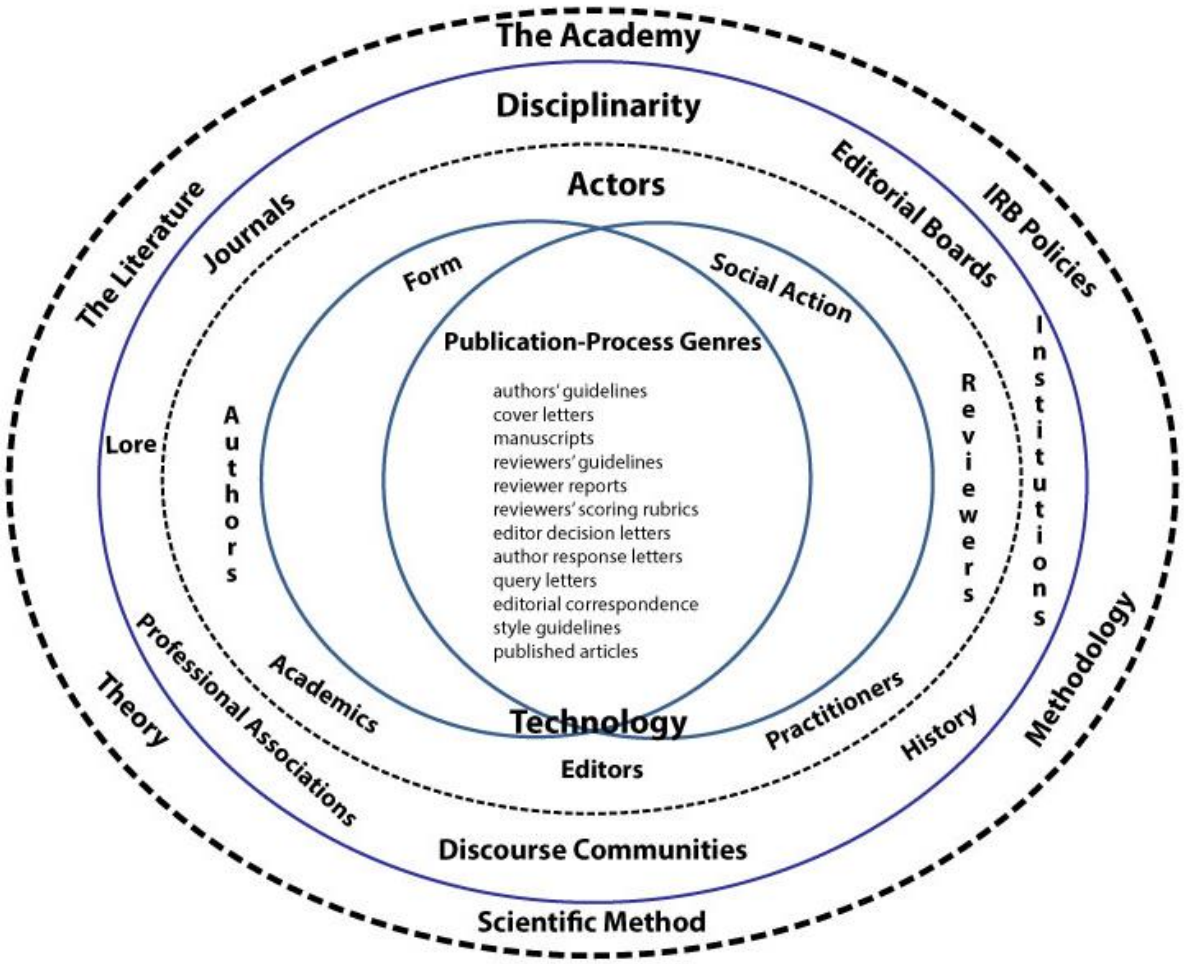


Figure 6.1: Peer review genre ecology as originally hypothesized.

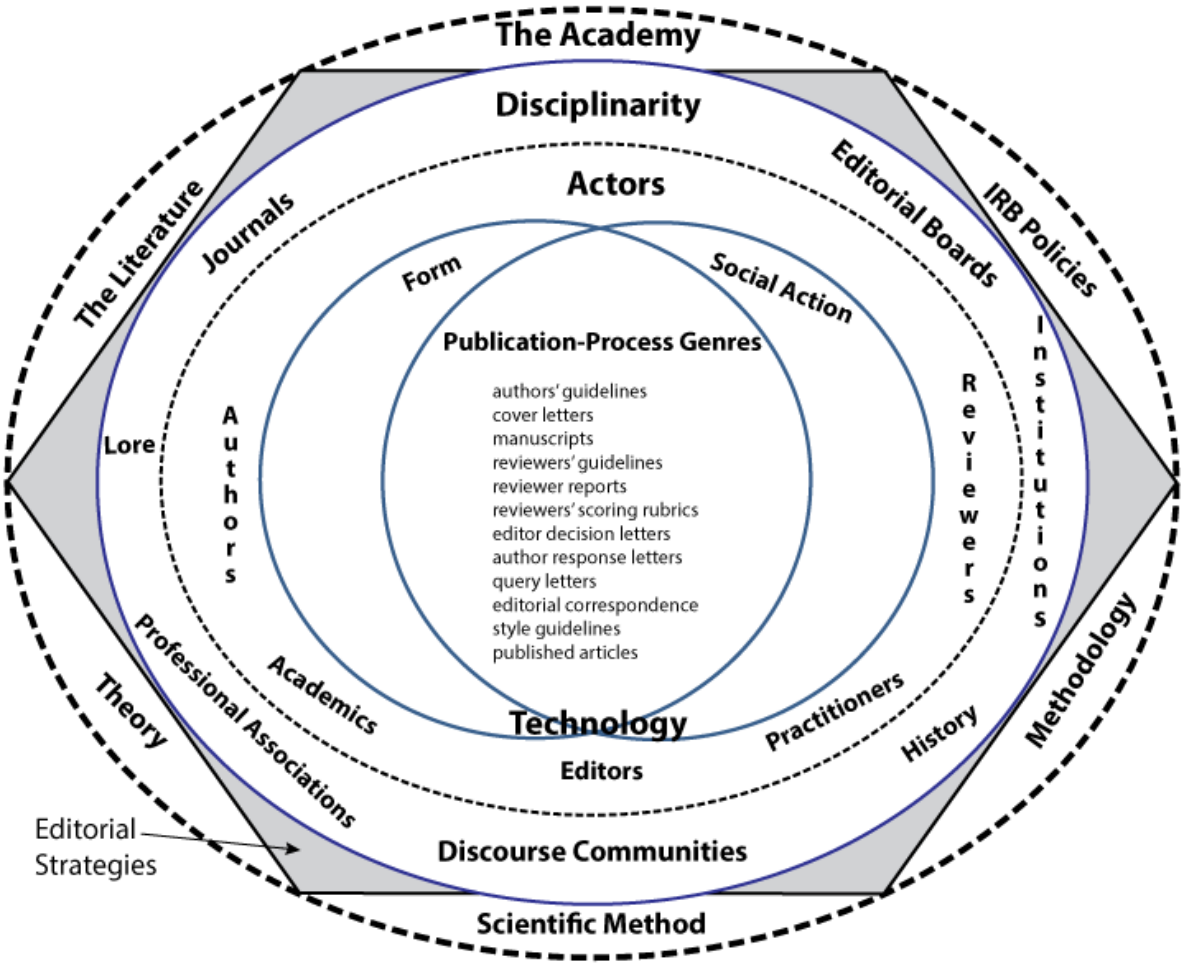


Figure 6.2: Revised peer review genre ecology.

Presently, several conclusions can be drawn from the analysis of these technical communication publication-process artifacts. (Conclusions that relate to the study's research questions are indicated in parentheses by the corresponding question number, e.g., RQ_1.)

- First of all, the analyses of these publication-process artifacts (e.g., reviewer reports, reviewers' guidelines, and editors' decision letters) indicate that, for these two journals, the peer review genre is functioning largely as social action (e.g., the feedback is being used for specific purposes—to inform decision-making and to improve manuscripts) rather than as form (e.g., structure and document design). The

disparate physical appearances of the reviewers' reports serve as evidence that few reviewers used the reviewers' guidelines for guidance on the form aspect of peer review. Form entails more than document design and structure; it also comprises discourse, content, and categories—the computer-aided text analysis revealed very little alignment (on average, less than a 10% overlap of terms) between the discourse and content of the reviewers' reports and the reviewers' guidelines (RQ_1).

- Though the reviewers' guidelines had limited impact in respect to form, most of the reviewers seem to have internalized the peer review genre conventions in terms of social action. In multiple respects, the data appear to support the conclusion that the peer review genre is shaping the discourse community's social structure (Luzón, 2005). Social action comprises disciplinarity, editorial practices, rhetorical actions, and knowledge-making. The data show, for instance, disciplinary gatekeeping at work (e.g., reviewers' comments about articles that do not fit TC journals) and reviewers acting as copyeditors. Some reviewers employed Socratic questioning techniques to help authors develop their ideas, while other reviewers prescribed ways of knowledge-making. The academy conceptualizes peer review as an objective evaluation practice, yet gatekeeping, editing, developing, and prescribing are not neutral acts. They are social actions embedded within the genre conventions of peer review.
- Peer review need not be objective to work—at least not if the purpose of peer review is to facilitate the publication of quality research. To that end, the data in this study support the conclusion that peer review generally does work, particularly when reviewers and editors play the roles of (1) problem spotters—roles with similarities to

those of usability testers (roles usually associated with design fields) and (2) developmental editors.

In the context of publishing, the knowledge product (i.e., the manuscript) is tested against various criteria (as outlined in authors' and reviewers' guidelines) by reviewers (problem spotters) much like a design is tested by representative users for usability issues. The goal of usability testing is to catch as many problems as possible. Extrapolating from usability research, an editor could expect two reviewers—representative users—to detect about half the problems in a manuscript, and three reviewers, about 65%; yet, 15 reviewers would be needed to find every problem (Nielsen, 2000, 2012).

Many reviewers go beyond "testing" and embrace the role of developmental editor. That is, the reviewers document the manuscript problems and then outline action steps for remedying the problems. By doing so, these reviewers are intervening—albeit sanctioned via editor mediation—in the knowledge-development process and crossing observer–participant boundaries. In that respect, those reviewers are no longer objective evaluators; however, those reviewers who act as pseudo-developmental editors are providing useful, usable feedback to the authors—resolving the disconnect noticed in some reviews (i.e., reviewers identified problems but did not offer solutions)—and in the process the reviewers facilitate the production of high quality scholarship.

- This study did not evaluate the number of problems in manuscripts, only the number of problems that reviewers discussed in their reports; however, the comparative content analysis showed that reviewers infrequently identified the same problems.

This finding can likely be explained by Sposato, Ovbiagele, Johnston, Fisher, and Saposnik's (2014) observation that different reviewers will detect different flaws that align with their areas of expertise. Similarly, reviewers may evaluate manuscripts from different zoom levels—that is one reviewer may look at the manuscript as a whole (wide zoom) and another may home in a specific aspect of a manuscript (narrow zoom; Spool, 2019).

- If quality is the desired outcome, then reviewers should look at manuscripts from different perspectives and different zoom levels. From the lens of usability testing, an effective peer review is one that not only identifies manuscript problems but also leads to successful revision of the manuscript (assuming it was not rejected by the editor). In this respect, the data point to a usability gap. According to Nielsen (2000), usability testing requires going beyond documenting weaknesses; the weaknesses need to be fixed. However, some peer review reports document manuscript problems and indicate the reviewer's publication recommendation. Nothing more. From the author's perspective, this approach is like notifying someone that they have a flat tire, but not offering to help them change the tire, not providing tools and instructions for changing the flat, and not offering them a ride. In that respect, peer review is not working optimally—reviewers' labor considerations aside. On the other hand, some peer reports provide step-by-step details to guide authors in their revisions.
- Reviewers play many roles (RQ_2) and adopt various strategies. Some reviewers may perceive themselves as playing the role of objective evaluator whose main objective is to certify knowledge as valid and reliable. Others, through editor-mediated

conversations with authors, cross the observer–participant boundary and through their evaluative comments actively work to shape content, and subsequently, knowledge.

- Based on the data from this study, I conclude that peer review operates as a type of social action in which reviewers internalize the generic conventions of journal scholarship and help authors shape content much like developmental editors do; the reviewers function as pseudo-developmental editors in the field's disciplinary knowledge-making processes (RQ_3). Within the genre ecology of peer review, these reviewers are indeed focused on the quality of the knowledge product, but their primary considerations appear to be higher level writing problems, knowledge-making, development of arguments, compliance with disciplinary conventions, and enculturation of junior scholars. These reviewers seem to be thinking of the end product in terms of journal readers who will use the information for specific purposes that meet their specific information needs.

Limitations

Sample-Related Limitations

Results from this study cannot be generalized due to the small sample size (61 initial manuscripts and 16 revised manuscripts); approximately 400 manuscripts are needed to obtain statistically significant results that can be extrapolated to other populations (Thayer et al., 2007). For perspective: 400 manuscripts is equivalent to the combined number of manuscripts published over five years by *IEEE Transactions on Professional Communication*, the *Journal of Business and Technical Communication*, the *Journal of Technical Writing and Communication*, *Technical Communication*, and *Technical Communication Quarterly*. This publishing history brings into

question the idea of needing 400 manuscripts for generalized results. Statistics assume an infinite population, which is not the case here. The results of this study may be stronger than they appear. However, only two of those five journals are represented in these results; results may differ when data from the other three journals are included.

The data may be skewed in ways that are not apparent due to various factors (e.g., permissions, sampling criteria, the anonymity of the peer review process, editors' reviewer selection process). Some reviewers may be overrepresented in the data set (i.e., the data set may include multiple reviews from the same reviewer—within and across journals). Likewise, authors may have multiple manuscripts in the data set. The distribution of authors and reviewers may have confounded results in ways that are not obvious. I could not control this distribution because of the blinding process.

As a former managing editorial assistant, my familiarity with these types of publication-process artifacts may have influenced how I coded and interpreted the data.

The decision to separate the reviewer reports into audience-segmented files (i.e., author and editor) may have affected some analyses because some of the author and editor sections were blank. The similar shape of the graph lines in Figure 4.2 suggest that the number of report forms with blank author or editor sections had limited impact when the author and editor sections of the report were viewed as a whole.

Coding-Related Limitations

Ideally the research assistants would have coded all of the review files and I would have merely analyzed the results. Due to limited resources, I coded all of the review files and the research assistants coded a subset of the sample to establish the reliability of the coding scheme. As

recommended by the literature, I reported the interrater reliability scores (IRR) scores at the beginning, middle, and end (Krippendorff, 2019).

The IRR calculated at two intervals suggest a learning effect occurred in some categories, particularly Research Design/Conception. One research assistant struggled with that coding category during training and pilot testing but improved significantly when coding the reliability sample (i.e., about 20% of the sample).

Fatigue effects were apparent in other categories; for instance, the research assistants sometimes miscoded Writing/Presentation (Lower Level) as Writing/Presentation (Higher Level). With theoretical manuscripts or ones that featured rhetorical analysis, the research assistants and I debated between coding certain reviewer comments as Discussion of Results or Writing/Presentation (Higher Level).

Order-related limitations.

Order effects may have had some impact on the coding reliability; after pilot testing, the research assistants leapfrogged files to cover more of the sample—they were coding different files, usually from different journals. In contrast, I coded all of the Journal 1 files and then coded all of the Journal 2 files.

Furthermore, the research design stipulated that the author portion of the reviewer report be coded before the editor portion of the reviewer report, a design that may have produced order effects, particularly with the Likert ratings of the editor portions of the reports. (The editor portions tended to be shorter and focused on a few key points.) That design was chosen with follow-up studies in mind to ensure compatibility; publication-process artifacts later collected directly from authors are unlikely to include the editor portion of the reviewers' reports.

Rating-related limitations.

Many of the Likert ratings from the review quality instrument (RQI) skewed toward the lower-to-middle end of the scale. This apparent trend can be attributed to two factors: the coding unit of analysis and the number of categories coded. The unit of analysis varied from a single word to several sentences to avoid overlapping codes because a quantitative content analysis does not permit the use of multiple codes on text (Boettger & Palmer, 2010). The decision to use a varied unit of analysis seems justified given that 63.3% of Journal 1 reviewers and 48.4% of Journal 2 reviewers used feedback approaches with praise and criticism in the same sentence.

My RQI ratings were informed by the number of codes for each category only; my ratings did not account for the length of the text. Depending on the unit of analysis, a 100-word review and a 1,000-word review could potentially have the same number of codes. Weighted counts that accounted for the number of codes per *X* words would have more accurately represented the results.

Similarly, the Likert ratings could potentially be skewed if each topic were discussed equally but the number of categories coded differed considerably (e.g., 2 categories versus 10 categories). I considered those factors when rating items but not in any scientifically controlled manner; nonetheless, the semi-objective raw coding counts tempered the subjectivity of the Likert ratings. Future studies should use weighted counts.

Recommendations

Based on the findings of this study, I recommend that TC editors take the following actions:

- Provide comprehensive definitions of each manuscript disposition option. In other words, define the terms *accept*, *minor revision*, *major revision*, and *reject* (or the

journal's equivalent terms) and describe the characteristics of each to help calibrate reviewers' publication recommendations and reduce uncertainty in decision-making.

- Add microcopy to the reviewer recommendation portion of peer review forms. The microcopy should briefly define each recommendation option (e.g., minor revision, major revision). Alternately, consider adding help text that appears when the cursor hovers over the terms, or include links to the comprehensive definitions and the reviewers' guidelines.
- Specify reviewers' expected role(s) and tasks. Alternately, list role(s) and tasks that exceed the scope of reviewing, such as copyediting. While reviewers' service is invaluable in developing the TC community and its body of knowledge, reviewers' designated roles and tasks should not pose undue burdens. Reviewers are volunteers and their service has limited value to them in terms of career advancement. Any efforts to improve peer review (e.g., its validity, reliability) must be tempered by considerations of labor commitments—those of editors and reviewers.

Future Research

As often happens, this study generated more questions than answers, providing fodder for both short-term and long-term research goals. Above all else, future research should include data from additional technical communication journals in order to provide a more comprehensive picture of disciplinary knowledge-making practices.

Approximately 400 manuscripts and the associated publication-process artifacts are needed to obtain generalizable results—my long-term goal. Obtaining such a large sample will be difficult given the occluded nature of the peer review genre, the number of articles published in the technical communication field, permissions, and participant response rates. (Depending on

journal acceptance rates and the desired time period, one could almost argue that 400 manuscripts would not be a representative sample but nearly the entire data set.) As it is, I tried to get manuscripts from the five major journals mentioned and three journals declined to participate in the study.

Given that I could not find any previous research to use as a basis for interpreting the CATA results, an analysis of a large corpus of reviews from technical communication journals would be useful for determining the distribution, frequencies, and patterns of peer review discourse in various publication-process artifacts.

Additional areas of future research are listed below by topic area.

Author-Related Studies

- **Literature Review Gaps:** An investigation of the gaps in author's literature reviews and the reasons for them is warranted considering the number of reviewer comments on this topic. Possible explanations for these shortcomings are outlined in Chapter 4. The study would require further analysis of reviewers' reports as well as analysis of initial and revised manuscripts and interviews of authors and perhaps reviewers and editors.
- **Manuscript Revisions:** The goal of this study is to understand how content is developed at the manuscript level. The study would examine what authors do with the feedback they receive from reviewers and editors and how authors respond to revision requests. This study would involve further analysis of reviewers' reports as well as an examination of editors' decision letters, authors' response letters, initial manuscript submissions, and revised manuscripts.

Reviewer-Related Studies

- **Reviewer Roles:** Another productive line of inquiry would be a multidisciplinary study of reviewers' view of their roles (e.g., empiricists, rhetoricians, and theorists) and the relationship between those roles, reviewers' areas of expertise, and manuscript genres (e.g., traditional research articles, theoretical articles rhetorical analyses). Such a study might involve textual analyses and interviews or surveys.
- **Reviewers' Use of Guidelines:** This study would investigate the use of journals' reviewers' guidelines. With adequate resources, the study could be designed to use screen capture, keyboard stroke logging, or other methods that would unobtrusively record reviewers' behaviors rather than their perceptions of their behaviors. Alternately, the study might ask reviewers about their use and how the guidelines could be improved.
- **Form Quirks:** This minor topic might be pursued as part of another study. Among other things, I am curious (1) why reviewers state their publication recommendations in the author section of review forms, (2) why reviewers submit the same text in the author and editor section of the form, (3) why reviewers tell authors they are uncertain in their publication recommendation, and (4) whether the form prevents reviewers from providing certain types of feedback or requires them to provide feedback that seems irrelevant for certain types of manuscripts. In other words, in what ways, if any, is the form shaping or constraining their feedback?
- **Private Conversations:** This study would compare the content of the editor and author sections of the review forms to assess the nature of the information that is not being shared with authors. What details do reviewers discuss with editors but not with authors?

What can be learned from those private conversations that might be used to improve manuscripts or disciplinary practices?

- **Reviewers as Copyeditors:** This study would investigate why so many reviewers copyedited the manuscript under review. Possible explanations are outlined in Chapter 5. The study would involve further analysis of the reviews as well as the corresponding manuscripts. Any reviewer interviews or surveys would have to be conducted in tandem with the collection of additional reviews so that the artifacts can be linked to the anonymous reviewer.

Editor-Related Studies

- **Reviewer Selection & Roles:** This study would be designed to identify the strategies that editors use when selecting reviewers and the reasons for each strategy. The study would also investigate editors' perceptions of reviewers' roles and the purpose of peer review and the usefulness of each reviewer role. This study would involve interviews, surveys, or focus groups of former and current editors.
- **Reviewer Reports:** This study would investigate the usefulness of reviews for editorial decision-making and identify the characteristics of useful, usable reviews. This study would involve further analysis of reviewers' reports, editorial decision letters, and interviews with editors.
- **Review Terminology:** This goal of this study is to determine how each journal defines the terms *accept*, *minor revision*, *major revision*, and *reject* (or the equivalent terms used by the particular journal). This study might involve interviews with editors and editorial board members as well as analysis of the journals' authors' guidelines, reviewers' guidelines, and other relevant artifacts.

Key Findings

Key findings from this study relate to genres, manuscript problems, peer review practices, and publication recommendations and decisions. (Findings from hypothesis testing are indicated in parentheses by the corresponding hypothesis number, e.g., H_1.)

Genres

- Within and between journals, the peer review reports varied considerably in structure. As a genre form, the reviews lacked consistent presentation. Reports ranged from a single paragraph to multiple pages with reviews reading like letters, reports, bulleted lists, to Q&As, etc.
- A significant difference was found between the content of the reviewers' evaluative comments and the content of the journals' reviewers' guidelines (H_1). Reviewers seemed to rely on prior knowledge of the peer review genre (as a social action) rather than the reviewers' guidelines.

Manuscript & Review Content

- A significant difference was found between the number of reviewer comments associated with higher level concerns (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) and the number of reviewer comments associated with lower level concerns (e.g., grammar, mechanics, style, citations; H_1).
- A significant difference was found between the types of problems that reviewers identified (H_4); however, no significant difference was found between the number of problems each reviewer identified in manuscripts (H_5). On average, each reviewer discussed 15 to 18 unique problems per review.

- Problems with literature reviews were common; issues ranged from lack of breadth or depth to timeliness. Reviewers were especially critical of manuscripts that lacked literature from the technical communication discipline, which indicates that reviewers sometimes functioned as disciplinary gatekeepers.
- Reviewers often functioned as copyeditors. About 40% of the reviewers' writing-related comments dealt with lower level writing concerns.

Publication Recommendations & Decisions

- As expected, reviewers' publication recommendations differed significantly from one another. In turn, the reviewers' publication recommendations differed significantly from the editors' publication decisions (H_3).
- No significant relationship was found between the number of problems each reviewer identified and the reviewer's publication recommendation (H_6).
- Reviewers usually evaluated different aspects of manuscripts; when reviewers did evaluate the same aspects, they rarely disagreed.

Implications

The findings from this study have several implications for the technical communication field as a whole.

Implications for the Technical Communication Discipline

The study results have implications relating to the roles that peer reviewers play and the reports that reviewers write. Reviewers usually focused on higher level problems (e.g., theoretical framework, argumentation, organization, data analysis, conclusions) rather than lower level problems (e.g., grammar, mechanics, style, citations). Within the realm of higher level problems,

reviewers often gravitated to writing-related problems rather than research-related problems. As technical communicators—experts in writing—it is natural to be bothered by writing-related problems; however, in the context of peer review, technical communicators should consider the role(s) they are expected to play as reviewer. Is the reviewer expected to evaluate the writing (the presentation of the research), the research (e.g., design, methods), both, or something else? The answer may vary from journal to journal, editor to editor, manuscript to manuscript. The question remains open: Which reviewer roles are most helpful to editors (who are also experts in writing)?

The study also has implications relating to the usefulness and usability of reviewer reports. The usefulness and usability of a report depends on the perceived purpose(s) of peer review—again, something that may vary from journal to journal and editor to editor—and the audience of the report (editor or author). How might the genre conventions of the peer review report incorporate user experience (UX) considerations (e.g., usefulness and usability)? How might considerations of the UX of peer review (the reports, forms, and the process itself) affect technical communication scholarship?

Implications for Technical Communication Pedagogy

The study findings could be applied to TC editing pedagogy practices, particularly regarding the order of edits. For example, students could consider the implications of copyediting manuscripts that have not yet been accepted for publication. Discussions might cover labor inefficiencies, editorial roles, and error severity. An activity (without a performance grade) could be designed that applies Nielsen's (2000) usability testing statistics to editing: how many editors are needed to find all errors in a manuscript? Which types of problems are easiest/hardest to detect and why? What are the consequences, if any, of failing to identify and correct certain types of problems?

In other technical communication classes, instructors might focus on professional peer review as a form of persuasive communication and help students develop rhetorical strategies for writing and responding to criticism. An awareness of common points of reviewer disagreement could help students reconcile conflicting advice and negotiate the revision process.

Implications for Technical Communication Practitioners

The results from this study have potential implications for workplace review practices, particularly in respect to the selection of subject matter experts (SMEs). The study data showed a significant difference between the types of problems that reviewers identified, which means SMEs should selected with that knowledge in mind. The SMEs should be selected to minimize coverage gaps. In other words, if each SME is focused on something different, try to identify the areas that no one is looking at; otherwise, quality control may suffer.

Few studies have analyzed the content of reviewers' evaluative comments (e.g., Bakanic et al. 1989) much less examined how reviewers' comments shape editorial decisions and content development. The results of this study disrupt positivistic notions of editorial peer review as an objective evaluation standard, yet help us begin to understand how reviewer (dis)agreement is beneficial for content development and disciplinary knowledge-making. Editorial peer review is a content-shaping mechanism that helps ensure technical communication journals publish quality scholarship. Editorial peer review operates as a form of social action in which reviewers internalize the generic conventions of journal scholarship and function as pseudo-developmental editors. These findings call for changes to the way we foster disciplinary knowledge-making in the technical communication field; among other things, we need to (1) define manuscript disposition terms to help calibrate reviewers' publication recommendations and reduce

uncertainty in decision-making, (2) add microcopy to reviewer forms that explains the differences between publication recommendation options, and (3) specify reviewer roles and tasks.

References

Alam, M. , Kim, N. A., Havey, J., Rademaker, A., Ratner, D., Tregre, B., . . . Coleman, W. P. (2011). Blinded vs unblended peer review of manuscripts submitted to a dermatology journal: A randomized multi-rater study. *British Journal of Dermatology, 165*(1), 563–7. doi:10.1111/j.1365-2133.2011.10432.x

Armstrong, J. S., & Hubbard, R. (1991). Does the need for agreement among reviewers inhibit the publication of controversial findings? *Behavioral and Brain Sciences*, *14*(1), 136–137.

Bailar, J. C. (1991). Reliability, fairness, objectivity and other inappropriate goals in peer review. *Behavioral and Brain Sciences*, *14*(1), 137–138.

Bakanic, V., McPhail, C., & Simon, R. J. (1989). Mixed messages: Referees' comments on the manuscripts they review. *The Sociological Quarterly*, *30*(4), 639–654.

Ball, C. (2017). Building a scholarly multimedia publishing infrastructure. *Journal of Scholarly Publishing*, *48*(2), 99–115.

Baruch, Y., Konrad, A. M., Aguinis, H., & Starbuck, W. H. (Eds.). (2008). *Opening the black box of editorship* (pp. 239–249). New York, NY: Palgrave Macmillan.

Berkenkotter, C. (1995). The power and the perils of peer review. *Rhetoric Review, 13*(2), 245–248.

Berkenkotter, C., & Huckin, T. N. (2004). Rethinking genre from a sociocognitive perspective. In J. M. Dubinsky (Ed.), *Teaching technical communication: Critical issues for the classroom* (pp. 283–309). Boston, MA: Bedford/St. Martin's.

Blakeslee, A. M. (2001). Bridging the workplace and the academy: Teaching professional genres through classroom-workplace collaborations. *Technical Communication Quarterly*, *10*(2), 169–192.

Blakeslee, A. M. (2009). The technical communication research landscape. *Journal of Business and Technical Communication*, *23*(2), 129–171. doi:10.1177/1050651908328880

Boettger, R. K. (2014). The technical communication editing test: Three studies on this assessment type. *Technical Communication*, *61*(1), 215–231.

Boettger, R. K., & Friess, E. (2016). Academics are from Mars, practitioners are from Venus: Analyzing content alignment within technical communication forums. *Technical Communication, 63*(4), 314–364.

Boettger, R. K., & Moore, L. E. (2018). Analyzing error perception and recognition among professional communication practitioners and academics. *Business and* Professional Communication Quarterly, 81(4), 462–484. doi:10.1177/2329490618803740

Boettger, R. K., & Palmer, L. A. (2010). Quantitative content analysis: Its use in technical communication. *IEEE Transactions on Professional Communication*, *53*(4), 346–357.

Bornmann, L., Nast, I., & Daniel, H.-D. (2008). Do editors and referees look for signs of scientific misconduct when reviewing manuscripts? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics*, *77*(3), 415–432. doi:10.1007/s11192-007-1950-2

Bornmann, L., Mutz, R., & Daniel, H.-D. (2010). A reliability-generalization study of journal peer reviews; A multilevel meta-analysis of inter-rater reliability and its determinants. *PLoS ONE, 5*(12), e14331. doi:10.1371/journal.pone.0014331

Bornmann, L., Weymuth, C., & Daniel, H.-D. (2010). A content analysis of referees' comments: How do comments on manuscripts rejected by a high-impact journal and later published in either a low- or high-impact journal differ? *Scientometrics*, *83*, 493–506. doi:10.1007/s11192-009-0011-4

Brown, P., & Levinson, S. (1987). *Politeness: Some universals in language use*. Cambridge, UK: Cambridge University Press.

Bruce, R., Chauvin, A., Trinquart, L., Ravaud, P., & Boutron, I. (2016). Impact of interventions to improve the quality of peer review of biomedical journals: A systematic review and meta-analysis. *BMC Medicine*, *14*(85). doi:10.1186/s12916-016-0631-5

Brumberger, E., & Lauer, C. (2017). The evolution of technical communication: An analysis of industry job postings. *Technical Communication*, *62*(4), 224–243.

Burbules, N. C. (2014). Philosophical reflections on editing. *Educational Theory*, *64*(4), 317–331.

Bureau of Labor Statistics. (2019). Technical writers. Occupational outlook handbook. *U. S. Department of Labor*. Retrieved from https://www.bls.gov/ooh/media-and-communication/technical-writers.htm

Callaham, M. L., Wears, R. L., & Waeckerle, J. F. (1998). Effect of attendance at a training session on peer reviewer quality and performance. *Annual of Emergency Medicine*, *32*(1), 318–322.

Carliner. S. (2015). Editor's note: The madness underlying the method. *IEEE Transactions on Professional Communication*, *58*(2), 126–134. doi:10.1109/TPC.2015.2429252

Ceci, S. J., & Williams, W. M. (2011). Understanding the current causes of women's underrepresentation in science. *Proceedings of the National Academy of Sciences*, *108*(1), 3157–62.

Chrisman, J. J., Sharma, P., & Chua, J. (2017). The mind-set of editors and reviewers. *Family Business Review*, *30*(3), 211–218. doi:10.1177/0894486517708377

Cobo, E. Cortes, K., Ribera, J., Cardellach, F., Selva-O'Callagham, R., Kostov, B., . . . Vilardell, M. (2011). Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: Masked randomised trial. *BMJ*, *343*, d6783. doi:10.1136/bmj.d6783

Cobo, E., Selva-O'Callagham, R., Ribera, J. M., Cardellach, F., Dominguez, R., & Vilardell, M. (2007). Statistical reviewers improve reporting in biomedical articles: A randomized trial. *PLoS One*, *2*(3), e322. doi:10.1371/journal.pone.0000332

Coppola, N. W. (2011). Professionalization of technical communication: Zeitgeist for our age: Introduction to this special issue (part 1). *Technical Communication*, *58*(4), 277–284.

Coppola, N. W. (2012). Professionalization of technical communication: Zeitgeist for our age: Introduction to this special issue (part 2). *Technical Communication*, *59*(1), 1–7.

Coppola, N. W., & Elliot, N. (2013). Conceptualizing the technical communication body of knowledge: Context, metaphor, and direction. *Technical Communication, 60*(4), 267–278.

COPE Council. (2013). COPE ethical guidelines for peer reviewers. Committee on Publication Ethics. Retrieved from http://publicationethics.org/files/Peer%20review%20guidelines_0.pdf

COPE Council. (2017). About COPE. Committee on Publication Ethics. Retrieved from https://publicationethics.org/about

COPE Council. (2017, September). Ethical guidelines for peer reviewers. Retrieved from https://publicationethics.org/files/Ethical_Guidelines_For_Peer_Reviewers_2.pdf

Cresswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Los Angeles, CA: Sage.

Das Sinha, S., Sahni, N., & Nundy, S. (1999). Does exchanging comments of Indian and non-Indian reviewers improve the quality of manuscripts? *National Med Journal of India*, *5*(1), 210–213.

Davy, D., & Valecillos, C. (2011). Qualitative research in technical communication: A review of articles published from 2003 to 2007. In J. Conklin & G. F. Hayhoe (Eds.), *Qualitative research in technical communication* (pp. 347–375). New York, NY: Routledge.

Dayton, D., & Bernhardt, S. A. (2004). Results of a survey of ATTW members, 2003. *Technical Communication Quarterly*, *13*(1), 13–43. doi:10.1207/S15427625TCQ1301_5

De Jong, M. D. T. (2009). Editing for the workplace, the academe, and the discipline. *Technical Communication*, 56(1), 1–2.

Devitt, A. J., Bawarshi, A., & Reiff, M. J. (2003). Materiality and genre in the study of discourse communities. *College English*, *65*(5), 541–558.

Donovan, S. K. (2011). Big journals, small journals, and the two peer reviews. *Journal of Scholarly Publishing*, *42*(4), 534–538.

Eberley, S., & Warner, K. W. (1990). Fields of subfields of knowledge: Rejection rates and agreement in peer review. *American Sociologist*, *21*(3), 217–231.

Eden, D. (2008). What authors need to know to navigate the review process successfully: Understanding and managing the editor's dilemma. In Y. Baruch, A. M. Konrad, H. Aguinis, & W. H. Starbuck (Eds.), *Opening the black box of editorship* (pp. 239–249). New York, NY: Palgrave Macmillan.

Elvik, R. (1998). Are road safety evaluation studies published in peer reviewed journals more valid than similar studies not published in peer reviewed journals? *Accident Analysis & Prevention*, *30*(1), 101–118.

Ernst, E., & Resch, K. L. (1999). Reviewer bias against the unconventional? *Complementary Therapy Medicine*, *7*(1), 19–23.

Fischer, C. (2011). A value-added role for reviewers in enhancing the quality of published research. *Journal of Scholarly Publishing*, *42*(2), 226–237. doi:10.1353/scp.2011.0000

Fisher, M., Friedman, S. B., & Strauss, B. (1994). The effects of blinding on acceptance of research papers by peer review. *JAMA*, *272*(1), 143–146.

Fiske, D. W., & Fogg, L. (1990). But the reviewers are making different criticisms of my paper! Diversity and uniqueness in reviewer comments. *American Psychologist*, *45*(5), 591–598.

Ford, E. (2013). Defining and characterizing open peer review: A review of the literature. *Journal of Scholarly Publishing*, *44*(4), 311–326.

Fortanet, I. (2008). Evaluative language in peer review referee reports. *Journal of English for Academic Purposes*, *7*, 27–37.

Gale, F. G. (1998). Composition journals and the politics of knowledge-making: A conversation with journal editors. *JAC*, *18*(2), 197–211.

Gardner, A., Willey. K., Jolly, L., & Tibbits, G. (2012). Peering at the peer review process for conference submissions. *Frontiers in Education Conference Proceedings* (pp. 1–6).

Garten, J., Hoover, J., Johnson, K. M., Boghrati, R., Iskiwitch, C., & Dehghani, M. (2018). Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis. *Behavior Research Methods, 50*(1), 334–361. doi:10.3758/s13428-017-0875-9

Gebhardt, R. C., Berkenkotter, C., Arrington, P., Hesse, D., Fontaine, S. I., & Hunter, S. (1995). Symposium on peer reviewing in scholarly journals. *Rhetoric Review*, *13*(2), 237–272. doi:10.1080/07350199509359186

Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method* (4th ed.). New York, NY: Routledge.

Getchell, K., & Amicucci, A. N. (2014). Responding with the golden rule: A cross-institutional peer review experiment. *Teaching English in the Two-year College*, *42*(1), 55–64.

Giammona, B. (2004). The future of technical communication: How innovation, technology, information management, and other forces are shaping the future of the profession. *Technical Communication*, *51*(3), 349–366.

Godlee, F. (2002). Making reviewers visible: Openness, accountability, and credit. *JAMA*, *287*(21), 2762–2765.

Goldberg, P. K. (Ed.). (2014). *American Economic Review*. *American Economic Review: Papers & Proceedings*, *104*(5), 621–631. doi:10.1257/aer.104.5.621

Gonzalez, M. E. (2006). The value(s) of editorial peer review. *On the Horizon*, *14*(3), 121–129.

Goodman, S. N., Berlin, J., Fletcher, S. W., & Fletcher, R. H. (1994). Manuscript quality before and after peer review and editing at *Annals of Internal Medicine*. *Annals of Internal Medicine*, *121*(1), 11–21. doi:10.7326/0003-4819-121-1-199407010-00003

Gosden, H. (2001). "Thank you for your critical comments and helpful suggestions": Compliance and conflict in authors' replies to referees' comments in peer reviews of scientific research papers. *Ibérica*, *3*, 3–17.

Gosden, H. (2003). "Why not give us the full story?": Functions of referees' comments in peer reviews of scientific research papers. *Journal of English for Academic Purposes*, *2*, 87–101.

Grod, O. N., Lortie, C. J., & Budden, A. E. (2010). Behind the shroud: A survey of editors in ecology and evolution. *Frontiers in Ecology and the Environment*, *8*(4), 187–192.

Guthrie, J., Parker, L. D., & Dumay, J. (2015). Academic performance, publishing and peer review: Peering into the twilight zone. *Accounting, Auditing, & Accountability Journal*, *28*(1), 2–13.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London, UK: Edward Arnold.

Halvorson, K., & Rach, M. (2012). *Content strategy for the web* (2nd ed.). Berkeley, CA: New Riders.

Hayhoe, G. (1996). Editorial policy and peer review process. *Technical Communication, 43*(3), 222–224.

Hayhoe, G. (2001). Something so right: Why I'm proud of our review process. *Technical Communication, 48*(3), 261–262.

Hayhoe, G. (2008). An editor's debts. *Technical Communication, 55*(1), 7–8.

Hayhoe, G. (2010). Editing a technical journal. In A. J. Murphy (Ed.), *New Perspectives on Technical Editing* (pp. 155–180). Amityville, NY: Baywood.

Henning, T., & Bemer, T. (2016). Reconsidering power and legitimacy in technical communication: A case for enlarging the definition of technical communicator. *Journal of Technical Writing and Communication, 46*(3), 311–341. doi:10.1177/0047281616639484

Henschel, S., & Meloncon, L. (2014). Of horsemen and layered literacies: Assessment instruments for aligning technical and professional communication undergraduate curricula with professional expectations. *Programmatic Perspectives, 6*(1), 3–26. Retrieved from http://www.cptsc.org/pp/vol6-1/henschel&meloncon.pdf

Henze, B. (2004). Emergent genres in young disciplines: The case of ethnological science. *Technical Communication Quarterly, 13*(4), 393–421.

Hirschauer, S. (2010). Editorial judgments: A praxeology of "voting" in peer review. *Social Studies of Science, 40*(1), 71–103.

Hirst, A., & Altman, D. G. (2012). Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PLoS ONE, 7*(4), e35621. doi:10.1371/journal.pone.0035621

Houry, D., Green, S., & Callaham, M. (2012). Does mentoring new peer reviewers improve review quality? A randomized trial. *BMC Medical Education, 12*(1), 83.

Howard, G. (2012). Peer review as boundary work. *Journal of Scholarly Publishing, 43*(3), 322–335.

Hughes, M. A., & Hayhoe, G. F. (2008). *A research primer for technical communication: Methods, exemplars, and analyses*. New York, NY: Routledge.

Hunter, S. (1995). The case for reviewing as collaboration and response. *Rhetoric Review, 13*(2), 265–272.

ICMJE. (2017). Responsibilities in the submission and peer-review process. *International Committee of Medical Journal Editors*. Retrieved from http://www.icmje.org/recommendations/browse/roles-and-responsibilities/responsibilities-in-the-submission-and-peer-peview-process.html#three

Jacobs, J. A. (2008). The case for an activist editorial model. In Y. Baruch, A. M. Konrad, H. Aguinis, & W. H. Starbuck (Eds.), *Opening the black box of editorship* (pp. 124–133). New York, NY: Palgrave Macmillan.

Jackson, J.L., Srinivasan, M., Rea, J., Fletcher, K.E., & Kravitz, R.L. (2011). The validity of peer review in a general medicine journal. *PLoS ONE*, *6*(7), e22475. doi:10.1371/journal.pone.0022475

Jefferson, T., Alderson, P., Wager, E., & Davidoff, F. (2002). Effects of editorial peer review: A systematic review. *JAMA*, *287*(21), 2784–2786.

Jefferson, T., Wager, E., & Davidoff, F. (2002). Measuring the quality of editorial peer review. *JAMA*, *287*(21), 2786–2789.

Johnson, D. M. (1992). Compliments and politeness in peer-review texts. *Applied Linguistics*, *13*(1), 51–71.

Johnson, R. R. (2009). Trajectories, Kairos, and tulips: A personal reflection and meditation on programs in rhetoric, technical, professional, and scientific communication. *Programmatic Perspectives*, *1*(1), 45–58. Retrieved from http://www.cptsc.org/pp/vol1-1/johnson.pdf

Johnson, S. C. Lowenstein, D. H., Ferriero, D. M., Messing, R. O., Oksenberg, J. R., & Hauser, S. L. (2007). Early editorial manuscript screening versus obligate peer review: A randomized trial. *Annual of Neurology*, *61*(4), A10–2. doi:10.1002/ana.21150

Johnson-Eilola, J., & Selber, S. A. (2001). Sketching a framework for graduate education in technical communication. *Technical Communication Quarterly*, *10*(4), 403–437. doi:10.1207/s15427625tcq1004_3

Johnson-Eilola, J., & Selber, S. A. (2004). *Central works in technical communication*. New York, NY: Oxford University Press.

Kain, D. (2005). Constructing genre: A threefold typology. *Technical Communication Quarterly*, *14*(4), 375–409. doi:10.1207/s15427625tcq1404_2

Koerber, A., & McMichael, L. (2008). Qualitative sampling methods: A primer for technical communicators. *Journal of Business and Technical Communication*, *22*(4), 454–473. doi:10.1177/1050651908320362

Kravitz, R. L., Franks, P., Feldman, M. D., Gerrity, M., Byrne, C., & Tierney, W. (2010). Editorial peer reviewers' recommendations at a general medical journal: Are they reliable and do editors care? *PLoS ONE*, *5*(4), e10072. doi:10.1371/journal.pone.0010072

Krippendorff, K. (2019). *Content analysis: An introduction to its methodology* (4th ed.). Los Angeles, CA: Sage.

Kynell-Hunt, T., & Savage, G. J. (Eds.). (2003). *Power and legitimacy in technical communication: The historical and contemporary struggle for professional status* (Vol. I). Amityville, NY: Baywood.

Kynell-Hunt, T., & Savage, G. J. (Eds.). (2004). *Power and legitimacy in technical communication: The historical and contemporary struggle for professional status* (Vol. II). Amityville, NY: Baywood.

Lakoff, R. T. (1979). *Language and women's place.* New York, NY: Harper & Row.

Lakoff, R. T., & Bucholtz, M. (2004). *Language and women's place: Text and commentaries* (2nd ed.). Cary, NC: Oxford University Press. Retrieved from http://www.ebrary.com

Lam, C. (2014). Where did we come from and where are we going? Examining authorship characteristics in technical communication research. *IEEE Transactions on Professional Communication, 57*(4), 266–285. doi:10.1109/TPC.2014.2363892

Lam, C. (2016). Correspondence analysis: A statistical technique ripe for technical and professional communication researchers. *IEEE Transactions on Professional Communication, 59*(3), 299–310. doi:10.1109/TPC.2016.2583279

Lauer, C., Brumberger, E., & Beveridge, A. (2018). Hand collecting and coding versus data-driven methods in technical and professional communication research. *IEEE Transactions on Professional Communication, 61*(4), 389–408. doi:10.1109/TPC.2018.2870632

Lay, M. M. (2004). Reflections on *Technical Communication Quarterly*, 1991–2003: The manuscript review process. *Technical Communication Quarterly*, *13*(1), 109–119. doi:10.1207/S15427625TCQ1301_10

Lee, C. J., Sugimoto, C. R., Zhang, G., & Cronin, B. (2013). Bias in peer review. *Journal of the American Society for Information Science and Technology*, *64*(1), 2–17. doi:10/1002/asi.22784

Lloyd, M. E. (1990). Gender factors in reviewer recommendations for manuscript publication. *Journal of Applied Behavior Analysis*, *23*(4), 539–543. doi:10.1901/jaba.1990.23-539

Locke, J. (1700). Of words or language in general. In *An essay concerning human understanding* (4th ed), Book III, Chapter 1. Retrieved from http://web.mnstate.edu/gracyk/courses/web%20publishing/LockeBook3_2.htm

Lockyer, S. (2012). Textual analysis. In L. M. Given (Ed.), *The SAGE encyclopedia of qualitative research methods* (pp. 865–866). Thousand Oaks, CA: Sage. doi:10.4135/9781412963909

Lowe, W. (2019). Yoshikoder, version 0.6.5.0. Gnu Public License. Retrieved from http://yoshikoder.sourceforge.net/index.html

Lowry, P. B., Humpherys, S. L., Malwitz, J., & Nix, J. (2007). A scientometric study of the perceived quality of business and technical communication journals. *IEEE Transactions on Professional Communication, 50*(4), 352–378. doi:10.1109/TPC.2007.908733

Luzón, M. J. (2005). Genre analysis in technical communication. *IEEE Transactions on Professional Communication*, *48*(3), 285–295.

McNealy, M. S. (1999). *Strategies for empirical research in writing*. Needham Heights, MA: Allyn and Bacon.

Marsh, H. W., & Ball, S. (1989). The peer review process used to evaluate manuscripts submitted to academic journals. *The Journal of Experimental Education, 57*(2), 151–169. doi:10.1080/00220973.1989.10806503

McNabb, R. (2001). Making the gesture: Graduate student submissions and the expectation of journal referees. *Composition Studies*, *29*(1), 9–26.

Meruane, O. S., Vergara, C. G., & Pina-Stranger, A. (2016). What we still don't know about peer review. *Journal of Scholarly Publishing*, *47*(2), 180–212.

Miller, C. R. (1984). Genre as social action. *Quarterly Journal of Speech, 70*(2), 151–167.

Miller, C. R., & Fahnestock, J. (2013). Genres in scientific and technical rhetoric. *Poroi: An Interdisciplinary Journal of Rhetoric Analysis and Invention*, *9*(1), 12. doi:10.13008/2151-2957.1161

National Communication Association. (2012). Scholarly publishing resources: CARD call, February 2012. Retrieved from https://dev.natcom.org/uploadedFiles/Teaching_and_Learning/CARD_Calls/PDF-CARDCalls-Publishing_Resources_February_2012.pdf

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). Los Angeles, CA: Sage.

Nielsen, J. (2000). Why you only need to test with 5 users. Nielsen Norman Group. Retrieved from https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

Nielsen, J. (2012). Usability 101: Introduction to usability. Nielsen Norman Group. Retrieved from https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Norton, S. (2009). *Developmental editing: A handbook for freelancers, authors, and publishers.* Chicago, IL: The University of Chicago Press.

Olson, C. M., Rennie, D., Cook, D, Dickersin, K., Flanagin, A., Hogan, J. W., . . . Pace, B. (2002). Publication bias in editorial decision making. *JAMA, 287*(21), 2825–2828.

Perreault, W. D., & Leigh, L. E. (1989). Reliability of nominal data based on qualitative judgments. *Journal of Marketing Research*, *26*(2), 135–148.

PhD on Track. (2016). Submitting articles: The peer-review process. Retrieved from http://www.phdontrack.net/share-and-publish/submitting-articles/

Pierie, J.-P., Walvoort, H. C., & Overbeke, A. J. (1996). Readers' evaluation of effect of peer review and editing quality of articles in the Nederlands Tijdschrift voor Geneeskunde. *The Lancet, 348*(9040), 1480–1483. doi:10.1016/S0140-6736(96)05016-7

Popper, K. (2002). *The logic of scientific discovery*. London, UK: Routledge.

Price, S., & Flach, P. A. (2017). Computational support for academic peer review: A perspective from artificial intelligence. *Communications of the ACM, 60*(3), 70–79. doi:10.1145/2979672

Rothwell, P. M., & Martyn, C. N. (2000). Reproducibility of peer review in clinical neuroscience: Is agreement between reviewers any greater than would be expected by chance alone? *Brain, 123*, 1964–1969.

Rude, C. D. (2009). Mapping the research questions in technical communication. *Journal of Business and Technical Communication, 23*(2), 174–215. doi:10.1177/1050651908329562

Rude, C. D., & Eaton, A. (2011). *Technical editing* (5th ed.). Boston, MA: Longman.

Sage Journals. (n.d.). Reviewer's guide. Retrieved from https://journals.sagepub.com/pb-assets/PDF/Author%20Gateway%20PDF/How%20to%20become%20a%20reviewer_NEW.pdf

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Los Angeles, CA: Sage.

Savage, G. J. (1999). The process and prospects for professionalizing technical communication. *Journal of Technical Writing and Communication, 29*(4), 355–381.

Sawin, P. E. (1999). Gender, context, and the narrative construction of identity: Rethinking models of "women's narrative." In M. Bucholtz, A. C. Liang, & L. A. Sutton (Eds.), *Reinventing identities: The gendered self in discourse* (pp. 241–258). Cary, NC: Oxford University Press. Retrieved from http://www.ebrary.com

Schroter, S., Black, N., Evans, S., Godlee, F., Osorio, L., & Smith, R. (2008). What errors do peer reviewers detect, and does training improve their ability to detect them? *Journal of the Royal Society of Medicine, 101*, 507–514. doi:10.1258/jrsm.2008.080062

Schuster, M. L. (2015). My career and the "rhetoric of" technical writing and communication. *Journal of Technical Writing and Communication, 45*(4), 381–391. doi:10.1177/0047281615585754

Selfe, C. L., & Hawisher, G. E. (2012). Methodologies of peer editorial review: Changing practices. *College Composition and Communication, 63*(4), 672–698.

Sheffield, J. P., Sparks, S. C., & Ianetta, M. (2014). Revaluing the work of the editor. *College English*, *77*(2), 146–164.

Smith, E. O. (2000). Strength in the technical communication journals and diversity in the serials cited. *Journal of Business and Technical Communication*, *14*(2), 131–184.

Spinuzzi, C., & Zachry, M. (2000). Genre ecologies: An open-system approach to understanding and constructing documentation. *ACM Journal of Computer Documentation*, *24*(3), 169–181.

Spool, J. M. (2019). Zooming in and out of UX design resolutions. UIE. Retrieved from https://articles.uie.com/zooming-in-and-out-of-ux-design-resolutions/

Sposato, L. A., Ovbiagele, B., Johnston, S. C., Fisher, M., & Saposnik, G. (2014). A peek behind the curtain: Peer review and editorial decision making at *Stroke*. *American Neurological Association, 76,* 151–158. doi:10.1002/ana.24218

Stemler, S. (2001). An overview of content analysis. *Practical Assessment, Research & Evaluatio*n, 7(17). Retrieved from http://pareonline.net/getvn.asp?v=7&n-17

Swales, J. M. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.

Swales, J. M. (1996). Occluded genres in the academy: The case of the submission letter. In E. Ventola & A. Mauranen (Eds.), *Academic writing: Intercultural and textual issues* (pp. 45–58). Amsterdam, The Netherlands: John Benjamins Publishing.

Swales, J. M. (2000). Languages for specific purposes. *Annual Review of Applied Linguistics*, *20*(1), 59–76.

Swales, J. M. (2004). *Research genres: Exploration and applications.* New York, NY: Cambridge University Press.

Tannen, D. (1994). *Gender and discourse*. Cary, NC: Oxford University Press. Retrieved from http://www.ebrary.com

Tannen, D. (2002). Agonism in academic discourse. *Journal of Pragmatics, 34*(10/11), 1651–1669. doi:10.1016/S0378-2166(02)00079-6

Taylor & Francis. (2019). A guide to becoming a peer reviewer. Retrieved from https://editorresources.taylorandfrancis.com/reviewer-guidelines/

TCBOK. (n.d.). Definition of technical communication. Society for Technical Communication. Retrieved from https://www.tcbok.org/wiki/about-tc/definition-technical-communication/

Thayer, A., Evans, M., McBride, A., Queen, M., & Spyridakis, J. (2007). Content analysis as a best practice in technical communication research. *Journal of Technical Writing and Communication, 37*(3), 267–279. doi:10.2190/TW.37.3.c

Thompson, I. (2001). Collaboration in technical communication: A qualitative content analysis of journal articles, 1990–1999. *IEEE Transactions on Professional Communication*, *44*(3), 161–173.

Thralls, C., & Blyler, N. R. (2004). The social perspective and professional communication: Diversity and directions in research. In J. Johnson-Eilola & S. A. Selber (Eds.), *Central works in technical communication* (pp. 124–145). New York, NY: Oxford University Press.

Van Buren, R., & Buehler, M. F. (1980). *The levels of edit* (2nd ed). JPL 80-1. Pasadena, CA: Jet Propulsion Laboratory, California Institute of Technology.

Van Rooyen, S., Black, N., & Godlee, F. (1999). Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. *Journal of Clinical Epidemiology*, *52*(7), 625–629.

Vardi, M. Y. (2017). Ten years at the helm of *Communications of the ACM*. *Communications of the ACM*, *60*(6), 5. doi:10.1145/3090801

Vinther, S., Nielson, O. H., Rosenberg, J., Keiding, N., & Schroeder, T. V. (2012). Same review quality in open versus blinded peer review in "Ugeskrift for Laeger." *Danish Medical Journal*, *59*(8), A4479.

Walsh, R., Rooney, M., Appleby, L., & Wilkinson, G. (2000). Open peer review. *British Journal of Psychiatry*, *176*(1), 47–51.

Yoshikoder. (2015). Reports. Comparing documents. Retrieved from https://github.com/conjugateprior/yoshikoder/wiki/Reports

Zuckerman, H. & Merton, R. K. (1971). Patterns of evaluation in science: Institutionalism, structure, and functions of the referee system. *Minerva*, *9*(1), 66–100.

Appendixes

Appendix A: IRB Approval Letters

Appendix B: Glossary

Appendix C: Codebook & Coding Form

Appendix D: CATA Custom Dictionary

Appendix E: Stop Words

Appendix A: IRB Approval Letters

UMCIRB 17-001261: Initial Approval

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board Office
4N-70 Brody Medical Sciences Building· Mail Stop 682
600 Moye Boulevard · Greenville, NC 27834
Office 252-744-2914 · Fax 252-744-2284 · www.ecu.edu/ORIC/irb

Notification of Initial Approval: Expedited

From: Social/Behavioral IRB
To: Suzan Flanagan
CC:
Michael Albers
Date: 7/11/2017
Re: UMCIRB 17-001261
Technical Communication Scholarship

I am pleased to inform you that your Expedited Application was approved. Approval of the study and any consent form(s) is for the period of 7/11/2017 to 7/10/2018. The research study is eligible for review under expedited category #5. The Chairperson (or designee) deemed this study no more than minimal risk.

Changes to this approved research may not be initiated without UMCIRB review except when necessary to eliminate an apparent immediate hazard to the participant. All unanticipated problems involving risks to participants and others must be promptly reported to the UMCIRB. The investigator must submit a continuing review/closure application to the UMCIRB prior to the date of study expiration. The Investigator must adhere to all reporting requirements for this study.

Approved consent documents with the IRB approval date stamped on the document should be used to consent participants (consent documents with the IRB approval date stamp are found under the Documents tab in the study workspace).

The approval includes the following items:

Name	Description
Flanagan-SIGDOC-data-permission-request-letter-5-26-2017-IRB.docx	Consent Forms
Flanagan-SIGDOC-research-proposal-5-26-17-IRB.docx	Study Protocol or Grant Application
permission.pdf	Dataset Use Approval/Permission
permission.pdf	Consent Forms

The Chairperson (or designee) does not have a potential for conflict of interest on this study.

UMCIRB 17-001261: 2018 Continuing Review

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board
4N-64 Brody Medical Sciences Building· Mail Stop 682
600 Moye Boulevard · Greenville, NC 27834
Office 252-744-2914 · Fax 252-744-2284 · www.ecu.edu/ORIC/irb

Continuing Review Approved

ID: CR00006998

Title: 2018 Review for UMCIRB 17-001261

Technical Communication Scholarship

Description: Your continuing review has been approved as of 6/24/2018. To navigate to the project workspace, click on the above ID.

UMCIRB 17-001261: 2019 Continuing Review

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board
4N-64 Brody Medical Sciences Building· Mail Stop 682
600 Moye Boulevard · Greenville, NC 27834
Office 252-744-2914 · Fax 252-744-2284 · www.ecu.edu/ORIC/irb

Continuing Review Approved

ID:	CR00007802
Title:	2019 Review for UMCIRB 17-001261 Technical Communication Scholarship
Description:	Your continuing review has been approved as of 6/7/2019. To navigate to the project workspace, click on the above ID.

UMCIRB 17-002615: Initial Approval

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board
4N-64 Brody Medical Sciences Building· Mail Stop 682
600 Moye Boulevard · Greenville, NC 27834
Office 252-744-2914 · Fax 252-744-2284 · www.ecu.edu/ORIC/irb

Notification of Initial Approval: Expedited

From: Social/Behavioral IRB
To: Suzan Flanagan
CC:
Michael Albers
Date: 2/20/2018
Re: UMCIRB 17-002615
Editorial Peer Review

I am pleased to inform you that your Expedited Application was approved. Approval of the study and any consent form(s) is for the period of 2/19/2018 to 2/18/2019. The research study is eligible for review under expedited category #5,6, 7. The Chairperson (or designee) deemed this study no more than minimal risk.

Changes to this approved research may not be initiated without UMCIRB review except when necessary to eliminate an apparent immediate hazard to the participant. All unanticipated problems involving risks to participants and others must be promptly reported to the UMCIRB. The investigator must submit a continuing review/closure application to the UMCIRB prior to the date of study expiration. The Investigator must adhere to all reporting requirements for this study.

Approved consent documents with the IRB approval date stamped on the document should be used to consent participants (consent documents with the IRB approval date stamp are found under the Documents tab in the study workspace).

The approval includes the following items:

Name	Description
[redacted]	Dataset Use Approval/Permission
Flanagan-dissertation-prospectus-short-version.docx	Study Protocol or Grant Application
[redacted]	Additional Items
Interview-Informed-Consent-Form-Feb-6-2018.docx.doc	Consent Forms
IRB Email Script.docx	Recruitment Documents/Scripts
IRB Interview Questions.docx	Interview/Focus Group Scripts/Questions

The Chairperson (or designee) does not have a potential for conflict of interest on this study.

UMCIRB 17-002615: 2019 Continuing Review

EAST CAROLINA UNIVERSITY
University & Medical Center Institutional Review Board
4N-64 Brody Medical Sciences Building· Mail Stop 682
600 Moye Boulevard · Greenville, NC 27834
Office 252-744-2914 · Fax 252-744-2284 · www.ecu.edu/ORIC/irb

Continuing Review Approved

ID: CR00007493

Title: 2019 Review for UMCIRB 17-002615

Editorial Peer Review

Description: Your continuing review has been approved as of 1/14/2019. To navigate to the project workspace, click on the above ID.

Appendix B: Glossary

anonymous	Shorthand for double-blind peer review. The author's and the reviewers' identities are unknown to one another during the review process; upon publication of a manuscript, the author's identity is consequently revealed to the reviewers.
comparative content analysis	A comparison of the results of two quantitative content analyses (e.g., a comparison of codes associated with Reviewer 1 with the codes associated with Reviewer 2).
content	May include text, data, visuals, videos, multimedia, or other knowledge-related artifacts.
content analysis	"A research technique for making replicable and valid inferences from data to their context" (Krippendorff, 2019, p. 21).
content development	The often iterative changes made to a manuscript after the author's initial submission to a peer reviewed journal; these changes include the author's revisions and the subsequent editing of accepted manuscripts—the changes may involve the text, data, visuals, multimedia, or other knowledge-related artifacts that comprise the manuscript.
content strategy	Rhetorically grounded, planned communication practices that recursively shape information into usable knowledge-based products for various uses, audiences, and media.
copyediting	A type of editing performed after a manuscript is completed that involves minor corrections to grammar, mechanics, and style.
developmental editing	A type of editing that involves major restructuring or changes to content (Norton, 2009).
double-blind peer review	A form of peer review in which neither the authors nor the reviewers know the identity of the other party.
editorial peer review	The International Committee of Medical Journal Editors defines peer review as "the critical assessment of manuscripts submitted to journals by experts who are not part of the editorial staff" (ICMJE, 2017).
editorial strategy	An element of a content strategy that encompasses audience needs, content priorities, roles, tasks, organization, journal policies, standards, publication goals, etc.

genre	A type of text characterized by its form and the actions performed in recurring situations (Miller, 1984).
initial manuscript	An article submitted to the journal for the first time.
interrater reliability (IRR)	Cicchetti defines interrater reliability as "the extent to which two or more independent reviews of the same scientific document agree" (p. 120, as cited in Bornmann, Mutz, & Daniel, 2010).
mean priority score	A scale is used to calculate perfect agreement and perfect disagreement between reviewers' publication recommendations.
occluded genre	A genre unavailable for public scrutiny (e.g., peer review reports).
peer review	The International Committee of Medical Journal Editors defines peer review as "the critical assessment of manuscripts submitted to journals by experts who are not part of the editorial staff" (ICMJE, 2017). (I use *peer review* as shorthand for editorial peer review.)
publication-process artifacts	Records that were created as part of the publication process (e.g., authors' manuscripts, cover letters, and response letters; journals' guidelines for authors and reviewers, style guides, scoring rubrics, reviewers' reports, editors' decision letters, and published articles).
reliability	Hughes & Hayhoe (2008) define reliability as "the likelihood that the results would be the same if the study were repeated" (p. 60).
reviewer agreement	Each reviewer evaluated a specific manuscript and, in his or her report to the editor, the reviewer recommended the same publication outcome as the other reviewer(s).
revised manuscript	An article that has been peer reviewed previously.
RQI	A review quality instrument that measures "the extent to which a peer reviewer has considered key aspects of a manuscript" (van Rooyen, Black, & Godlee, 1999, p. 628).
structural analysis	A research method used to describe genres and to identify document design elements.
technical communication	Communication related to technology, technical subject matters (e.g., science, engineering), and technical procedures.
textual analysis	A research method that may include discourse analysis, narrative analysis, genre analysis, structural analysis, or content analysis.

validity	In quantitative studies, validity requires "measur[ing] the concept you wanted to study" in a "test environment" that reflects "the real world" (Hughes & Hayhoe, 2008, p. 59), whereas, in qualitative studies, validity requires "trustworthiness, authenticity, and credibility" (Cresswell, 2014, p. 201).

Appendix C: Codebook & Coding Form

CODEBOOK

Editorial Peer Review

Background Information

The principal investigator (PI) for this research study is Suzan Flanagan. The study has been approved by East Carolina University's IRB. The PI and research assistants were required to sign nondisclosure agreements prior to accessing the study data. The study data must be kept confidential and stored securely.

Each research assistant (RA) has been assigned a coder identification number (Coder ID). Your Coder ID is _____. Contact the PI with any questions about the research study or coding procedures.

This phase of the research study involves the analysis of reviewer reports and manuscript evaluation instruments (i.e., reviewer's guidelines, rubrics, and/or guiding questions). Each reviewer report has been assigned a manuscript identification number (MS ID), which can be found in the data file's page headers. The page headers also include estimates of the reviewer report word count.

Coding Instructions

1. Code each Assigned Reviewer Report independently of the other research assistant(s).
    a. Code reports with V0 and AU file designations first (J#-####-**V0**-R#-**AU**).
    b. Code reports with R1 file designations before reports with R2 file designations (J#-####-V0-**R#**-AU).
        i. For each R1 and R2 file pair (e.g., J#-####-V#-**R1**-AU and J#-####-V#-**R2**-AU), identify the number and types of problems that both reviewers agreed on (reviewer agreements).
        ii. For each R1 and R2 file pair, identify the number and types of contradictory comments (reviewer contradictions).
    c. After all the reports with VO and AU file designations have been coded, follow the same procedures to code the reports with VO and ED file designations (J#-####-**V0**-R#-**ED**).
    d. After all the reports with VO and ED file designations have been coded, proceed to the V1 and AU files and continue in the same manner until all files have been coded.
2. Use a separate Coding Form for each reviewer report coded.
3. Enter your Coder ID on each form.
4. Enter the MS ID on the coding form.
5. Check the Word Count using Microsoft Word's "word count" feature and enter the number on the coding form. If the number differs significantly from the estimated count, notify the principal investigator (PI).
    a. If the reviewer report Word Count equals zero, the coding form is complete. Save the file and submit it to the PI.

b. If the reviewer report Word Count is greater than zero, code the reviewer report using the Code Definitions and Coding Form provided. Each unit of text must be coded in one category only; select an appropriate unit of analysis to avoid double coding text.

6. Complete each section of the coding form unless instructed to skip a section.
7. Contact the PI after every ten forms completed so that the intercoder reliability can be calculated.

RESEARCH STUDY TERMINOLOGY

Assigned Reviewer Reports: A list of the documents you are to code. The documents are listed by manuscript identification number and arranged by journal, audience (author or editor), manuscript version (initial submission, first revision, etc.).

Coder Identification Number (Coder ID): Two-digit number assigned by the principal investigator to each research assistant.

Intercoder Reliability: A measurement of agreement between the research assistants in their analyses of the documents.

Manuscript Evaluation Instrument: A tool for comparing various aspects of the reviewer reports (i.e., content/topics, structure/format/order, and language/wording) to the journals' reviewer guidelines, rubric, or guiding questions.

- **Content/topic** refers to what the reviewers are saying about the manuscript in relation to the reviewer guidelines, etc.
- **Structure/format/order** refers to how reviewers present and organize their evaluative comments in relation to the reviewer guidelines, etc. For example, is a guiding question or rubric prompt followed by the reviewer's response to the question or prompt?
- **Language/wording** refers to how reviewers express their evaluative comments in relation to the reviewer guidelines, etc. For example, is language used verbatim or nearly verbatim from the reviewer guidelines, rubric, or guiding questions?

Manuscript Identification Number (MS ID): an alphanumeric string assigned by the principal investigator to indicate various details about the manuscript and its associated publication process artifacts (e.g., reviewer reports). In this study, the MS IDs take one of two forms:

J#-####-V#-R#-AU
J#-####-V#-R#-ED

J#	denotes journal (randomly assigned consecutive number)
####	denotes manuscript number (randomly assigned number between 1,000 and 5,000)
V#	denotes manuscript version, where V0 = initial submission, V1 = first revision, etc.
R#	denotes reviewer number (as assigned by the corresponding journal)
AU	denotes author as the primary review audience
ED	denotes editor as the primary review audience

Number of Manuscript Problems: The quantity of unique problems that reviewers mention in their reviews. For example, if the author misspelled the same word multiple times, count the mistake as one problem; if the author misspelled five different words, count the misspellings as five problems.

Report Comparison: A tool for identifying points of reviewer agreement and disagreement.

Reviewer Agreements: Aspects of a manuscript that both Reviewer 1 and Reviewer 2 discussed in the same or similar way. For example, Reviewer 1 said the literature review lacked current scholarship and Reviewer 2 said the literature review relied on works from 30 years ago; in other words, both reviewers have identified an in-common problem with the literature review—that is, they agree that the author needs to update the literature review.

Reviewer Contradictions: Aspects of a manuscript that both Reviewer 1 and Reviewer 2 discussed in opposite ways. For example, Reviewer 1 said the method was appropriate and Reviewer 2 said the method was inappropriate.

Reviewer's Evaluative Comments: The reviewer's remarks and assessment of various aspects of the manuscript, such as relevance of contribution, writing/presentation, design/conception, methods/statistics, discussion of results, reference to the literature and documentation, theory, and ethics. The remarks may also include the reviewer's publication recommendation.

Reviewer Evaluations: A tool designed to identify and quantify the manuscript elements reviewers have discussed.

Reviewer Recommendations: Indicates whether the reviewer stated their publication recommendation within the evaluative comment section of the review and whether the reviewer seemed confident in their recommendation.

Review Structure: Refers to the ways in which the reviewer presents and organizes the reviewer report and the absence or presence of various genre conventions.

Review Quality Instrument: A tool designed to assess the reviewer's report and the depth and breadth of their evaluation of the manuscript.

Unit of Analysis: In this study, the unit of analysis ranges from a word to several sentences.

Word Count: Number of words in each reviewer report section as calculated by Microsoft Word's "word count" feature. (Note: These word counts may vary slightly from the actual reviews because the text has been redacted by the respective journals' staff members, and the researcher made additional redactions.)

CODING FORM—PART 1

Coder ID:

MS ID:

Word Count:

{SKIP LOGIC: If WORD COUNT = 0, then stop here [save and submit form], else, continue to next section.}

Reviewer's Evaluative Comments

Elements discussed in review

CODE:
1 = relevance of contribution
2 = writing/presentation (higher order)
3 = writing/presentation (lower order)
4 = design/conception
5 = methods/statistics
6 = discussion of results
7 = reference to the literature and documentation
8 = theory
9 = ethics
10 = other—briefly describe: _____

Number and types of elements discussed in review

Indicate the types & quantities of each element (positive, negative, or neutral) discussed in the review									
contribution	writing: higher	writing: lower	design	methods	results	lit	theory	ethics	other
1	2	3	4	5	6	7	8	9	10
Indicate the types & quantities of each unique problem (negative element) discussed in the review									
contribution	writing: higher	writing: lower	design	methods	results	lit	theory	ethics	other
1	2	3	4	5	6	7	8	9	10

{DISPLAY LOGIC: If R = 2 in MS ID (J#-####-V0-**R#**-AU), then answer Report Comparison questions; else, save and submit form and proceed to Coding Form—Part 2.}

Report Comparison

Indicate the types & quantities of problems (negative elements) both reviewers discussed (i.e., problems reviewers agreed upon)									
contribution	writing: higher	writing: lower	design	methods	results	lit	theory	ethics	other
1	2	3	4	5	6	7	8	9	10

Indicate the types & quantities of contradictory comments about the same elements (i.e., reviewers' points of disagreement)									
contribution	writing: higher	writing: lower	design	methods	results	lit	theory	ethics	other
1	2	3	4	5	6	7	8	9	10

CODING FORM—PART 2

Reviewer Recommendation

Publication Recommendation

CODE: 1 = the reviewer does not state their publication recommendation
2 = the reviewer states their publication recommendation

{SKIP LOGIC: If Publication Recommendation code = 1, then skip to Review Structure}

Reviewer Indecision

CODE: 1 = the reviewer states their publication recommendation without expressing uncertainty or indecision about the recommendation
2 = the reviewer expresses uncertainty or indecision about their publication recommendation (e.g., minor revision vs. major revision; major revision vs. reject; minor revision vs. accept)

Review Structure

Summarizing judgment [select all that apply]

CODE: 1 = as opening remarks only
2 = as closing remarks only
3 = as opening and closing remarks
4 = in a file attachment
5 = in combination with a file attachment
6 = none
7 = other—briefly describe: _____

Outline of article [select one]

CODE: 1 = review includes an outline of article
2 = review does not include an outline of article
3 = other—briefly describe: _____

Conclusion [select one]

CODE: 1 = review includes conclusion paragraph(s)
2 = review does not include conclusion paragraph(s)
3 = other—briefly describe: _____

Comment presentation [select all that apply]

CODE: 1 = numbered point-by-point comments
2 = unnumbered point-by-point comments
3 = combination of numbered/unnumbered point-by-point comments
4 = page-by-page comments (location-based comment)
5 = section-by-section comments (location-based comment)
6 = combination of location-based comments
7 = combination of point-by-point and location-based comments
8 = other—briefly describe: _____

Feedback approaches [select all that apply]

CODE: 1 = praise/criticism pairs (e.g., A is effective; however, B needs work.)
2 = hedged praise (e.g., A *seems* to support your argument.)
3 = hedged criticism (e.g., B is confusing *but* that could be my reading of the text.)
4 = praise attributed to journal criteria (e.g., C meets the journal's guidelines.)
5 = criticism attributed to journal criteria (e.g., C does not follow the guidelines.)
6 = direct praise (e.g., A is useful to the field.)
7 = direct criticism (e.g., B is irrelevant to your argument.)
8 = other—briefly describe: _____

Manuscript Evaluation Instrument

The **content/topics** of the review aligns with the journal's manuscript evaluation instrument.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1	2	3	4	5	6

The **structure/format/order** of the review aligns with the journal's manuscript evaluation instrument.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1	2	3	4	5	6

The **language/wording** of the review aligns with the journal's manuscript evaluation instrument.

Strongly Disagree	Disagree	Somewhat Disagree	Somewhat Agree	Agree	Strongly Agree
1	2	3	4	5	6

Review Quality Instrument

To what extent did the reviewer discuss the importance of the research question?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer discuss the originality of the paper?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer clearly identify the strengths of the study design/conception?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer clearly identify the weaknesses of the study design/conception?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer clearly identify the strengths of the methods/statistics?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer clearly identify the weaknesses of the methods/statistics?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer comment on the author's interpretations of the results?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer make specific useful comments on the writing/presentation (higher order)?

None At All				A Great Deal
1	2	3	4	5

To what extent did the reviewer make specific useful comments on the writing/presentation (lower order)?

None At All				A Great Deal
1	2	3	4	5

Were the reviewer's comments constructive?

Not At All Constructive				Very Constructive
1	2	3	4	5

Did the reviewer supply appropriate evidence using examples from the paper to substantiate their comments?

No Comments Substantiated		Some Comments Substantiated		All Comments Substantiated
1	2	3	4	5

How would you rate the quality of this review overall?

Poor				Excellent
1	2	3	4	5

ASSIGNED REVIEWER REPORTS: JOURNAL 1

Initial Submissions Journal 1 [30 mss]			
Author		**Editor**	
J1-1171-V0-R1-AU	J1-1171-V0-R2-AU	J1-1171-V0-R1-ED	J1-1171-V0-R2-ED
J1-1235-V0-R1-AU	J1-1235-V0-R2-AU	J1-1235-V0-R1-ED	J1-1235-V0-R2-ED
J1-1481-V0-R1-AU	J1-1481-V0-R2-AU	J1-1481-V0-R1-ED	J1-1481-V0-R2-ED
J1-1818-V0-R1-AU	J1-1818-V0-R2-AU	J1-1818-V0-R1-ED	J1-1818-V0-R2-ED
J1-1884-V0-R1-AU	J1-1884-V0-R2-AU	J1-1884-V0-R1-ED	J1-1884-V0-R2-ED
J1-1959-V0-R1-AU	J1-1959-V0-R2-AU	J1-1959-V0-R1-ED	J1-1959-V0-R2-ED
J1-2023-V0-R1-AU	J1-2023-V0-R2-AU	J1-2023-V0-R1-ED	J1-2023-V0-R2-ED
J1-2063-V0-R1-AU	J1-2063-V0-R2-AU	J1-2063-V0-R1-ED	J1-2063-V0-R2-ED
J1-2190-V0-R1-AU	J1-2190-V0-R2-AU	J1-2190-V0-R1-ED	J1-2190-V0-R2-ED
J1-2213-V0-R1-AU	J1-2213-V0-R2-AU	J1-2213-V0-R1-ED	J1-2213-V0-R2-ED
J1-2348-V0-R1-AU	J1-2348-V0-R2-AU	J1-2348-V0-R1-ED	J1-2348-V0-R2-ED
J1-2716-V0-R1-AU	J1-2716-V0-R2-AU	J1-2716-V0-R1-ED	J1-2716-V0-R2-ED
J1-2717-V0-R1-AU	J1-2717-V0-R2-AU	J1-2717-V0-R1-ED	J1-2717-V0-R2-ED
J1-2779-V0-R1-AU	J1-2779-V0-R2-AU	J1-2779-V0-R1-ED	J1-2779-V0-R2-ED
J1-2975-V0-R1-AU	J1-2975-V0-R2-AU	J1-2975-V0-R1-AU	J1-2975-V0-R2-ED
J1-3011-V0-R1-AU	J1-3011-V0-R2-AU	J1-3011-V0-R1-ED	J1-3011-V0-R2-ED
J1-3207-V0-R1-AU	J1-3207-V0-R2-AU	J1-3207-V0-R1-ED	J1-3207-V0-R2-ED
J1-3303-V0-R1-AU	J1-3303-V0-R2-AU	J1-3303-V0-R1-ED	J1-3303-V0-R2-ED
J1-3357-V0-R1-AU	J1-3357-V0-R2-AU	J1-3357-V0-R1-ED	J1-3357-V0-R2-ED
J1-3382-V0-R1-AU	J1-3382-V0-R2-AU	J1-3382-V0-R1-ED	J1-3382-V0-R2-ED
J1-3567-V0-R1-AU	J1-3567-V0-R2-AU	J1-3567-V0-R1-ED	J1-3567-V0-R2-ED
J1-3745-V0-R1-AU	J1-3745-V0-R2-AU	J1-3745-V0-R1-ED	J1-3745-V0-R2-ED
J1-4068-V0-R1-AU	J1-4068-V0-R2-AU	J1-4068-V0-R1-ED	J1-4068-V0-R2-ED
J1-4086-V0-R1-AU	J1-4086-V0-R2-AU	J1-4086-V0-R1-ED	J1-4086-V0-R2-ED
J1-4233-V0-R1-AU	J1-4233-V0-R2-AU	J1-4233-V0-R1-ED	J1-4233-V0-R2-ED
J1-4308-V0-R1-AU	J1-4308-V0-R2-AU	J1-4308-V0-R1-ED	J1-4308-V0-R2-ED
J1-4430-V0-R1-AU	J1-4430-V0-R2-AU	J1-4430-V0-R1-ED	J1-4430-V0-R2-ED
J1-4475-V0-R1-AU	J1-4475-V0-R2-AU	J1-4475-V0-R1-ED	J1-4475-V0-R2-ED
J1-4910-V0-R1-AU	J1-4910-V0-R2-AU	J1-4910-V0-R1-ED	J1-4910-V0-R2-ED
J1-4995-V0-R1-AU	J1-4995-V0-R2-AU	J1-4995-V0-R1-ED	J1-4995-V0-R2-ED

ASSIGNED REVIEWER REPORTS: JOURNAL 2

Initial Submissions Journal 2 [31 mss]			
Author		**Editor**	
J2-1040-V0-R1-AU	J2-1040-V0-R2-AU	J2-1040-V0-R1-ED	J2-1040-V0-R2-ED
J2-1228-V0-R1-AU	J2-1228-V0-R2-AU	J2-1228-V0-R1-ED	J2-1228-V0-R2-ED
J2-1234-V0-R1-AU	J2-1234-V0-R2-AU	J2-1234-V0-R1-ED	J2-1234-V0-R2-ED
J2-1316-V0-R1-AU	J2-1316-V0-R2-AU	J2-1316-V0-R1-ED	J2-1316-V0-R2-ED
J2-1979-V0-R1-AU	J2-1979-V0-R2-AU	J2-1979-V0-R1-ED	J2-1979-V0-R2-ED
J2-2103-V0-R1-AU	J2-2103-V0-R2-AU	J2-2103-V0-R1-ED	J2-2103-V0-R2-ED
J2-2416-V0-R1-AU	J2-2416-V0-R2-AU	J2-2416-V0-R1-ED	J2-2416-V0-R2-ED
J2-2432-V0-R1-AU	J2-2432-V0-R2-AU	J2-2432-V0-R1-ED	J2-2432-V0-R2-ED
J2-2480-V0-R1-AU	J2-2480-V0-R2-AU	J2-2480-V0-R1-ED	J2-2480-V0-R2-ED
J2-2605-V0-R1-AU	J2-2605-V0-R2-AU	J2-2605-V0-R1-ED	J2-2605-V0-R2-ED
J2-2732-V0-R1-AU	J2-2732-V0-R2-AU	J2-2732-V0-R1-ED	J2-2732-V0-R2-ED
J2-2739-V0-R1-AU	J2-2739-V0-R2-AU	J2-2739-V0-R1-ED	J2-2739-V0-R2-ED
J2-2861-V0-R1-AU	J2-2861-V0-R2-AU	J2-2861-V0-R1-ED	J2-2861-V0-R2-ED
J2-3059-V0-R1-AU	J2-3059-V0-R2-AU	J2-3059-V0-R1-ED	J2-3059-V0-R2-ED
J2-3238-V0-R1-AU	J2-3238-V0-R2-AU	J2-3238-V0-R1-ED	J2-3238-V0-R2-ED
J2-3509-V0-R1-AU	J2-3509-V0-R2-AU	J2-3509-V0-R1-ED	J2-3509-V0-R2-ED
J2-3565-V0-R1-AU	J2-3565-V0-R2-AU	J2-3565-V0-R1-ED	J2-3565-V0-R2-ED
J2-3575-V0-R1-AU	J2-3575-V0-R2-AU	J2-3575-V0-R1-ED	J2-3575-V0-R2-ED
J2-3668-V0-R1-AU	J2-3668-V0-R2-AU	J2-3668-V0-R1-ED	J2-3668-V0-R2-ED
J2-3833-V0-R1-AU	J2-3833-V0-R2-AU	J2-3833-V0-R1-ED	J2-3833-V0-R2-ED
J2-3928-V0-R1-AU	J2-3928-V0-R2-AU	J2-3928-V0-R1-ED	J2-3928-V0-R2-ED
J2-4006-V0-R1-AU	J2-4006-V0-R2-AU	J2-4006-V0-R1-ED	J2-4006-V0-R2-ED
J2-4120-V0-R1-AU	J2-4120-V0-R2-AU	J2-4120-V0-R1-ED	J2-4120-V0-R2-ED
J2-4355-V0-R1-AU	J2-4355-V0-R2-AU	J2-4355-V0-R1-ED	J2-4355-V0-R2-ED
J2-4522-V0-R1-AU	J2-4522-V0-R2-AU	J2-4522-V0-R1-ED	J2-4522-V0-R2-ED
J2-4531-V0-R1-AU	J2-4531-V0-R2-AU	J2-4531-V0-R1-ED	J2-4531-V0-R2-ED
J2-4549-V0-R1-AU	J2-4549-V0-R2-AU	J2-4549-V0-R1-ED	J2-4549-V0-R2-ED
J2-4657-V0-R1-AU	J2-4657-V0-R2-AU	J2-4657-V0-R1-ED	J2-4657-V0-R2-ED
J2-4715-V0-R1-AU	J2-4715-V0-R2-AU	J2-4715-V0-R1-ED	J2-4715-V0-R2-ED
J2-4731-V0-R1-AU	J2-4731-V0-R2-AU	J2-4731-V0-R1-ED	J2-4731-V0-R2-ED
J2-4964-V0-R1-AU	J2-4964-V0-R2-AU	J2-4964-V0-R1-ED	J2-4964-V0-R2-ED

ASSIGNED REVIEWER REPORTS: REVISIONS

JOURNAL 1 REVISIONS

First Revisions Journal 1 [1 mss]			
Author		**Editor**	
J1-4068-V1-R1-AU	J1-4068-V1-R2-AU	J1-4068-V1-R1-ED	J1-4068-V1-R2-ED

JOURNAL 2 REVISIONS

First Revisions Journal 2 [13 mss]			
Author		**Editor**	
J2-1040-V1-R1-AU	J2-1040-V1-R2-AU	J2-1040-V1-R1-ED	J2-1040-V1-R2-ED
J2-1228-V1-R1-AU	J2-1228-V1-R2-AU	J2-1228-V1-R1-ED	J2-1228-V1-R2-ED
J2-1316-V1-R1-AU	J2-1316-V1-R2-AU	J2-1316-V1-R1-ED	J2-1316-V1-R2-ED
J2-2732-V1-R1-AU	J2-2732-V1-R2-AU	J2-2732-V1-R1-ED	J2-2732-V1-R2-ED
J2-2861-V1-R1-AU	J2-2861-V1-R2-AU	J2-2861-V1-R1-ED	J2-2861-V1-R2-ED
J2-3059-V1-R1-AU	J2-3059-V1-R2-AU	J2-3059-V1-R1-ED	J2-3059-V1-R2-ED
J2-3575-V1-R1-AU	J2-3575-V1-R2-AU	J2-3575-V1-R1-ED	J2-3575-V1-R2-ED
J2-3668-V1-R1-AU	J2-3668-V1-R2-AU	J2-3668-V1-R1-ED	J2-3668-V1-R2-ED
J2-3833-V1-R1-AU	J2-3833-V1-R2-AU	J2-3833-V1-R1-ED	J2-3833-V1-R2-ED
J2-4120-V1-R1-AU	J2-4120-V1-R2-AU	J2-4120-V1-R1-ED	J2-4120-V1-R2-ED
J2-4355-V1-R1-AU	J2-4355-V1-R2-AU	J2-4355-V1-R1-ED	J2-4355-V1-R2-ED
J2-4522-V1-R1-AU	J2-4522-V1-R2-AU	J2-4522-V1-R1-ED	J2-4522-V1-R2-ED
J2-4715-V1-R1-AU	J2-4715-V1-R2-AU	J2-4715-V1-R1-ED	J2-4715-V1-R2-ED
Second Revisions Journal 2 [2 mss]			
Author		**Editor**	
J2-1316-V2-R1-AU	J2-1316-V2-R2-AU	J2-1316-V2-R1-ED	J2-1316-V2-R2-ED
J2-3668-V2-R1-AU	J2-3668-V2-R2-AU	J2-3668-V2-R1-ED	J2-3668-V2-R2-ED

CODE DEFINITIONS

Code ID#	Code	Definition	Examples†
1	**relevance of contribution**	Reviewer comments that relate to potential knowledge gains, relevance of the research topic or manuscript topic to the audience (journal readers or discipline), "practical usefulness of the findings," and the study's "importance, newness, and originality" (Bornmann, Weymuth, et al., 2010, p. 497).	"The subject is appropriate to the audience . . ." (Lay, 2004, p. 118). "I don't see much that is new in the solutions that are offered" (Lay, 2004, p. 113).
2	**writing / presentation (higher order elements)**	Reviewer comments that discuss elements such as purpose, thesis, tone, definition of terms, organization, clarity, completeness, development of ideas (e.g., examples, details, description, explanation), argumentation (e.g., claims, logic, fallacies, evidence).	"Deletion of this section and its sweeping coverage of a variety of tangential issues . . . would lead the reader more quickly and directly to the heart of the article" (Lay, 2004, p. 117). "This piece gets off to a good running start, establishing a clear and interesting thesis early . . . (Lay, 2004, p. 114).
3	**writing / presentation (lower order elements)**	Reviewer comments that discuss the manuscript's grammar and mechanics (e.g., sentence structure, punctuation, spelling, capitalization), word choice, style, citation formatting* (e.g., deviations from APA style), or professional appearance (e.g., document design). *Note: Comments related to the cited literature itself (e.g., selection, quality, quantity, interpretation, or timeliness of sources) or failure to cite the literature should be coded as **reference to the literature and documentation**.	". . . that section seems coarsely written" (Lay, 2004, p. 117). ". . . it is well written . . ." (Lay, 2004, p. 114).

Code ID#	Code	Definition	Examples†
4	**design / conception**	Reviewer comments that discuss the study's conceptual framework, research design, internal consistency, rigor, appropriateness for answering the research questions, sampling size and quality, limitations, generalizability, or replicability (Bornmann, Weymuth, et al., 2010).	"I do not believe that the manuscript is . . . fully researched enough . . ." (Lay, 2004, p. 117). "sample too small or biased" . . . (Bornmann et al., 2008, p. 421). "pilot study research with little evidence of generalizability" (Bornmann et al., 2008, p. 421).
5	**methods / statistics**	Reviewer comments that discuss the "correctness, appropriateness, and newness of methods or statistical analyses . . . operationalization of key constructs and . . . the measurement of data" (Bornmann, Weymuth, et al., 2010, p. 497).	"He/she mentions that Y and Z were not available for interviews . . ." (Lay, 2004, p. 115). "My main concern is with the single subject case study . . . (Lay, 2004, p. 115).
6	**discussion of results**	Reviewer comments that discuss the scope of the discussion, the interpretation of the data, the objectiveness and correctness the conclusion, the takeaways, implications, future research, etc. (Bornmann, Weymuth, et al., 2010).	"It appears your data may be richer . . ." (Lay, 2004, p. 116). "There are a number of concepts represented in the tables that are not discussed in the article at all" (Lay, 2004, p. 116).
7	**reference to the literature and documentation**	Reviewer comments that discuss the selection, quality, quantity, interpretation, or timeliness of cited sources; the thoroughness of the literature review; or "whether the research study . . . is embedded in the framework of the relevant literature" (Bornmann, Weymuth, et al., 2010, p. 497).	". . . there were no citations of [noted scholar's] work" (Lay, 2004, p. 113). "The author does not evaluate [the literature] for credibility, reliability, nature of arguments, and supporting data" (Lay, 2004, p. 115).

Code ID#	Code	Definition	Examples†
8	**theory**	Reviewer comments that discuss "whether the manuscript contributes to theory development or whether the theory underlying the research study seems complete and sound" (Bornmann, Weymuth, et al., 2010, p. 499).	". . . this manuscript does address the application [of theory] and does so well" (Lay, 2004, p. 116). "It almost appears that the theory was 'sitting there' . . ." (Lay, 2004, p. 115).
9	**ethics**	Reviewer comments that discuss ethical issues related to scientific or disciplinary standards (e.g., IRB, consent, plagiarism, multiple publication of same research, falsification of data)	"[Two people] were not available [to participate]; however, . . . a good deal is attributed to them" (Lay, 2004, p. 115). "'This paper is essentially identical to the paper of 1998'" (Bornmann, Weymuth, et al., 2010, p. 498).
10	**other**	Reviewer comments that do not fit the existing codes should be coded as **other**, and, if able, the coder should provide a suggested category on the coding form or reasons why the comments do not fit the existing codes.	

†Note: These published review excerpts have been interpreted outside their original context and, in some cases, have been modified slightly for illustrative purposes.

Appendix D: CATA Custom Dictionary

Broad Technical Communication Peer Review Version

Note: The journal specific dictionaries have intentionally been omitted from the appendix.

abide*
abilit*
able
absent
abstract
academ*
accept*
access*
accommodat*
accompan*
account*
accredit*
accura*
accusation*
achieve*
acknowledge*
act
activ*
add
address*
adequate*
adhere*
adjust*
adopt*
advance*
advantage*
advice
advise*
affiliat*
affirm*
age*
agree*
aid*
aim*
alert*
align*
allot*
allow*
alphabet*
alter*
ambig*
amend*
america*
amount*
analys*
analytic*
ancillary
annotat*
anonymous*
answer*
appeal*
appear*
application*
applie*
apply*
appoint*
approach*
appropriate*
approv*
approximate*
archiv*
area*
argument*
article*
ask*
aspect*
aspir*
assert*
assess*
assign*
assist*
associat*
assum*
attach*
attempt*
attend*
attention
attribut*
audience*
author*
automat*
availab*
avoid*
aware*
axis
back
backed
background
bar
base*
basic*
beginning
behave*
belief*
benefit*
best
bias*
bibliograph*
blind*
board*
bodies
body
bolt*
book*
boost*
bottom
box*
breach*
brief*
broad*
build*
call*
career*
careful*
carry*
cascad*
case*
categor*
center*
central*
certificate*
challeng*
chance*
change*
chapter*
characteristic*
chart*
check*
choose
circumstan*
citation*
cite*
claim*
clarif*
clarity
classroom*
clear*
close*
code*
coding
cogent
collaborat*
colleague*
collect*
combination*
commenc*
comment*
commercial*
committed
committee*
common*
communication
communicator*
communit*
compar*
compel*

compet*
complet*
complex*
complian*
complicat*
comply
comprehensive
compromis*
concept*
concern*
conclusion*
concurrent
conduct*
conference*
confidence
confidential*
confirm*
conflat*
conflict*
congruent
connect*
consider*
consist*
constrain*
construct*
construe*
contact*
contain*
content*
contest*
continu*
contradict*
contribut*
control*
cooperat*
copies
copy
copyright*
correspond*
council
count*
course*
coverage
create*
creation*
creativ*
credibility
credible
criteria
critic*
critique*
crucial
cultivate*
cultur*
current*
data
database
date*
day*
deadline*
decide*
decision*
declare*
decline*
define*
definition*
delay*
delete*
deletion*
demographic*
demonstrat*
denigrat*
department*
depend*
depth
derivat*
derogatory
describe*
descript*
design
designed
desk
detail*
determine*
develop*
differenc*
different
differentiate*
difficult*
direct*
disadvantage*
disagree
disclose*
discount*
discover*
discredit*
discusse*
discussion
display*
distinguish*
document*
draft*
duplicat*
earn*
economic*
edit
edited
editing
editor*
edits
education*
effective*
effort*
elaborat*
electronic
element*
employ*
encourage*
endorse*
engag*
english
enhance*
ensure*
error*
essay*
essential
establish*
ethic*
evaluat*
evidence
exact*
examin*
example*
execut*
exist*
expand*
expect*
experience*
experiment*
expert
expertise
experts
explain*
explanation*
explore*
extend*
extension
extensive
extent
extreme*
facilitat*
fact
factor*
facts
factual*
fail*
fair
false
familiar*
feature*
feedback
field*
figure*
file*
filter*
finding*
fit*
flaw*
focal
follow*
forecast*
form
formal
format*
forms
found
frame*
framing
full*
fundament*
future
gain*
gap*
gatekeeper

gateway
gender
general
generaliz*
given
global*
goal*
good
google
grammar
grammatical
grant*
graph*
group*
guidance
guide
guideline*
hand
handl*
head*
health
help*
hesistat*
heuristic*
hidden
highlight*
honest*
hostil*
house
idea
ideas
identif*
identit*
ignore*
illustration*
image*
impact*
implement*
implication*
importance
important
impression*
improv*
impugn*
includ*
incomplet*
increas*
independent*
indicat*
inference*
inferential
inflammatory
inform
information
informed
informs
initial*
innovative
insight*
inspire*
institution*
instruction*
insufficient
integrative
integrity
intellect*
intended
intention*
interaction*
interest*
international
interpret*
introduce*
introduction
investigate*
investigation*
invitation*
involv*
irregular*
issue*
item*
jargon
job*
journal*
judge*
justif*
keep
key
keyword*
know
knowledge
known
label*
language*
large*
late*
law*
lead*
learn*
legal*
length*
lesson*
letter*
level*
libel*
license*
light
likely
limit*
linguistic*
link*
list*
literature
logic*
logs
mail
main
major
majority
manner
manuscript*
material*
matter*
mean*
mediate*
meet*
member*
mentee*
mention*
mentor*
merit*
message*
method
methodology
methods
milestone*
mind
minor
minority
misconduct
missing
mission
mistake*
mixed
model*
name*
narrow*
national*
native
nature
necessary
need*
negative*
network*
new*
notat*
note*
novel*
number*
objectiv*
obligation*
observation*
observe*
offer*
online
open*
operate*
opinion*
opportunit*
option*
order
organization*
organize*
others
outcome*
outlin*
outside
overview
ownership
page*
paper*
paragraph*
paraphrase*
part

participant*
particular
partner*
party
pay
pedagog*
peer
people
perform*
permission*
person*
perspective*
pertinent
phrase*
place*
plagiari*
platform*
plausible
play*
please*
point*
policies
policy
political
poor
portfolio*
positive*
possib*
post*
practic*
practitioner*
prefer*
preliminary
prepare*
present*
prevent*
previous*
principle*
prior
problem*
procedure*
proceedings
process*
produce*
product*
profession
professional*
proficien*
profile*
program*
project*
prolong
promot*
proof
proofread*
proper*
proposal*
propose*
protocol*
provide*
psychological
public
publication*
publish*
purpose*
pursuit
qualitative
quality
quantify
quantitative
question*
quotation*
reaction*
read
readable
reader*
reading
real
reason*
receipt
recent*
recognition
recognize*
recommend*
record*
referee*
reference*
reflect
refrain
reject*
relate*
relation*
relevan*
remember
remove*
replicab*
replicate*
report*
represent*
request*
require*
research
researcher*
resource*
respect
respond*
response*
responsib*
resubmit*
result*
return*
review*
revise*
revision*
rewrit*
right
rigor*
robust
role
rubric*
sample*
scheme*
scholar*
science
scope
section*
secure*
see
segment*
select*
send*
sense*
sensitiv*
sent
sequence*
serious
serve*
service*
setting*
short*
show*
significan*
similar*
situat*
size
software
solution*
source*
special*
specific
specified
specify
standard*
state*
statistic*
status
step*
strateg*
strength
strengthen
strengths
strong*
structure*
student*
studie*
study
style*
subheading*
subject
subjective
submission*
submit*
substantial
substantiate*
sufficient*
suggest*
suitab*
summarize*
summary
support*
surround*
synthesize*
teach
teacher*

teaching
team*
technical
technique*
technolog*
term*
test*
text*
thematic
theme*
theor*
think
tie
time*
title
titled
tone
topic*
track*
train*
transfer*
transpar*
treatment*
trust
trustworth*
tutorial*
type*
typical*
unbiased
unclear
underlying
unfamilar*
unusual
usability
usable
useful*
usual*
valid*
valuable
value
verif*
version*
view*
visib*
way*
weigh*
well
whole
word*
work
workplace
world
writ*

Appendix E: Stop Words

a	his	[prepositions]
about	however	primarily
all	I	[pronouns]
along	[identifying terms]	[proper nouns]
am	in	put
an	it	rather
another	its	really
are	just	she
around	last	some
at	made	somewhat
beyond	make	the
come	[modal verbs]	their
[conjunctions]	much	they
even	my	thus
every	nevertheless	under
few	next	[URLs]
finally	nonetheless	various
first	[numerals (1, 2, 3, etc.)]	very
generally	obviously	was
he	on	we
[helping verbs]	one	were
hence	our	you
her	over	your

* The file with the complete list of stop words was corrupted; this list was reconstructed from memory. The brackets indicate general categories of stop words.