

Use of Politeness Strategies in Signed Open Peer Review

Syavash Nobarany, Kellogg S. Booth

Department of Computer Science, University of British Columbia

201-2366 Main Mall, Vancouver, BC, V6T 1Z4, Canada

1-778-994-7127, 1-604-822-8193

{nobarany, ksbooth}@cs.ubc.ca

Please send all correspondence to Syavash Nobarany at nobarany@cs.ubc.ca

Abstract

Scholarly peer review is a complex collaborative activity that is increasingly supported by web-based systems. Yet little is known about how reviewers and authors interact in such environments, how criticisms are conveyed, or how the systems may affect the interactions and use of language of reviewers and authors. We looked at one aspect of the interactions between reviewers and authors, the use of politeness in reviewers' comments. Drawing on Brown and Levinson's politeness theory, we analyzed how politeness strategies were employed by reviewers to mitigate their criticisms in an open peer review process of a special track of a Human-Computer Interaction conference. We found evidence of frequent use of politeness strategies and that open peer review processes hold unique challenges and opportunities for using politeness strategies. Our findings revealed that less experienced researchers tended to express unmitigated criticism more often than did experienced researchers, and that reviewers tended to use more positive politeness strategies (e.g., compliments) toward less experienced authors. Based on our findings, we discuss implications for research communities, and the design of peer reviewing processes and the information systems that support them.

Keywords: Politeness, peer review, open process, anonymity, reviewer

Introduction

Scholarly peer review processes have played a central role in assessment of manuscripts for publication, assessment of researchers for hiring, tenure, and promotion decisions, and allocation of research funds (Bornmann, 2011; Chubin & Hackett, 1990). In this article, we are concerned with the peer-review of scholarly scientific manuscripts for publication. This process has been one of the cornerstones of science for centuries, but it has also been subject to criticism for decades (Benos et al., 2007; Campanario, 1998; Smith, 2006). Alternative forms for the process, such as open peer-review, post-publication peer-review, and hybrid peer-review processes have been explored in various research fields (Lee, Sugimoto, Zhang, & Cronin, 2013). Evaluation of manuscripts is a social process (Merton, 1973, pp. 339–340), even in formal sciences (De Millo, Lipton, & Perlis, 1979); like other social processes, it is affected by the behavior of participating agents. We looked at behavior and language use in an open signed peer review system, where reviews were signed by reviewers and published on the web. A primary criticism of signed review is that removing anonymity of reviewers could result in individuals refusing to serve as reviewers. Critics further suggest that junior reviewers might fear criticizing senior authors in a signed or open process, particularly if submissions are not anonymized (Baggs, Broome, Dougherty, Freda, & Kearney, 2008; Khan, 2010; Smith, 1999). With many possible processes to choose from, we believe it is important to understand reviewers' behavior when using signed or open peer review processes, and (for example) to find out if junior reviewers do tend to avoid reviewing the papers of senior researchers in these situations.

Avoiding criticizing someone else is a politeness strategy, only used when a person is extremely concerned with the negative effect of criticism on another person, often the *addressee* or *recipient*, but possibly others as well (Brown & Levinson, 1987, p. 69). We therefore chose a theory of politeness as the lens through which our questions could be analyzed. We utilized Brown and Levinson's politeness theory (Brown & Levinson, 1987) to analyze reviews of alt.chi 2007, a special submission track of the annual Conference on Human Factors in Computing (CHI) that is sponsored by the Association for Computing Machinery (ACM). These were signed reviews, available to anyone interested in seeing them. Reviewers

were able to freely choose the papers they reviewed. This presented an opportunity for us to test whether junior reviewers exhibited bias in their choices of papers to review and to analyze the use of politeness strategies by a variety of reviewers for whom we could obtain relevant demographic information to use in our analysis.

Fairness is a goal of any system of peer-view. Disrespect constitutes a breach of fairness, because people believe they are entitled to polite treatment from others (Miller, 2001). In addition, polite treatment by a community influences an individual's perception of fairness and his/her commitment and loyalty toward the community (Lind & Tyler, 1988, p. 239). Gilliland and Beckstein (1996) found that authors' perception of reviewers' interpersonal sensitivity predicts authors' perception of fairness of editorial decisions and their reaction to the decisions. Politeness strategies mitigate criticisms that are being made and thus help maintain smooth relationships and interactions. Moreover, mitigated advice is found to be more effective than blunt advice (MacGeorge, Lichtman, & Pressey, 2002).

Reviewers might thus be expected to use politeness strategies to minimize the potential for conflict with authors, while still conveying their criticisms of the authors' work. Indeed, Walsh et al. (2000) found that signed reviews were more polite than anonymous reviews in a study of reviewers for a psychiatry journal. Rather than directly compare different review processes, we wanted to first better understand how politeness strategies are used in an open signed peer review process supported by a web-based system. We also examined whether junior reviewers actually do avoid reviewing papers authored or co-authored by senior researchers. Our findings include a discussion of how the design of future systems to support peer review might incorporate politeness. This is important because, although lack of civility has been a serious concern in various domains (Hadjistavropoulos & Bieling, 2000; Nicholls, 1999; Souder, 2011; Sternberg, 2002), excessive politeness has also been a concern (Baggs et al., 2008; Melero & López-Santoveña, 2001; Rose & Stevens, 1925). A proper balance needs to be struck.

The primary contributions of our work are fourfold. We conducted the first analysis (to our knowledge) of politeness strategies in an open peer review process. We observed overwhelming usage of these strategies, and we identified common combinations of the strategies and observed that several

instances of politeness strategies are particularly well facilitated by open review. We found no evidence that junior reviewers avoided reviewing papers co-authored by senior researchers, but we did find that less experienced researchers tended to express unmitigated criticism more often than did experienced researchers. We discuss how visibility of reviews to various parties such as authors, other reviewers, or the community may affect use of politeness strategies, and we identify design implications both for peer review processes and the systems that support those processes.

Related Work

We first review an influential theory of politeness, which was developed by Brown and Levinson, and we summarize some results from previous studies of politeness usage in computer-mediated communication and peer review. We then offer a brief overview of prior research on peer-review to provide a broader context for our study.

Politeness Theory

Politeness is a well-recognized concept that has been studied for over three decades in various areas of linguistics, particularly sociolinguistics and pragmatics (Eelen, 2001). It is a “*system of interpersonal relations designed to facilitate interaction by minimizing the potential for conflict and confrontation inherent in all human interchange*” (R. T. Lakoff, 1990, p. 34).

Politeness is closely related to the concept of *civility*: politeness is consideration of others and offering of good intentions; civility is withholding bad intentions and not showing hostility (R. T. Lakoff & Ide, 2005). While we might accept just civility or even tolerate low-levels of hostility in political contexts (Tracy, 2008), less contentious relationships and interactions are preferred in academic settings such as within the peer review process (K. Hyland, 1996; Myers, 1989).

Several theories have been developed to analyze and understand politeness in language usage and social interaction. One of the first studies of politeness was done by Goffman, on expressions of deference and demeanor. He studied how rules of propriety are broken by mental patients (Goffman, 1956) and later offered various conceptual frameworks for understanding politeness based on the notion of *face-work*, which he defined as behavior that “*counteracts the threats to face by avoidance or corrective processes*”

(Goffman, 1955, p. 213), and the notion of territories of the self that people claim for themselves and expect others to avoid violating (Goffman, 1971). Brown and Levinson's theory (Brown & Levinson, 1987), while debated by critics (Eelen, 2001; Kerbrat-Orecchioni, 1997; Watts, 2003), is the most influential and most widely used sociolinguistic theory of politeness (Eelen, 2001, p. 3; Watts, 2003, p. 25). It builds on Goffman's work and provides a taxonomy of universal politeness strategies and communicative choices that interpret social interactions as expressions of social relationships built through strategic use of language.

Inspired by Goffman's conceptualization of politeness, Brown and Levinson's politeness theory is built around the notion of *face*, "*the public self-image that every member [of society] wants to claim for himself*" (Brown & Levinson, 1987, p. 61). There are two types of face. *Negative face* is the basic claim to freedom from imposition (autonomy). *Positive face* is the positive self-image and the desire for approval and appreciation of that image by other people. Positive and negative face roughly correspond to Goffman's notions of *face* and *territory* (Kerbrat-Orecchioni, 1997; Watts, 2003). *Face-threatening acts (FTAs)* are acts that intrinsically attack face. Suggestions or advice threaten an addressee's negative face because they potentially impose work on the addressee. Criticism, however, suggests that the speaker does not approve, like, or accept the addressee's wants, beliefs, or products and thus threatens positive face. Brown and Levinson assume that people's actions are systematically related to their intentions, and that those intentions are reconstructible by observers or recipients of the actions (Brown & Levinson, 1987, p. 7). Based on this and inspired by Goffman's notion of face-work, Brown and Levinson define a *redressive* action as any action that can "*counteract the potential face damage of the FTA by doing it in such a way, or with such modifications or additions that indicate clearly that no such face threat is intended or desired, and that the speaker in general recognizes the hearer's wants and himself wants them to be achieved*" (Brown & Levinson, 1987, p. 70). See Table 1 for a summary of four politeness concepts.

Table 1. Summary of concepts in Brown and Levinson's theory of politeness.

Term	Definition
Face	The public self-image that every member of society wants to claim for himself
Positive Face	Desire for approval and appreciation of that image by other people
Negative Face	Basic claim to freedom from imposition (autonomy)
Face-threatening act (FTA)	Acts that intrinsically attack face, such as advice and criticism

We found Brown and Levinson's politeness theory a useful lens for understanding politeness strategies in peer review processes. Peer review is full of suggestions for improvement and criticisms about possible problems, or shortcomings of a paper written by one or more researchers, so it can be important to balance clarity and politeness using suitable politeness strategies. Brown and Levinson's taxonomy includes five strategy groups (*super-strategies*) that we use in our analysis.

1. *Baldly performing an FTA* without redressive action is a strategy that can be appropriate in situations where maximum communicative efficiency is desired, such as when communication is urgent, the communication channel is noisy, or when the focus of interaction is task-oriented.

2. *Positive politeness strategies* are oriented toward the addressee's positive face. They try to indicate that the speaker wants or approves the addressee's wants, products, or beliefs. Examples are giving compliments and expressing agreement as part of a FTA.

3. *Negative politeness strategies* are oriented toward the addressee's negative face. They try to indicate that the speaker will only minimally interfere with the addressee's freedom of action. Examples are *apologies* and *hedging*.

4. *Off-record strategies* are when a speaker's potential FTA can be interpreted in multiple ways, so the speaker cannot be seen to have committed an FTA. For example, saying "*Damn! I'm out of cash, I forgot to go to the bank today*" could be interpreted as a request for money, which threatens negative face through imposing, but it does not hold the speaker to having made an actual request.

5. *Not performing an FTA* is also a strategy. It may be employed when a face threat is so severe that avoiding it is preferred even if it disrupts the desired communication.

We recognize that use of language is influenced by several factors, so it is possible that some of the structures and strategies we observe as politeness strategies might be being used for purposes other than politeness. Brown and Levinson's theory postulates intentional use of strategies, so even normative use that might be considered unintentional is rooted in reasons and intentions that are repeatedly made. This theory allows us to identify use of politeness strategies, and to analyze the concerns about junior reviewers avoiding criticizing senior authors. According to Brown and Levinson's theory, mutual perceptions of power difference and social distance influence the weightiness of a face threat and the need for mitigating it. The higher the power of addressee over the speaker, or the farther the social distance is, the weightier the face threat becomes. When both authors' and reviewers' names are revealed, reviewers have a fairly accurate estimate of both power difference and social distance. Therefore, the concerns can be justified by arguing that less experienced (and consequently less powerful) reviewers might feel a higher need for mitigating their criticisms or avoiding them altogether.

Analysis of Politeness in Peer Review and Other Scholarly Situations

Study of the social behavior patterns of scientists dates back to seminal work of Merton (1973), who looked at patterns of collaboration, evaluation, and competition. Following Merton, numerous sociologists of science deepened and expanded our understanding of the social side of the process of knowledge production. Here, for the sake of brevity, we limit our brief review to studies that looked at politeness behavior in scholarly situations. Various discourse genres have been analyzed based on politeness strategies, but studies dedicated exclusively to referee reports or conference reviews are rather scarce. Fortanet (2008) examined structural aspects of reviews, defining a taxonomy for them based on the form of request: Question, Criticism, or Recommendation. Kourilova (1998) considered both structural aspects and politeness in reviews by Slovak researchers in a quantitative description of the reviews based on frequency and organization of compliments and criticisms, perceived politeness, and perceived competence of non-native reviewers. Johnson (1992) analyzed reviews written by students of other students' midterm papers. The students were asked to explain what they liked about a paper and to suggest ways that it could be improved, focusing on partners as their audience rather than the instructor.

Johnson's emphasis on positive aspects differs substantially from academic peer review. Moreover, Johnson focused on analyzing compliments but not other politeness strategies. Hyland and Hyland (2001) analyzed written feedback given by two teachers to students, and explored pedagogical reasons for mitigating criticisms and the risk of miscommunication. Myers (1989) described and analyzed use of politeness strategies in published articles, and discussed the face threats raised when interactions between scientists are visible to the research community (e.g., citing and praising or criticizing previous work). Salager-Meyer et al. (2007) analyzed hedging in book reviews and found hedged criticisms are more frequent than un-hedged criticisms.

Politeness theories have also been applied to online knowledge exchange for creation of an automatic politeness checker, and to understand effects of politeness on reply rates in online discussion forums and effects of using online media on community norms (Burke & Kraut, 2008; Graham, 2007). In some communities more politeness was rewarded with more answers, while in others impoliteness was more effective (Burke & Kraut, 2008).

Studies of anonymity in peer-review

In most peer review processes, either the identities of both authors and reviewers are unknown to each other (double-blind), or only reviewers' identities are withheld from authors (single-blind). Reviews are normally only seen by reviewers and authors. Over the last few decades these processes have been widely scrutinized and criticized (Armstrong, 1997; Benos et al., 2007; Smith, 2006) and alternative processes have been considered. Several studies have investigated the effects of anonymity of authors and reviewers. Most of the previous studies, but not all (McNutt, Evans, Fletcher, & Fletcher, 1990), found no significant effect on review quality of blinding reviewers to the identities of authors, or on time spent composing reviews (Justice AC, 1998; van Rooyen, Godlee, Evans, Black, & Smith, 1999; van Rooyen, Godlee, Evans, Smith, & Black, 1998). However, blinding reviewers to identities of authors sometimes affected recommendation scores (Godlee, Gale, & Martyn, 1998; Isenberg, Sanchez, & Zafran, 2009). Several studies identified non-blind reviewers' biases in recommending papers for acceptance, favoring authors from English-speaking countries and from prestigious institutions (Blank, 1991; Peters & Ceci,

1982; Ross et al., 2006). Walsh et al. (2000) found that signed reviews were more polite, higher quality, and more likely to recommend papers for publication, whereas van Rooyen et al. (1998) did not identify any effect on the quality of reviews. Van Rooyen et al. also found that unmasking reviewers increased the chance of reviewers declining review requests (23% vs. 35%) and that most authors preferred knowing the identity of reviewers (55% in favor vs. 26% against). In contrast, Melero, and López-Santoveña (2001) found that 75% of reviewers were in favor of masking reviewers and 17% were against it. Similarly, a large survey of attitudes toward various peer-review practices, which was a follow up to an earlier similar survey (Ware & Monkman, 2008), found that despite the difficulties in implementing double-blind peer-review, 76% of researchers consider it to be effective, but only 25% considered open and published peer-review to be effective (Mulligan, Hall, & Raphael, 2012).

While several studies have investigated the effect of anonymity of reviewers on the quality of reviews, the potential effect on other types of biases such as biases in favor of positive results (Mahoney, 1977) or of results that match the reviewers' stance (Abramowitz, Gomes, & Abramowitz, 1975) or that corroborate their previous work (Ernst & Resch, 1994) is largely unknown. Revealing the identity of reviewers might reduce some of the possible misbehaviors over the course of the peer review process, such as requesting reviews from competing researchers or from researchers with conflicting stances, or delaying the publication of research that competes with the reviewers' research (Campanario, 1998). However, these problems are probably inevitable in narrow research areas if only a few researchers are qualified to serve as reviewers (Riggs, 1995). Shatz (2004) offers a comprehensive analysis of arguments for and against anonymity of authors and reviewers.

Opening up peer review so that every reader can review papers through post-publication peer review or review of preprints has moved into the spotlight over the last decade. Van Rooyen et al. (2010) found that telling reviewers that their signed reviews might be published on the web did not affect the quality of reviews, but it did increase the time spent on composing reviews. Bornmann and Daniel (2010) analyzed public reviews submitted to the journal *Atmospheric Physics and Chemistry* and found that the level of inter-reviewer reliability was low, comparable to those of traditional peer-review processes.

Bingham et al. (1998) found that post-publication reviews by online readers can provide valuable feedback, but those reviews are often short and specific, and thus do not adequately replace editorial peer review. The journal *Nature* found similar results and received little interest in this type of review from either authors or reviewers; only 5% of papers were made available by their authors for open review (sometimes for fear of being scooped or compromising patent applications) and only 54% of those received any comments, most of which were not of editorial value (“Nature’s trial of open peer review,” 2006).

These studies have identified a number of expected pros and cons of open and signed peer-review processes. We also have evidence from several journals that have adopted signed peer review processes, in which reviewers sign their reviews so authors know who the reviewers are, in contrast with an anonymous process. The *British Medical Journal (BMJ)* is one of the more prestigious journals to adopt signed peer review. Other journals have adopted open peer review processes, in which reviews are published (e.g., placed on the web as supplementary material or as part of the history of a paper). The *European Molecular Biology Organization (EMBO)* journals use an anonymous open process where anonymized reviews of accepted papers are published on the web (Pulverer, 2010). Journals such as *BioMed Central (BMC) Medicine* and *BMJ Open* use both a signed and open process.

In the Human-Computer Interaction research community, alt.chi, a special submission track of the annual ACM Conference on Human Factors in Computing (CHI), uses a signed open process. It is one of the early attempts at adopting an open and signed peer-review process within the HCI community and thus offers an opportunity to understand various aspects of authors and reviewers’ behavior in this evolving type of peer review process. According to the 2007 call for papers, alt.chi recognizes “controversial ideas, novel prototypes, failed but valuable user studies, bold experiments, and anything else that can give a fresh perspective on CHI [... alt.chi in particular invites] work that would otherwise not have been presented at CHI 2007 - whether because it has been rejected from other venues, or because it is too different to be submitted in the first place” (“Alt. CHI,” 2007, para. 3). In the remaining sections, we describe the analysis of a corpus of alt.chi reviews from 2007, with the goal of increasing our

understanding of the use of language, especially the use of politeness mechanisms by reviewers in open and signed peer-reviewing environments as it used in practice. Politeness has been understudied in the context of open and signed peer review and has been subject to assumptions in the literature. Our study looks for evidence for the use of politeness and for some of the other expected pros and cons predicted in the literature that can be examined through the lens of politeness theory.

Method

We investigated how reviewers use politeness strategies to achieve their desired level of politeness in an open signed process, what factors might affect the use of these strategies, and whether junior reviewers actually avoid reviewing papers co-authored by senior researchers. We employed both qualitative and quantitative analysis methods. Brown and Levinson's theory of politeness guided our analysis. The primary reason for using a linguistic analysis approach to this research was that identifying the nuances of use of language that reviewers might employ to mitigate their criticisms could have been very difficult within a standard open coding corpus analysis. We needed a reliable way of identifying and understand usage of politeness strategies. We decided that we would benefit from using Brown and Levinson's theory in our analysis. In the following subsections we describe our data collection and data analysis methods, as well as a profile of the participants in the alt.chi 2007 peer review process.

Data Collection

The complete set of 165 reviews for the 44 papers submitted to alt.chi 2007 was used as the corpus for this study. All reviews were signed, so we were able to identify all of the reviews submitted by each reviewer. All submitters were required to review at least three other submissions. This reduced self-selection bias for participating as a reviewer. Our primary reason for looking at alt.chi reviews instead of other open publication venues was our familiarity with the research domain and the research methods used within the alt.chi community. This meant that we were better able to understand the criticisms made by reviewers.

Profile of the alt.chi 2007 Reviewers

The corpus of reviews had 65 reviewers (21 females) who were industry researchers, faculty members (at all ranks), postdoctoral fellows, and students. We extracted years-of-publishing-experience for all reviewers and for the most experienced co-author of each paper using a variety of sources including Google Scholar, the DBLP Computer Science Bibliography, and reviewers' home pages. The proceedings of the conference provided the primary affiliations of the reviewers. Table 2 describes the profile of all reviewers, contact authors, and senior co-authors of 44 submissions. Of these, 39 reviewers were contact authors for one or more submissions, one reviewer was a co-author of a submission but not a contact author, and the other 25 reviewers were neither contact authors nor co-authors of submissions. The contact author for a submission was the co-author who made the submission, not necessarily the first author or the most experienced author. All but one contact author served as a reviewer. The senior co-author was the co-author who had the most years of publishing experience based on our assessment. Twenty of the 38 senior co-authors were the contact authors for their submissions.

Table 2. Profile of the 65 reviewers, the 38 senior authors, and the 40 contact authors

<i>Profile of Reviewers</i>		<i>Profile of Senior Authors</i>		<i>Profile of Contact Authors</i>	
Publishing Experience		Publishing Experience		Publishing Experience	
1 year or less	21	1 year or less	6	1 year or less	16
2-4 years	19	2-4 years	7	2-4 years	10
5-9 years	13	5-9 years	9	5-9 years	8
10-19 years	8	10-19 years	9	10-19 years	3
20 years or more	4	20 years or more	7	20 years or more	3
Type of Affiliation		Type of Affiliation		Type of Affiliation	
Academia	43	Academia	28	Academia	29
Industry	10	Industry	5	Industry	4
Other	12	Other	5	Other	7
Language of Country of Affiliation		Language of Country of Affiliation		Language of Country of Affiliation	
English	48	English	31	English	33
Other	17	Other	7	Other	7
Gender		Gender		Gender	
Female	21	Female	13	Female	11
Male	44	Male	25	Male	29

Data Analysis

Determining the appropriate unit of analysis was a challenging step in the analysis process (Zhang & Wildemuth, 2009). Deciding the unit of analysis can be based on the theory guiding the analysis, on the purpose of the study, or on the nature of the corpus. Brown and Levinson do not impose a specific unit of analysis; hence we considered a single sentence, a criticism, and a full review as potential units, but ultimately decided that a criticism best suited the purpose of our study. This allowed associations between types of criticisms and politeness strategies and also enabled counting the prevalence of politeness strategies in mitigating criticisms. We defined a criticism to be a topical chain (Geisler, 2003) from a review that expressed a critical observation on an issue. A total of 532 criticisms were extracted from 165 reviews conducted for alt.chi 2007. The relevant context around each criticism was included in the excerpts to facilitate interpretation and to ensure that any compliments surrounding criticisms were included in the excerpts. We also added various metadata such as review ratings and expertise ratings as codes to the criticisms, and we analyzed the co-occurrence matrix of codes to explore relationships between politeness strategies, scores, and criticized issues.

A first coder (the first author) coded and examined 20% of the criticisms for occurrences of politeness strategies according to categories adapted from Brown and Levinson's theory and for the type of issues that were criticized. This initial coding process resulted in a codebook of politeness strategies and criticized issues. Criticisms accompanied with no politeness strategy were coded as bald-on-record, unless the review was in the form of bulleted lists of pros and cons, because although the bulleted lists separated compliments from criticisms we thought it reasonable to consider this to be a mitigation because both positive and negative were presented. The criticized issues were categorized into six major categories: (1) writing issues, (2) adequacy of the literature review, (3) technical and methodological errors, (4) level of contribution or novelty, (5) clarity or missing information, and (6) premature submission or need for more work. The descriptions and examples of the codes were discussed with a second coder (not an author) who independently coded 10% of the corpus. Inter-coder reliability (considering both inductive codes and strategy codes), calculated using a Pooled Cohen's Kappa measure

(Vries, Elliott, Kanouse, & Teleki, 2008), was 0.95 and falls within the range 0.81–1.00 indicating almost perfect agreement according to the guidelines recommended by Landis and Koch (1977). As a secondary statistic, we also looked at the average Cohen's Kappa for comparison, which was 0.70 and falls within the range 0.61–0.80 indicating substantial agreement according to Landis and Koch (1977). After establishing reliability, the first coder completed the coding without a second coder.

The only one of Brown and Levinson's politeness super-strategies that did not appear in our coding was "*Don't do the FTA*". The alt. chi reviewing form does not allow for confidential comments to the program committee (editors), therefore we were not able to identify when reviewers avoided making a comment to authors that they might make to if it were confidential, nor were we able to determine if anyone chose not to write a review. In order to get some insight into possible usage of this strategy, we instead looked for clues that might help us understand if some reviewers had avoided reviewing some of the papers. Because alt.chi allows reviewers to select the papers that they want to review, and based on the concerns raised in the literature about junior reviewers having fear of criticizing senior researchers (Baggs et al., 2008; Khan, 2010; Smith, 1999), we suspected that junior researchers might be less likely to review submissions that were co-authored by senior researchers. To examine this potential use of the "*Don't do the FTA*" strategy and to ascertain whether junior authors avoided reviewing work by senior authors, we analyzed the relation between reviewers' experience versus authors' experience to see if there was a pattern.

We also investigated associations between usage of politeness strategies and factors including reviewers' experience, authors' experience, rating of the paper, the final decision about the paper, and the self-rated expertise levels of reviewers. We used a mixed effects logistic regression model (Baayen, Davidson, & Bates, 2008) to account for variation across reviewers and across papers.

Findings

We identified frequent usage of politeness in the corpus. Approximately 85% of the criticisms were mitigated by at least one politeness strategy. We describe the use of each strategy and illustrate with examples from the corpus.

Criticisms with no redressive action

Approximately 15% of the criticisms were made with no redressive action. A mixed effects logistic regression analysis was used to assess main effects of reviewers’ experience, reviewers’ gender, reviewers’ type of affiliation (academia, industry, other), official language of reviewers’ country of affiliation (English or other), authors’ experience, reviewers’ self-rated expertise levels, reviewers’ summary ratings of papers, and acceptance decisions on the use of redressive action, while controlling for random effects of reviewers and papers. The analysis showed that years of publishing experience of reviewers, and the reviews’ corresponding ratings were significant predictors of using no redressive action (see Figure 1 and 2). More experience decreased the odds of criticizing with no redressive action ($z=-2.38, p<.05$), and criticisms mentioned in reviews with lower ratings were more likely to use no redressive action ($z=-2.74, p<.01$). There was no interaction between authors’ experience and reviewers’ experience on the use of redressive action ($p=.53$).

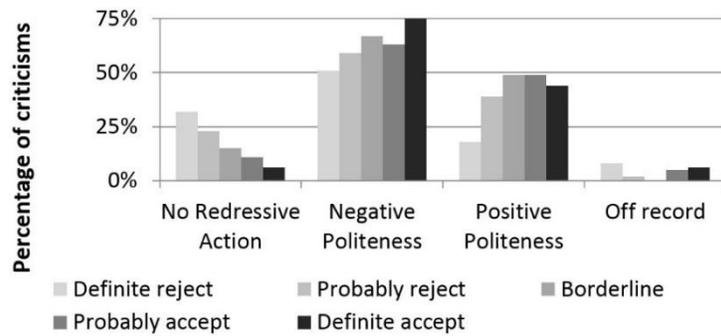


Figure 1. Relation between use of politeness super-strategies and the ratings assigned to papers by the reviewers

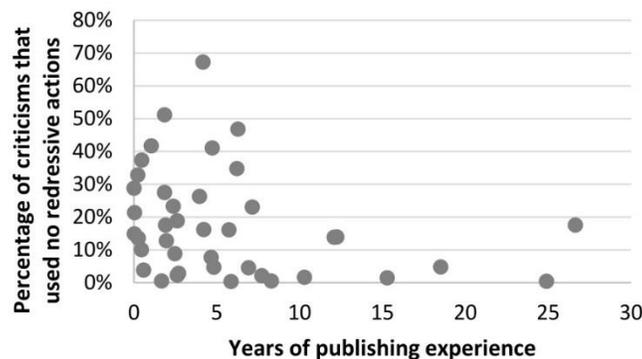


Figure 2. Relation between reviewers' years of publishing experience, and their not using redressive actions. Each point represents a reviewer. Only reviewers with more than 5 criticisms are depicted (jitter is added to reduce overlap).

We looked at various categories of criticized problems and found that 47% (8 out of 17) of the criticisms of minor writing issues were made with no redressive action. This suggests that some reviewers considered mitigating criticisms related to minor writing issues to be unnecessary. Moreover, 33% (5 out of 15) of criticisms related to major errors in considering related work were made with no redressive action. One possibility was that they were so easy to fix that they could be ignored, while another possibility was that they were considered so important that they should not be mitigated. Based on the scores associated with those reviews (three were rated as 2 and two were rated as 1) and their tone (R32 wrote "*There are also no references whatsoever to back up claims made...*"), the second explanation seems more likely. Seriousness of the problem, such as omission of a large part of the literature, seems to be enough to prevent reviewers from using politeness strategies to mitigate face threats (an apparent preference for clarity over politeness). However, some reviewers managed to clearly convey criticisms about similar problems while still using various politeness strategies. For example, R3 wrote: "*The authors argue that 'very little research is being done ... to evaluate healthcare informatics systems.' There may be very little, but there is some, [2 examples are mentioned]. Reference to, and possibly discussion of, this and similar studies seems at the very least relevant, if not necessary, in the context of this paper.*" This excerpt attempts to avoid disagreement (a positive politeness strategy) by saying "*there may be little*" and tries to further minimize the imposition by hedging using "*possibly*" and "*seems*". Unsupported claims, one of the codes in the category of methodological errors, were also frequently criticized with no redress (25% of them had no redress). Unsupported claims are often considered to be serious issues by reviewers. In these cases efficiency appears to be the goal. The reviewer probably wants to make sure there is clear communication and is less worried about politeness.

Criticisms mitigated through negative politeness.

Negative politeness strategies (redressing imposition) were the most commonly used strategies for mitigating criticisms, regardless of the score that accompanied the review or the subject of the criticism (Figure 1 and Figure 3). Approximately 66% of the criticisms were mitigated by at least one negative politeness strategy. Regression analysis (similar to the one described for criticisms with no redressive action) did not identify any significant predictors for using negative politeness strategies. Hedging was the most commonly used negative politeness strategy; 55% of all criticisms used at least one hedging device. For example, R63 wrote the following to criticize the clarity of a paper: “*Sorry, It was hard for me to follow the paper. I get lost to understand its discourse in snippets of the novel.*” In this criticism, two negative politeness strategies including apologizing and using a self-deprecating hedging device (i.e., “*hard for me*”), are put together to minimize the face threat. In a different example, R4 criticized clarity of another paper and its lack of examples, by saying “*It was hard to tell how it should be practiced in the real-world. Why didn’t the paper describe some examples of how the method is applied? Shouldn’t that be a part of every method paper?*” In this example, questioning and stating the problem as a general rule in the form of a rhetorical question are the negative politeness strategies that were used. Although rhetorical questions often serve as an off-record strategy, it is clear in this case that the criticism is on record and the rhetorical question is acting as a negative politeness strategy.

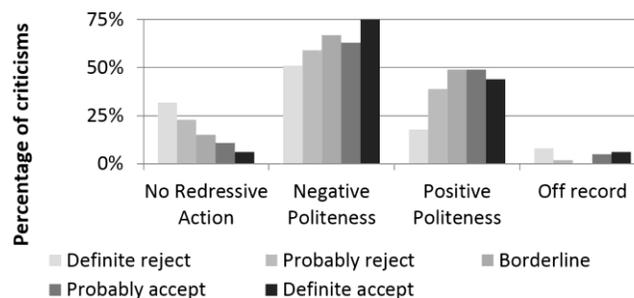


Figure 3. Relation between categories of criticized issues and politeness strategies.

Using multiple hedging techniques was also common. R60 wrote “*In my reading of the paper, it seemed as though there were times when the metaphor seemed to be forced.*” This uses hedging to

mitigate criticism by both framing it as R60's understanding and opinion, and twice expressing doubt about the existence of the problem (using "seemed"). Another example is R9's criticism on lack of a justification for the chosen control condition: "How is it that writing formula's [sic] on paper relates to the building of a knowledge base? If there is a strong connection then I think the authors need to be clearer at [sic] to what it is." In this example, questioning, mentioning it as an opinion ("I think"), and toning down the lack of clarity by saying "clearer" instead of "clear" were the negative politeness strategies that were used. Sometimes, hedging might be reflecting the reviewer's actual doubt (Myers, 1991), but in many cases it was clear that it was used to tone down a criticism. Figure 4 suggests that confidence of reviewers in their expertise (i.e., self-reported expertise levels) does not affect hedging.

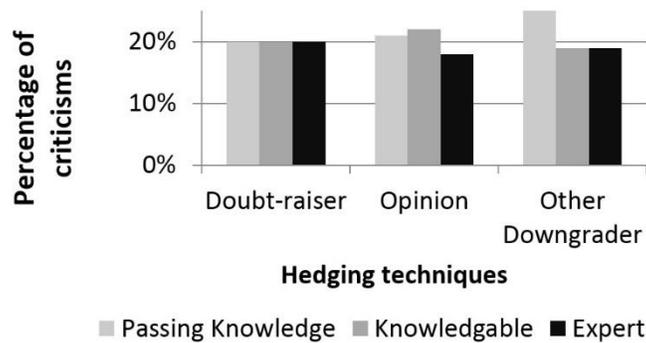


Figure 4. Relation between use of hedges and self-reported levels of expertise to the reviews.

Another negative politeness strategy is to signal a lack of desire to express a criticism. For example, R14 wrote "Having worked on this particular project [cited and discussed in the paper], I am afraid to say that this in [sic] inaccurate." Saying "I am afraid" serves this purpose. In addition, R14 framed the error in describing previous work as an inaccuracy in order to tone down her serious criticism.

Criticisms mitigated through positive politeness

Positive politeness strategies (approval of the addressee) were the second most frequently used set of strategies (Figure 1 and 3). Approximately 45% of criticisms were mitigated by at least one positive politeness strategy. Regression analysis (similar to the one described for criticisms with no redressive action) suggested that criticisms mentioned in reviews of articles by less experienced authors ($z=-3.55$, $p<.001$), reviews accompanied with higher ratings ($z=2.88$, $p<.01$), or reviews accompanied with higher

confidence ratings ($z=1.9, p<.05$) were more likely to be mitigated by positive politeness strategies. There was no interaction between authors' experience and reviewers' experience on the use of positive politeness ($p=.57$).

Complimenting (noticing or attending to the hearer, according to Brown and Levinson's terminology) was the most frequently used positive politeness strategy; 24% of all criticisms included at least one compliment. The most common structure was connecting two separate statements, one complimenting the paper and one criticizing it. For example, R3 wrote "*Overall, there are a lot of great ideas going on in this paper. However, a number of them could benefit from much more in-depth exploration.*" There were also a variety of non-standard uses of compliments. For example, R41 wrote "*I would strongly consider editing this paper down to a coherent and tight thirteen pages to fit within the chairs' recommendations. it would give a chance to focus on the excellent meat of the argument without distraction.*" In this statement, in addition to the use of a compliment ("*excellent meat*"), writing "*I would strongly consider*" demonstrates use of a personal-centre switching strategy (i.e., speaker talking as if he is the addressee), which is another positive politeness strategy. By writing "*the paper does not compare the tool to many other similar tools. I understand that this may not be feasible, or even necessary, as the authors classify their article as a short-paper,*" R24 mitigated his criticism about the lack of comparative evaluation by sympathizing about the difficulty of performing it and by downgrading the need for it, which combines both positive and negative politeness. In another excerpt, R3 wrote "*Given that the authors seem to be trying to keep the submission short, it might be beneficial to try to narrow the scope somewhat, as they touch on many ideas that they cannot fully pursue. That said, many of the points raised in the paper merit consideration and discussion in the CHI community.*" Here, R3 attends to the want of the authors regarding the length of the paper; however, his suggestion for narrowing the scope implies removing some of the ideas mentioned. He thus continues with complimenting those ideas to further mitigate his criticism.

Expressing agreement and avoiding disagreement is another positive politeness strategy. The following two examples illustrate criticism prefaced by stating agreement with related issues addressed in

a paper. R41 wrote *“There are certainly similarities between these types of negative results, but there are important, if subtle, differences that should be discussed in a paper about negative results.”* R3 wrote *“It is indeed one approach to understanding ways of creating and validating knowledge in HCI, but it is far from the sole approach.”*

Including both the reviewer and the author in a criticism is another way of suggesting that they are cooperators. For example, R46 wrote *“All we have done is start to catch up with real design”* to include himself and perhaps the research community in the target of his criticism. This could also be considered a form of personal center switch, where the statement is mostly about the addressee, but by including the reviewer and the community, the reviewer shares the blame. R15 used a similar structure: *“I think we might also have to do some work to resituate (sorry!) that critique within contemporary conditions.”*

Off-record strategy

Only 3.6% of the criticisms were mitigated by off-record politeness strategies. Regression analysis (similar to the one described for criticisms with no redressive action) suggested that criticisms in reviews of more experienced reviewers were more likely to be mitigated by off-record strategies ($z=3.18$, $p<.01$). There was no interaction between authors' experience and reviewers' experience on the use of off-record strategies ($p=.52$). There appeared to be a spectrum of being off-record. Some statements suggested what they were hinting at more clearly than did others. For example, by writing *“It is a descriptive scheme that presents how ONE metaphor was used to evaluate ONE system”*, R60 rather clearly pointed to the limited contribution of the work, whereas R30 writing *“I feel that the paper is more of a grant proposal”* does not clarify which aspect of the paper is being criticized, but it can be guessed that the paper does not report any work that is already completed. There is a risk that off-record criticisms might be misunderstood by readers and similarly that some statements that are not criticisms might be interpreted as off-record criticisms in the context of a peer review. For example, R64 wrote *“While the authors do not mention this work, it does remind me of [a reference]”* This could be interpreted as an off-record criticism of omission of an important reference, or it may be suggesting an interesting connection

between two works. Some examples are accompanied by sufficient discussion in the reviews to recover from the disadvantages of being off-record.

Combining strategies

Two or more politeness strategies can be combined to strike a reviewer's desired balance of clarity and politeness. R3 wrote "*there is a conspicuous lack of any sort of evaluation section. It could be that this paper is intended more as an artistic submission than a research submission, but, as the author notes, it is very well positioned to explore some import [sic] research questions*". This is a great example of multiple politeness strategies. R3 offers a possible justification for the addressee's decision to not evaluate a system by showing his understanding (or an attempt towards it). In addition, he gives a reason why he is criticizing the decision. Giving reasons is a way of suggesting that the speaker and the addressee are cooperating. In addition, the reasoning embeds a compliment to further mitigate the face threat.

Positive politeness strategies were frequently combined with negative politeness strategies. For example, R9's comment "*I think, perhaps, that the paper would be best presented as a demo as I suspect it would help make their achievements more accessible*" combines multiple hedges, and a compliment, and indicates caring about the author's work and making it more accessible. Framing a problem as an area for improvement or work to be done was a common strategy that combines positive and negative politeness. For example, R3 wrote "*It would be highly interesting to install this system in a public place for an extended period of time.*" This frames the problem of lack of evaluation as exciting future work. In doing so, the imposition is mitigated through impersonalization (saying "*it would be*" instead of "*I would like you*"). R3's interest in the system described in the paper becomes apparent by saying "*highly interesting*" instead of "*necessary*", a positive politeness strategy.

No evidence of using an avoidance strategy

While we expected to see junior reviewers avoid reviewing papers co-authored by senior researchers, as was predicted in the literature (Baggs et al., 2008; Khan, 2010; Smith, 1999), we did not observe any evidence of that behavior or in fact any selection pattern related to the experience of

reviewers and authors, which was operationalized as years of publishing experience (Pearson's $r = 0.02$, $p = .75$, see Figure 5). To examine the interaction effect between experience of authors and reviewers on ratings, we conducted a mixed effects linear regression analysis assessing main effects on ratings of reviewers' experience, reviewers' gender, reviewers' type of affiliation (academia, industry, other), official language of the reviewers' country of affiliation (English or other), authors' experience, and self-rated expertise levels, while controlling for random effects of reviewers and papers. Higher self-rated expertise levels ($F = 12.23$, $p < .01$) were associated with higher ratings. Interaction between authors' and reviewers' experience was not significant ($F = 0.2$, $p = .64$). One possible explanation is that not only did junior researchers avoid reviewing papers' coauthored by senior researchers, but so did other senior researchers. To investigate this hypothesis we examined the correlation between the number of reviewers per paper and publishing experience of the most senior coauthor of the papers. We did not observe any significant correlation (Pearson's $r = -0.01$, $p = .94$), suggesting that seniority of authors played little or no role in determining the number or seniority of reviewers.

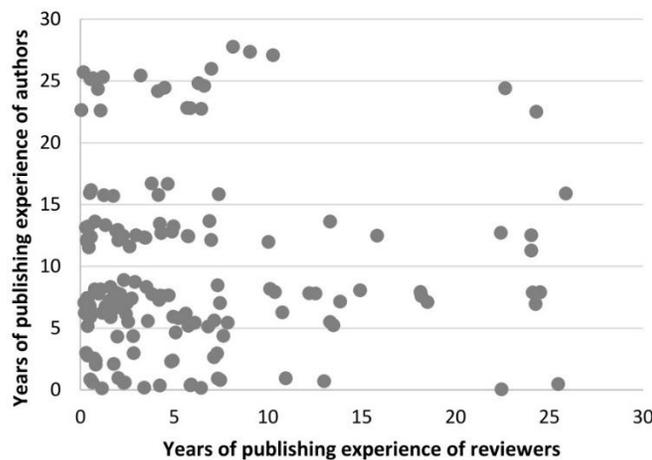


Figure 5. Relation between years of publishing experience of reviewers and authors. Each point represents a review (jitter is added to reduce overlap).

Open signed processes create new needs for and offer new ways of using politeness strategies

We found evidence that a signed non-blink process offers new ways for asserting common ground and agreement. For example, R6 explained his interest in the issues discussed by a paper by writing “Indeed, I am currently writing a paper on this theme.” Similarly, R52 wrote “I couldn't agree

more, in fact, my own alt.chi submission supports this thought very well” and R46 showed his awareness of the author’s previous work (showing care for the author’s work and attending to it), while criticizing the lack of novelty by writing “*some of the authors have already published similar material, e.g., [Reference] covers a lot of the same ground.*”

A signed process enables reviewers to show how much they care about the work under review, and why their comments should be taken seriously. For example, by writing “*Having worked on this particular project [cited and discussed in the paper], I am afraid to say that this in [sic] inaccurate,*” R14 clarifies that she is sure about what she is saying because she was part of the cited project. This reveals a knowledge difference, and may affect the assessment of weightiness of the face threat. In another review, R41 wrote “*As co-organizer of the [a workshop on the topic of the paper], I am of course pleased to see [author]’s engagement with the issue, but I’m disappointed to see an absence of any consideration of the literature.*” Mentioning the qualifications of the reviewer again reveals knowledgeability and high scholarly power of the reviewer (affecting the mutual perception of power difference) which could reduce the weightiness of the face threat, and consequently reduce the need for mitigating subsequent criticisms. In this case, the reviewer also used his relevance to the topic to welcome the author as a newcomer into the research area discussed in the paper, which shows that he attends to the author and her work (a positive politeness strategy). In another example, R38 wrote “*I don’t understand. I am left with a feeling that there must be something more deep than I can imagine... When reading about the author on the net I kind of find the author amusing and interesting.*” In this case knowledge about the background of the author enabled the reviewer to compliment him.

We also found evidence of politeness strategies used to preserve other reviewers’ face. For example, R53 mentioned “*I would agree with [another reviewer’s name]’s concerns regarding the instructions and counter-balancing during the evaluation,*” and R64 wrote “*I agree with [another reviewer’s name] that the authors present their tugging metaphor well...*” Expressing agreement with other reviewers is possible in anonymous processes as well; however, when names are attached to the

reviews, mutual perception of power difference and social distance can affect the use of politeness strategies.

It is possible to use politeness strategies to preserve face for any member of the community. For example, R60 wrote “*While the authors do not mention this work, it does remind me of [name of the author of a related work],*” which reveals the reviewer’s awareness of and attention to the work of the author of the cited related work. In this case the author of the cited work was the reviewer’s coauthor in another submission, but more generally in an open review process third parties (such as an author of a cited work) might see reviews, so preserving their face (in this case by not failing to recognize their work’s relevance) becomes important. Myers (1989) discusses how who to cite can be a type of politeness strategy. Bornmann and Daniel (2008), Brooks (1986), Camacho-Miñano and Núñez-Nickel (2009), Harwood (2009), and Sugimoto and Cronin (2013) provide analyses of citation behavior and motivations of authors and reviewers.

Finally, politeness strategies can be used to preserve face for the community as whole. For example, R46 wrote “*The main conclusions are obvious to anyone with a background in philosophy/sociology of knowledge, to the extent that undergraduate education texts on curriculum theory cover this sort of ground. However, psychology curricula steadfastly avoid such corrupting influences.*” R46 starts by praising the part of the community with a background in philosophy/sociology of knowledge (positive politeness), and then sarcastically criticizes “*psychology curricula*” instead of those with a background in psychology to mitigate the face threat to psychologists as a group.

Discussion

We had three goals for our research: (a) to understand the literature on politeness strategies, (b) to use that understanding as a framework to investigate whether politeness strategies are used by reviewers in open peer review, and (c) to identify how peer review support systems could be improved based on that understanding. In this section we return to the framework and the literature to make recommendations for how future support systems for peer review could incorporate recognition and support for the politeness strategies that were observed in our study.

If all criticisms were expressed bald-on-record (non-redressed) we could conclude that the communication is task-focused and that reviewers do not have any concerns about face threats. Quite to the contrary, we identified overwhelming use of positive and negative politeness strategies in the alt.chi corpus. This indicates a high level of concern about mitigating face threats by reviewers and presents an opportunity to use the literature on the use of politeness strategies to identify how support systems can explicitly incorporate these strategies into the peer-review process.

While reviewing is an inherently face-threatening act, the openness of the review process for alt.chi 2007 may have contributed to the frequent use of politeness strategies. According to the theory of politeness, mutual perception of power difference and social distance between authors and reviewers affect use of politeness strategies. These two variables are in part determined by knowing the identities of authors and reviewers, which were disclosed in the review process for alt.chi 2007. In a previous study by Walsh et al. (2000), signed reviews were found to be more polite than anonymous reviews.

We observed instances of politeness towards other reviewers, toward other researchers (that are not directly involved in the interaction), and toward the community as a whole. When a communication is published, readers become part of the discourse and some face threats become more severe because of the high power the community as a whole has over its members (Myers, 1989). In addition, studies of interactional justice have found that some criticisms may only be appropriate if delivered in private (Miller, 2001). Making criticisms in public, therefore, might require more deliberation on the part of reviewers to consider the involved parties and to ensure that the face threats are appropriately mitigated. We therefore conjectured that the level of use of politeness strategies could be higher in open signed reviews, and that composing them could be more time-consuming compared to reviews that are not signed or are not published. Indeed, in a study by van Rooyen et al. (2010), the reviewers who were told that their signed reviews might be published on the web spent more time writing their reviews, but the quality of their reviews was judged to be normal. We suspect the extra time spent composing reviews might have been spent addressing politeness concerns and mitigating criticisms because of the increased weightiness of face threats implicit in publication on the web. This is an important avenue for future

research, particularly because increased load on reviewers has long been a concern in the design of peer review processes and support systems for peer review.

To our surprise, the only effect that the experience of authors had on the use of politeness strategies was that less experienced researchers received more positive politeness (e.g., compliments) from reviewers regardless of the experience of the reviewers. In addition, we did not observe any evidence of junior reviewers avoiding reviewing papers authored or co-authored by senior researchers. This does not mean that there was no selection bias. Nevertheless, the fact that there was no indication of such bias calls for further research on this issue, particularly because this is one of the primary concerns about using signed peer review processes (Baggs et al., 2008; Khan, 2010; Smith, 1999). Our preliminary analysis of open peer reviewing supported these concerns, because reviewers that are less experienced (and consequently perhaps less powerful within an academic community) might feel a higher need for mitigating their criticisms or avoiding them altogether. However, what was missing from that analysis was that in an open process the research community plays the role of an audience or a bystander and changes the balance of power. What we observed in the corpus was a high level of mitigation of criticisms from most researchers, which could be attributed to the presence of a powerful audience. We conjecture that the presence of a powerful audience may have an equalizing effect on the power differences between authors and reviewers.

Another surprising finding was the highly significant correlation of ratings and self-rated expertise levels. One possible explanation is that experts are able to select better papers for reviewing. Another possible explanation is that reviewers tend to highly value the contributions to their own area of expertise. It also might mean that self-rated expertise levels are used as communication acts, for example to affect how much reviewers' ratings should influence the fate of a paper. While this finding was not related to the core of our study of use of politeness, it highlights another important avenue for future research.

We found that positive politeness strategies (compliments, agreement, etc.) were more likely to be used by more confident reviewers (based on self-rated expertise levels), and towards less experienced

authors. This corroborates previous research in pragmatics that found compliments often flow between those with equal status, or from higher status to lower status (Herbert, 1989), and that complimenting requires the complimenter to have some level of confidence (Holmes, 1986). Not surprisingly, positive politeness strategies were more likely to appear in papers with higher ratings, which could be because those papers have more aspects that deserve compliments.

Interestingly, based on previous work (Holmes, 1988; Tannen, 1998) we expected to see some gender differences, but we were not able to identify any in our analyses. This could be due to the fact that the formal and public context, and the nature of the activity required high levels of use of politeness strategies from everyone. A similar discrepancy with the literature was found in an analysis of voicemails in a legal context (Hobbs, 2003). Another possible contributing factor could be the characteristics of the HCI community, which has valued gender equality and has had more participation by female researchers in comparison with some other communities within computer science (Cphoon, Nigai, & Kaye, 2011).

Potential implications for the design of peer review processes and support systems

Based on our observations, combined with what we know from the literature on the use of politeness strategies, we recommend four ways to improve the design of peer-review processes and the software systems that support them. Assessment of the effectiveness of our design recommendations requires further studies; hence, our recommendations are primarily to suggest avenues for future research.

Enhancing support for interactive communication during peer review process. It is known that use of redressive actions reduces clarity (Brown & Levinson, 1987, pp. 90–91), so balancing between politeness and clarity is a challenge (Mackiewicz & Riley, 2003). For example, in our observations hedging was one of the most frequently used strategies. Hedging devices can be used to convey fuzziness, such as degree of certainty or precision (G. Lakoff, 1973), or as a manipulative strategy for saying less than what is meant (Hübler, 1983, p. 23). The dual function of hedges diminishes clarity of communication (Cutting, 2007) and enables using them as a politeness strategy. In the context of peer-reviewing this lack of clarity can make it difficult to assess the severity of criticisms. Consequently,

editors and program committee members might be misled in their judgments, and authors might be misled in formulating their responses to the reviews.

Politeness strategies are only one of the reasons for ambiguity in criticisms. Various other complexities of human communication, in addition to individual and cultural differences in use of language, can impose ambiguity. What makes it even more challenging is that in many peer review venues this communication often consists of only one or two steps (authors receive reviews only, or authors see reviews plus additional feedback from reviewers after the reviewers see the authors' rebuttal to the original reviews). Having only one or two steps, instead of a more extended multi-step process may not provide sufficient opportunity for negotiation and clarification. We recommend that peer review support systems enable more interactive discussions to facilitate richer communication between authors, reviewers, and program committee members (for conferences) or editors (for journals) for examining and ensuring how to interpret comments and criticisms in reviews. The peer review process plays an important role in the advancement of science. It is thus desirable to avoid misinterpretations that may lead to misjudgments.

The alt. chi process allows participation of authors in the process prior to the final decision-making steps by a jury of experts. The *Frontiers* series of journals is one of the publication venues that has implemented another type of interactive discussion forum: after finishing an initial round of traditional independent reviews, authors and reviewers (while remaining anonymous) participate in an online discussion forum until reviews converge ("Frontiers review system," 2012). One potential problem in implementing such forums is the possibility of increasing the already long delay between submission and editorial decisions. Particularly, conferences that have to follow strict timelines might have difficulty enforcing active participation of authors and reviewers over a short period.

Designing review forms for separation of concerns. We found that lower ratings accompanying reviews were a significant predictor of criticisms with no redressive actions. This suggests that reviewers tend to make unmitigated criticism more often when their perception of the quality of a paper is lower. One possible explanation for this behavior is that reviewers want to ensure that the tone of their review

matches their ratings so editors or program committees are not misled by positive comments in the review. Although reviews are aimed at helping editors and program committees in their decision-making, they can (and some argue they should) help authors to improve their work. Authors who receive negative reviews might benefit more than others from encouragement and politeness that could make them more receptive to suggestions made by reviewers. One possible solution is to structure review forms to separate suggestions for improvement from justifications of the rating and to make both visible to the authors. The section asking for suggestions for improvement aims at encouraging reviewers to focus on helping authors going forward, while the other parts of the form aim at conveying reviewers' judgments considering the current state of the work. Future studies are needed to assess the effectiveness of this and other review form modifications.

Open peer review venues for familiarizing newcomers with community norms. We found that higher use of criticisms with no redressive action, and lower use of off-record strategies (often used for achieving the highest level of mitigation) were associated with less experienced reviewers. We thus suggest that a research community might benefit from having at least one open peer reviewed venue to provide exposure to reviewing practices within the community. We were not able to compare our observations with reviewers' behavior in closed systems, to be certain about how well behavior in an open peer-review environment reflects common behavior in an anonymous setting; however, based on the literature, we expect that the publicly demonstrated behavior will reflect valued and appropriate behavior, or behavioral norms (Terry & Hogg, 1999, p. 97). Open peer review could thus facilitate newcomers becoming familiar with linguistic norms that allow for constructive feedback while avoiding insult. Interactions and conversations rely on the set of assumptions that interactants share (Watts, 2009). In research communities where all publication venues use anonymous closed peer review processes, there is little chance for newcomers, including new researchers, researchers in neighboring fields, and perhaps even amateur or citizen scientists (although they currently are little engaged in research tasks (Wiggins & Crowston, 2012)) to effectively understand, respond to, or compose peer reviews. The existence of open peer reviewed venues could also help non-native reviewers who are less familiar with strategies that rely

on language-specific and grammatical structures. Linguistic expertise is considered to be one of the factors that reduce early productivity because it gives more experienced members a competitive edge over others (K. Hyland, 1996). Although asking reviewers to be polite (“Guide to reviewing CHI papers and notes,” 2012) and providing guidelines and describing expectations can alleviate this effect, these cannot substitute for exposure to what actually happens in the reviewing community. We believe that exposure to reviews composed by others can accelerate the distribution of new or improved practices, and promote evolution of the process when this is needed. Persistence of reviews on the web or other archival repositories enables using them as self-learning material or for teaching peer review processes. It also facilitates creation of more effective instruction documents for reviewers that illustrate guidelines by pointing at published reviews. This could accelerate the distribution, refinement, and evolution of effective reviewing practices and related linguistic structures. Similarly, providing support for program committee members and editors to annotate reviews and to identify and share exemplars of constructive, polite criticism could facilitate and promote those practices and structures.

Alternative approach: Structured reviews. Efficiency of composing reviews is another goal of peer review support systems, but efficiency of reviewers could be in conflict with the communicative goals of reviews because thoughtful and perhaps time consuming effort is required to compose clear and polite reviews. An alternative approach to what we have suggested so far for enabling effective polite interactions between reviewers and authors is using structured reviews. Current systems often allow a large degree of flexibility for reviewers to express their opinion in any way that they want, and to frame them the way they prefer. This flexibility perhaps requires a great deal of linguistic skill to achieve the multiple goals of informing an editor or a program committee, while also conveying constructive criticisms to authors in a polite and effective manner. Providing a standard structure to reviewers, a structure designed to achieve those goals, might enable achieving clarity and politeness in an efficient way. Increasing the structure of scientific communication has been studied for decades and structured abstracts have shown to be of higher quality in comparison with unstructured ones (Budgen et al., 2008; Taddio et al., 1994). A structured review form could be an optional way to guide reviewers in shaping

their reviews, for example by explicitly eliciting opinion on both pros and cons of primary aspects of submissions. Such a review form could help reviewers to focus more on responding to what is asked for, instead of making them manage communicative and social aspects of their review. Assessment of effectiveness of this approach in comparison with alternative approaches that aim at helping authors better manage their communicative and social aspects of their review is another important avenue for future research.

Limitations

This study was designed to provide a snapshot of reviewers' behaviour in an open signed peer-review process conducted in a special track of a Human-Computer Interaction conference. Because of the small scale of this study, generalizability is a concern. Further observation in other communities and over time is needed to assess the extent similar behavior can be observed in other contexts and research communities. This study was conducted in a context where participants were accustomed to peer review processes that mask the identities of reviewers. We do not know if the findings will hold when reviewers get used to open processes. If future reviewers develop their reviewing voices when signed reviewing is typical, they may adopt an even more polite tone; on the other hand, if reviewers have habituated to this mode, the behavior may fade out. Future studies will need to examine this, especially if signed reviewing becomes the dominant mode for peer review.

This study only looked at a limited aspect of reviewers' behavior, and did not assess how authors respond to such an open peer-reviewing environment. Previous research has found that masking identity of authors could increase participation of female authors (Budden et al., 2008). An important avenue for future research is to assess the effect of using public peer-review on participation of various groups of authors. One possibility is that a public environment could help in assuring authors that are concerned about reviewers' bias, by bringing transparency to the process. However, only future studies can determine how different types of environments and different types of masking identities interact.

The primary reason that we used Brown and Levinson's taxonomy of politeness strategies was to enable identification of the diverse ways that reviewers mitigate criticisms. We believe the theory was

extremely helpful in this regard. We were only able to take limited advantage of separation of positive and negative politeness strategies in our analysis of implications for design. As critics of Brown and Levinson have mentioned, there is some fuzziness in the definitions of positive and negative politeness (Kerbrat-Orecchioni, 1997), and both types of strategies could be used to redress both negative and positive face. For example, a negative politeness strategy can be used to redress negative face by softening a suggestion for change, but it also can be used to soften a criticism. In addition, in the context of peer-review face threatening acts often threaten both negative and positive faces by mentioning a problem (threatening positive face) and asking for a change in the direction suggested by the reviewer (threatening negative face).

It was often challenging to understand if the goal of reviewers was to be polite, or just to convey a level of certainty about their statement or the severity of a problem. This was particularly challenging for negative politeness strategies. In all cases, we coded those instances as an occurrence of politeness strategies. To get a sense of how often these strategies were actually used as politeness devices, rather than to convey doubt, we looked at the relation between self-assessment of expertise and use of hedges that convey uncertainty. If hedges were used to convey uncertainty, it seems reasonable to expect more hedges when reviewers are less certain about their expertise, but we found no relation between the two: 20% of criticisms made by each of the groups of *passing knowledge*, *knowledgeable*, and *expert* used an uncertainty hedge (only 2 reviews were self-rated as *no knowledge*, which was the only other rating available). This suggests that there is no relation between confidence of reviewers in their knowledge and their use of uncertainty hedges.

The problem of differentiating literal usage from some of the politeness devices is a common challenge in studies that try to count the occurrence of politeness strategies, which is one reason to focus on drawing conclusions based on qualitative analysis rather than relying on quantitative analysis of the number of occurrences of strategies. Sometimes criticisms used an amalgamation of multiple strategies that was hard to untangle. Therefore some of these strategies might have gone unnoticed.

We used regression analysis to assess effects of various background variables such as experience, gender, type and country of affiliation on use of language. However, a common limitation of our analyses is that we were not able to control for several other variables that could affect use of language, such as English proficiency and cultural variables. This is particularly important because previous research has shown that culture, language, and language proficiency affects use of politeness strategies (Callahan, 2011; Holtgraves & Yang, 1992; Maier, 1992). In addition, we looked at effects of years of publishing experience of authors to assess if seniority affects reviewers' use of politeness strategies; however, years of reviewing experience might not accurately represent "academic capital" or a position of power. Sometimes young stars in research communities are able to build a reputation in a short period of time due to their high productivity. Also, researchers could build a reputation by offering a high level of service to the community without achieving a superior publication record. Future studies are needed that examine other measures of reputation and academic capital to be able to better understand their effects on reviewers' behavior.

Another limitation of our study was that the unit of analysis we employed might not capture some of the politeness strategies in some peer review structures. In our dataset, one reviewer used a special formatting that included bulleted lists of pros and cons, which made it unsuitable for our analysis by separating compliments from criticisms. That data was not considered in the quantitative analysis.

A final limitation is that our study was not able to directly compare open signed review with more traditional single- or double-blind reviewing to determine if politeness strategies were employed differently. Moreover, different research communities appear to have different norms for conveying criticism (Hadjistavropoulos & Bieling, 2000; Rose & Stevens, 1925). To our knowledge there has been no study on a comparable community that can be used for comparison. Although conducting a comparative study would have been desirable, the ethical considerations surrounding how this could be done precluded using any pre-existing corpus. Because the alt.chi 2007 corpus was publicly available and both authors and reviewers were fully identified, we were able to obtain fairly complete demographic information including (for example) the degree of authoring experience and associate this with particular

papers (and ratings). Were we to attempt to do this with an existing corpus (such as for the normal CHI papers) we either would not be able to identify participants or, if we did have comparable demographic information, we would compromise the guaranteed anonymity of the double-blind peer review process. This could not be done without obtaining prior consent from everyone involved, a daunting task at best and certainly one that would be next to impossible after the fact. We leave this for a future study, hoping that our current work will in fact provide some of the justification required to obtain the necessary ethical consent so that we might in fact carry out this unquestionably valuable comparison.

Conclusion

Computer-supported peer review systems have become increasingly popular over the last few decades, at first only to make reviewing requests and distribute reviews through email. Later, dedicated peer review support systems such as Manuscript Central and Editorial Manager, or conference management systems such as ConfMaster, CyberChair, EasyChair, and Precision Conference Solutions, were developed to conduct the entire process. Our research is a first step in understanding design options in this important and under-explored area. As Lessig (2000, para. 28) points out, systems that support processes are not neutral: “*The code regulates. It implements values, or not.*” Studies of use of technology in peer-review are needed to make sense of how reviewers and authors interact in the various existing peer review processes to inform the design of systems that support current and future peer review processes.

We identified high frequency and variety of politeness strategies in the alt.chi peer review corpus. While these strategies affect clarity of criticisms, they also provide a means for maintaining relationships and ensuring that authors will not get discouraged. It is important for two reasons to understand the writing styles, conventions, and strategies that are used to convey criticisms both effectively and politely. First, it is important that reviewers convey clearly their opinions about various aspects of submissions, while maintaining the desired level of politeness according to the norms of their community. Second, it is important that authors understand the language used by the reviewers to effectively take into account the

criticisms provided when revising their work, or responding to the reviews. Based on our findings, we proposed implications for the design of peer review support systems, and for research communities.

Different communities may have different attitudes toward politeness. It may not be a goal to be more polite. Excessive politeness could become problematic when necessary FTAs are avoided rather than mitigated (Rose & Stevens, 1925). Design of peer review support systems should be mindful of the norms of the research communities that will use the systems, so that design decisions related to anonymity, persistence of reviews, and visibility of reviews to other reviewers, to the research community, or to the general public are made based on the values and norms of the particular community being served.

References

- Abramowitz, S. I., Gomes, B., & Abramowitz, C. V. (1975). Publish or Politic: Referee Bias in Manuscript Review1. *Journal of Applied Social Psychology*, 5(3), 187–200. doi:10.1111/j.1559-1816.1975.tb00675.x
- Alt. CHI. (2007). ACM. Retrieved December 17, 2013, from <http://www.chi2007.org/submit/altchi.php>
- Armstrong, J. S. (1997). Peer review for journals: evidence on quality control, fairness, and innovation. *Science and Engineering Ethics*, 3(1), 63–84.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412. doi:10.1016/j.jml.2007.12.005
- Baggs, J., Broome, M., Dougherty, M., Freda, M., & Kearney, M. (2008). Blinding in peer review: the preferences of reviewers for nursing journals. *Journal of Advanced Nursing*, 64(2), 131–138. doi:10.1111/j.1365-2648.2008.04816.x
- Benos, D. J., Bashari, E., Chaves, J. M., Gaggar, A., Kapoor, N., LaFrance, M., ... Zotov, A. (2007). The ups and downs of peer review. *Advances in Physiology Education*, 31(2), 145–152. doi:10.1152/advan.00104.2006
- Bingham, C. M., Higgins, G., Coleman, R., & Van Der Weyden, M. B. (1998). The Medical Journal of Australia internet peer-review study. *The Lancet*, 352(9126), 441–445. doi:10.1016/S0140-6736(97)11510-0
- Blank, R. M. (1991). The Effects of Double-Blind versus Single-Blind Reviewing: Experimental Evidence from The American Economic Review. *The American Economic Review*, 81(5), 1041–1067. doi:10.2307/2006906
- Bornmann, L. (2011). Scientific peer review. *Annual Review of Information Science and Technology*, 45(1), 197–245. doi:10.1002/aris.2011.1440450112
- Bornmann, L., & Daniel, H.-D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45–80.

- Bornmann, L., & Daniel, H.-D. (2010). Reliability of reviewers' ratings when using public peer review: a case study. *Learned Publishing*, 23(2), 124–131. doi:10.1087/20100207
- Brooks, T. A. (1986). Evidence of complex citer motivations. *Journal of the American Society for Information Science*, 37(1), 34–36.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some Universals in Language Usage*. Cambridge University Press.
- Budden, A. E., Tregenza, T., Aarssen, L. W., Koricheva, J., Leimu, R., & Lortie, C. J. (2008). Double-blind review favours increased representation of female authors. *Trends in ecology & evolution*, 23(1), 4–6. doi:10.1016/j.tree.2007.07.008
- Budgen, D., Kitchenham, B. A., Charters, S. M., Turner, M., Brereton, P., & Linkman, S. G. (2008). Presenting software engineering results using structured abstracts: a randomised experiment. *Empirical Software Engineering*, 13(4), 435–468. doi:10.1007/s10664-008-9075-7
- Burke, M., & Kraut, R. (2008). Mind your Ps and Qs: the impact of politeness and rudeness in online communities. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work* (pp. 281–284). New York, NY, USA: ACM. doi:10.1145/1460563.1460609
- Callahan, L. (2011). Asking for a Letter of Recommendation in Spanish and English: A Pilot Study of Face Strategies. *Hispania*, 94(1), 171–183.
- Camacho-Miñano, M.-M., & Núñez-Nickel, M. (2009). The multilayered nature of reference selection. *Journal of the American Society for Information Science and Technology*, 60(4), 754–777. doi:10.1002/asi.21018
- Campanario, J. M. (1998). Peer Review for Journals as it Stands Today—Part 1 and 2. *Science Communication*, 19(3), 181–211. doi:10.1177/1075547098019003002
- Chubin, D. E., & Hackett, E. J. (1990). *Peerless science: peer review and U.S. science policy*. Albany, N.Y.: State University of New York Press.
- Cohoon, J. M., Nigai, S., & Kaye, J. “Jofish.” (2011). Gender and computing conference papers. *Communications of the ACM*, 54(8), 72. doi:10.1145/1978542.1978561

- Cutting, J. (2007). *Vague language explored*. Basingstoke, New York: Palgrave Macmillan.
- De Millo, R. A., Lipton, R. J., & Perlis, A. J. (1979). Social processes and proofs of theorems and programs. *Communications of the ACM*, 22(5), 271–280. doi:10.1145/359104.359106
- Eelen, G. (2001). *A Critique of Politeness Theories*. St. Jerome Pub.
- Ernst, E., & Resch, K. L. (1994). Reviewer bias: a blinded experimental study. *The Journal of Laboratory and Clinical Medicine*, 124(2), 178–182.
- Fortanet, I. (2008). Evaluative language in peer review referee reports. *Journal of English for Academic Purposes*, 7(1), 27–37. doi:10.1016/j.jeap.2008.02.004
- Frontiers review system. (2013). Frontiers Media. Retrieved December 17, 2013, from <http://www.frontiersin.org/about/reviewsystem>
- Geisler, C. A. (2003). *Analyzing Streams of Language: Twelve Steps to the Systematic Coding of Text, Talk, and Other Verbal Data*. Longman.
- Gilliland, S. W., & Beckstein, B. A. (1996). Procedural and distributive justice in the editorial review process. *Personnel Psychology*, 49(3), 669–691. doi:10.1111/j.1744-6570.1996.tb01589.x
- Godlee, F., Gale, C. R., & Martyn, C. N. (1998). Effect on the Quality of Peer Review of Blinding Reviewers and Asking Them to Sign Their Reports. *JAMA: The Journal of the American Medical Association*, 280(3), 237–240. doi:10.1001/jama.280.3.237
- Goffman, E. (1955). On face-work: an analysis of ritual elements in social interaction. *Psychiatry: Journal for the Study of Interpersonal Processes*, 18, 213–231.
- Goffman, E. (1956). The nature of deference and demeanor. *American Anthropologist*, 58(3), 473–502. doi:10.1525/aa.1956.58.3.02a00070
- Goffman, E. (1971). The territories of the self. *Relations in public*, 28–61.
- Graham, S. L. (2007). Disagreeing to agree: Conflict, (im)politeness and identity in a computer-mediated community. *Journal of Pragmatics*, 39(4), 742–759. doi:10.1016/j.pragma.2006.11.017
- Guide to reviewing CHI papers and notes. (2012). *CHI 2012*. Retrieved December 17, 2013, from <http://chi2012.acm.org/cfp-reviewers-guide.shtml>

- Hadjistavropoulos, T., & Bieling, P. J. (2000). When reviews attack: ethics, free speech, and the peer review process. *Canadian Psychology = Psychologie Canadienne*, 41(3), 152–159.
- Harwood, N. (2009). An interview-based study of the functions of citations in academic writing across two disciplines. *Journal of Pragmatics*, 41(3), 497–518. doi:10.1016/j.pragma.2008.06.001
- Herbert, R. K. (1989). The ethnography of English compliments and compliment responses: A contrastive sketch. In W. Oleksy (Ed.), *Contrastive pragmatics* (pp. 3–35). Amsterdam, The Netherlands: Benjamins.
- Hobbs, P. (2003). The medium is the message: politeness strategies in men's and women's voice mail messages. *Journal of Pragmatics*, 35(2), 243–262. doi:10.1016/S0378-2166(02)00100-5
- Holmes, J. (1986). Compliments and Compliment Responses in New Zealand English. *Anthropological Linguistics*, 28(4), 485–508. doi:10.2307/30028355
- Holmes, J. (1988). Paying compliments: A sex-preferential politeness strategy. *Journal of Pragmatics*, 12(4), 445–465. doi:10.1016/0378-2166(88)90005-7
- Holtgraves, T., & Yang, J. (1992). Interpersonal underpinnings of request strategies: General principles and differences due to culture and gender. *Journal of Personality and Social Psychology*, 62(2), 246–256. doi:10.1037/0022-3514.62.2.246
- Hübler, A. (1983). *Understatements and Hedges in English*. John Benjamins Publishing.
- Hyland, F., & Hyland, K. (2001). Sugaring the pill: Praise and criticism in written feedback. *Journal of Second Language Writing*, 10(3), 185–212. doi:10.1016/S1060-3743(01)00038-8
- Hyland, K. (1996). Talking to the academy: forms of hedging in science research articles. *Written Communication*, 13(2), 251–281. doi:10.1177/0741088396013002004
- Isenberg, S. J., Sanchez, E., & Zafran, K. C. (2009). The effect of masking manuscripts for the peer-review process of an ophthalmic journal. *British Journal of Ophthalmology*, 93(7), 881–884. doi:10.1136/bjo.2008.151886
- Johnson, D. M. (1992). Compliments and politeness in peer-review texts. *Applied Linguistics*, 13(1), 51–71. doi:10.1093/applin/13.1.51

- Justice AC, C. M. (1998). Does masking author identity improve peer review quality?: A randomized controlled trial. *JAMA: The Journal of the American Medical Association*, 280(3), 240–242.
- Kerbrat-Orecchioni, C. (1997). A multilevel approach in the study of talk-in-interaction. *Pragmatics*, 7(1), 1–20.
- Khan, K. (2010). Is open peer review the fairest system? No. *BMJ*, 341(nov16 2), c6425–c6425. doi:10.1136/bmj.c6425
- Kourilova, M. (1998). Communicative characteristics of reviews of scientific papers written by non-native users of English. *Endocrine Regulations*, 32, 107–114.
- Lakoff, G. (1973). Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic*, 2(4), 458–508.
- Lakoff, R. T. (1990). *Talking power: the politics of language in our lives*. Basic Books.
- Lakoff, R. T., & Ide, S. (2005). *Broadening the Horizon of Linguistic Politeness*. John Benjamins Publishing Company.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174. doi:10.2307/2529310
- 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
- Lessig, L. (2000). Code is law: on liberty in cyberspace. *Harvard Magazine*, (Jan-Feb). Retrieved December 17, 2013, from <http://harvardmagazine.com/2000/01/code-is-law-html>
- Lind, E. A., & Tyler, T. R. (1988). *The Social Psychology of Procedural Justice*. Springer.
- MacGeorge, E. L., Lichtman, R. M., & Pressey, L. C. (2002). The evaluation of advice in supportive interactions: facework and contextual factors. *Human Communication Research*, 28(3), 451–463. doi:10.1111/j.1468-2958.2002.tb00815.x
- Mackiewicz, J., & Riley, K. (2003). The technical editor as diplomat: linguistic strategies for balancing clarity and politeness. *Technical Communication*, 50(1), 83–94.

- Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory bias in the peer review system. *Cognitive Therapy and Research*, 1(2), 161–175. doi:10.1007/BF01173636
- Maier, P. (1992). Politeness strategies in business letters by native and non-native English speakers. *English for Specific Purposes*, 11(3), 189–205. doi:10.1016/S0889-4906(05)80009-2
- McNutt, R. A., Evans, A. T., Fletcher, R. H., & Fletcher, S. W. (1990). The effects of blinding on the quality of peer review. A randomized trial. *JAMA: the journal of the American Medical Association*, 263(10), 1371–1376.
- Melero, R., & López-Santoveña, F. (2001). Referees' attitudes toward open peer review and electronic transmission of papers. *Food Science and Technology International*, 7(6), 521–527. doi:10.1106/0MXD-YM6F-3LM6-G9EB
- Merton, R. K. (1973). *The Sociology of Science: Theoretical and Empirical Investigations*. University of Chicago Press.
- Miller, D. T. (2001). Disrespect and the experience of injustice. *Annual Review of Psychology*, 52, 527–553. doi:10.1146/annurev.psych.52.1.527
- Mulligan, A., Hall, L., & Raphael, E. (2012). Peer review in a changing world: An international study measuring the attitudes of researchers. *Journal of the American Society for Information Science and Technology*, 64(1), 132–161. doi:10.1002/asi.22798
- Myers, G. (1989). The pragmatics of politeness in scientific articles. *Applied Linguistics*, 10(1), 1–35. doi:10.1093/applin/10.1.1
- Myers, G. (1991). Politeness and certainty: the language of collaboration in an ai project. *Social Studies of Science*, 21(1), 37–73.
- Nature's trial of open peer review. (2006, December). *Nature*. Retrieved December 17, 2013, from <http://www.nature.com/nature/peerreview/debate/nature05535.html>
- Nicholls, R. D. (1999). Peer review under review. *Science*, 286(5446), 1853–1853. doi:10.1126/science.286.5446.1853b

- Peters, D. P., & Ceci, S. J. (1982). Peer-Review Practices of Psychological Journals: The Fate of Published Articles, Submitted Again. *Behavioral and Brain Sciences*, 5(02), 187–195.
doi:10.1017/S0140525X00011183
- Pulverer, B. (2010). A transparent black box. *The EMBO Journal*, 29(23), 3891–3892.
doi:10.1038/emboj.2010.307
- Riggs, J. E. (1995). Priority, Rivalry, and Peer Review. *Journal of Child Neurology*, 10(3), 255–256.
doi:10.1177/088307389501000325
- Rose, D. H., & Stevens, N. E. (1925). The excessive politeness of american botanists. *Science*, 61(1591), 656–657. doi:10.1126/science.61.1591.656
- Ross, J. S., Gross, C. P., Desai, M. M., Hong, Y., Grant, A. O., Daniels, S. R., ... Krumholz, H. M. (2006). Effect of blinded peer review on abstract acceptance. *JAMA: the journal of the American Medical Association*, 295(14), 1675–1680. doi:10.1001/jama.295.14.1675
- Salager-Meyer, F., Alcaraz Ariza, M. Á., & Pabón Berbesí, M. (2007). Collegiality, critique and the construction of scientific argumentation in medical book reviews: A diachronic approach. *Journal of Pragmatics*, 39(10), 1758–1774. doi:10.1016/j.pragma.2006.06.003
- Shatz, D. (2004). *Peer review: a critical inquiry*. Rowman & Littlefield.
- Smith, R. (1999). Opening up BMJ peer review. *BMJ*, 318(7175), 4–5.
- Smith, R. (2006). Peer review: a flawed process at the heart of science and journals. *JRSM*, 99(4), 178–182. doi:10.1258/jrsm.99.4.178
- Souder, L. (2011). The ethics of scholarly peer review: a review of the literature. *Learned Publishing*, 24(1), 55–72. doi:10.1087/20110109
- Sternberg, R. J. (2002). On civility in reviewing. *APS Observer*, 15(1), 3.
- Sugimoto, C. R., & Cronin, B. (2013). Citation gamesmanship: testing for evidence of ego bias in peer review. *Scientometrics*, 95(3), 851–862. doi:10.1007/s11192-012-0845-z
- Taddio, A., Pain, T., Fassos, F. F., Boon, H., Ilersich, A. L., & Einarson, T. R. (1994). Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal,

- the Canadian Medical Association Journal and the Journal of the American Medical Association. *Canadian Medical Association Journal*, 150(10), 1611–1615.
- Tannen, D. (1998). *Talking from 9 to 5: Women and Men in the Workplace : Language, Sex and Power*. Virago Press.
- Terry, D. J., & Hogg, M. A. (1999). *Attitudes, Behavior, and Social Context: The Role of Norms and Group Membership*. Psychology Press.
- Tracy, K. (2008). Reasonable hostility: Situation-appropriate face attack. *Journal of Politeness Research*, 4(2), 169–191.
- Van Rooyen, S., Delamothe, T., & Evans, S. J. W. (2010). Effect on peer review of telling reviewers that their signed reviews might be posted on the web: randomised controlled trial. *BMJ*, 341(7782), c5729–c5729. doi:10.1136/bmj.c5729
- Van Rooyen, S., Godlee, F., Evans, S., Black, N., & Smith, R. (1999). Effect of open peer review on quality of reviews and on reviewers' recommendations: a randomised trial. *BMJ*, 318(7175), 23 – 27.
- Van Rooyen, S., Godlee, F., Evans, S., Smith, R., & Black, N. (1998). Effect of Blinding and Unmasking on the Quality of Peer Review. *JAMA: The Journal of the American Medical Association*, 280(3), 234 –237. doi:10.1001/jama.280.3.234
- Vries, H. D., Elliott, M. N., Kanouse, D. E., & Teleki, S. S. (2008). Using pooled kappa to summarize interrater agreement across many items. *Field Methods*, 20(3), 272–282. doi:10.1177/1525822X08317166
- Walsh, E., Rooney, M., Appleby, L., & Wilkinson, G. (2000). Open peer review: a randomised controlled trial. *The British Journal of Psychiatry*, 176(1), 47–51. doi:10.1192/bjp.176.1.47
- Ware, M., & Monkman, M. (2008). *Peer review in scholarly journals: perspective of the scholarly community—an international study*. Publishing Research Consortium.
- Watts, R. J. (2003). *Politeness*. Cambridge University Press.

- Watts, R. J. (2009). Relevance and relational work: linguistic politeness as politic behavior. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, 8(2-3), 131–166.
- Wiggins, A., & Crowston, K. (2012). Goals and Tasks: Two Typologies of Citizen Science Projects. In *Hawaii International Conference on System Science (HICSS)* (pp. 3426 –3435). Maui, Hawaii: IEEE Computer Society. doi:10.1109/HICSS.2012.295
- Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In *Applications of social research methods to questions in information and library science* (pp. 308–319). Westport, CT: Libraries Unlimited.