

Closed versus open reviewing of journal manuscripts: how far do comments differ in language use?

Lutz Bornmann, Markus Wolf, Hans-Dieter Daniel

Corresponding and first author:

Lutz Bornmann
Max Planck Society
Administrative Headquarters
Hofgartenstr. 8
80539 Munich
Tel. +49-(0)89-2108-1265
E-mail: bornmann@gv.mpg.de

Second author:

Markus Wolf
Center for Psychotherapy Research
University Hospital Heidelberg
Bergheimer Str. 54
69115 Heidelberg
Germany
Tel. +49-(0)6221-567381
E-mail: markus.wolf@med.uni-heidelberg.de

Third author:

Hans-Dieter Daniel
Evaluation Office, University of Zurich and
Professorship for Social Psychology and Research on Higher Education, ETH Zurich
Mühlegasse 21, MUG
8001 Zurich

Abstract

Whereas in traditional, closed peer review (CPR) a few, selected scientists (peers) are included in the process of manuscript review, public peer review (PPR) includes, in addition to invited reviewers, a wider circle of scientists who are interested in a manuscript and wish to write a comment on it. In this study, using the data of two comprehensive evaluation studies on the CPR process at *Angewandte Chemie International Edition* and the PPR process at *Atmospheric Chemistry and Physics*, we examined the language characteristics in comments that were written by invited reviewers in CPR and by invited reviewers *and* interested members of the scientific community in PPR. We used Linguistic Inquiry and Word Count (LIWC), a text analysis software program that counts words in meaningful categories (e.g., positive or negative emotions) using a standardized dictionary. We examined 599 comments from the reviews of 229 manuscripts. The results show that the comments in PPR are much longer than the comments in CPR. This is an indication that PPR reviewing has more of an improvement function and CPR reviewing has more of a selection function. The results also show that CPR is not, as might be expected, more susceptible to the expression of negative emotions than PPR is. On the contrary, positive emotion words are used statistically significantly more frequently in CPR than in PPR.

Keywords

Journal peer review, Text analysis, Linguistic Inquiry and Word Count, *Angewandte Chemie International Edition*, *Atmospheric Chemistry and Physics*

Introduction

Whereas in traditional, closed peer review (CPR) a few, selected scientists (peers) are included in the process of manuscript review, public peer review (PPR) includes, in addition to invited reviewers, a wider circle of scientists who are interested in a manuscript and wish to write a comment on it. The aim with PPR is to initiate specialist discussions on individual scientific contributions – discussions that are possible only to a limited extent in CPR.

The advantage of PPR is seen in the fact that by using comments by interested members of the scientific community, more valid decisions on whether to publish manuscripts can be expected, and manuscripts accepted for publication can be even more improved prior to publication than with traditional CPR, in which only reviewer comments are used. PPR is supposed to bring a new openness to the reviewing process that will enhance its accuracy and fairness (Bingham, Higgins, Coleman, & Van Der Weyden, 1998; Koonin & Lipman, 2006). However, in addition to the benefits of PPR, some drawbacks have been pointed out. For instance, Harnad (2000) raised the question as to whether it can really be expected that the self-appointed commentators will be qualified specialists: “The expert population in any given speciality is a scarce resource, already overharvested by classical peer review, so one wonders who would have the time or inclination to add journeyman commentary services to this load on their own initiative” (Harnad, 2000, p. 9). The discussion on PPR became intensified in recent years due to the results of a field trial at the journal *Nature* that tested the use of this new form of review (Anon, 2006). In the PPR trial launched by *Nature*, authors whose manuscripts were selected for traditional closed peer review could opt to have their manuscripts posted to an open website for public comment. *Nature* rated the results of the trial as hardly encouraging. Very few authors opted to have their manuscripts hosted for public comment, and their manuscripts received very few comments.

In this study we conducted a comparative text analysis of comments that were written in the context of CPR and PPR. Here, and for the first time in peer review research, we worked with a text analysis software program that counts words in meaningful categories using a standardized dictionary: Linguistic Inquiry and Word Count (LIWC) (Pennebaker, Chung, Ireland, Gonzales, & Booth, 2007). The program was used earlier by Hartley, Pennebaker, and Fox (2003) to assess the Abstracts, Introductions, and Discussions in 80 journal articles in the field of educational psychology. The few content analyses that have been conducted up to now in the area of peer review worked with category

systems that were developed inductively from the analyzed text material and specifically for the area of peer review (see here Bornmann, Nast, & Daniel, 2008; Bornmann, Weymuth, & Daniel, 2010; Mungra & Webber, 2010). In those studies, several coders assigned the text passages to the categories. The advantage of using LIWC as opposed to inductively or specifically developed category systems is that the content analysis is done by the program automatically and the result is free of subjective influences due to coders (Pennebaker, Mehl, & Niederhoffer, 2003; Popping, 2000). Further, the use of a standardized dictionary for the word analysis makes possible direct comparison of the results of different text analyses.

In this study we compared three types of texts from journal peer review. For one, we examined agreement and difference between comments that were written by invited reviewers (ACP RC) and interested members of the scientific community (ACP SC) on manuscripts that had been submitted to the journal *Atmospheric Chemistry and Physics* (ACP). ACP has a two-stage publication process, including PPR and interactive discussion (Pöschl, 2004). For another, we examined agreement and difference between reviewer comments that were written in the context of PPR at ACP (ACP RC) and in CPR at the journal *Angewandte Chemie – International Edition* (AC-IE) (AC-IE RC). As this study was an exploratory one, we did not test specific hypotheses when comparing these three types of texts. The aim of the study was to investigate the extent to which language use in peer reviews varies as a function of the mode of the peer review process. To do so, the ACP RCs, ACP SCs, and AC-IE RCs were selected for the analysis as texts that were written for a similar purpose (for evaluative purposes) but can be expected to show differences in word usage.

The data used in this study were compiled for two studies on journal peer review: One study examined the CPR process at AC-IE (Bornmann & Daniel, 2008a, 2008b, 2009a, 2009b, 2010a, 2010c; Bornmann et al., 2009), and the other examined the PPR process at ACP (Bornmann & Daniel, 2010b; Bornmann, Marx, Schier, Thor, & Daniel, 2010; Bornmann, Neuhaus, & Daniel, 2011; Bornmann, Schier, Marx, & Daniel, 2011).

Methods

Manuscript review at ACP

ACP was launched in September 2001. It is published by the European Geosciences Union (EGU; <http://www.egu.eu>) and Copernicus Publications (<http://publications.copernicus.org/>). ACP is freely accessible via the Internet (www.atmos-chem-phys.org). The ACP two-stage publication process is described at the ACP website as follows: In the first stage, manuscripts that pass a rapid pre-screening process (access review) are immediately published as ‘discussion papers’ on the journal’s *Atmospheric Chemistry and Physics Discussions* (ACPD) website: “The access review (pre-selection by the editor with optional advice from referees) is meant to avoid a potential overload of the discussion forum with papers that are clearly deficient or out of scope” (Pöschl, 2004, p. 107).

The discussion papers are then made available for ‘interactive public discussion,’ during which the comments of designated reviewers (ACP RC, usually, the reviewers that already conducted the access review) (anonymous or attributed), additional comments by other interested members of the scientific community (ACP SC) (attributed), and the authors’ replies are published alongside the discussion paper (Pöschl, 2010). Pöschl (2004) explained:

The interactive comments are published without peer review and revision, but can be censored in case of abusive commenting (personal offence, etc.) ... To ensure publication precedence for authors and to provide a lasting record of scientific discussion, the discussion papers and interactive comments are permanently archived and fully citable. (P. 107)

According to Pöschl (2010):

The interactive peer review and public discussion offer direct feedback and public recognition for high-quality papers (authors’ advantage); they prevent or minimize the opportunity for hidden obstruction and plagiarism (authors’ advantage); they provide complete and citable documentation of critical comments, controversial arguments, scientific flaws and complementary information (referees’ and readers’ advantage); they reveal deficiencies and deter submissions of carelessly prepared manuscripts, thus helping to avoid/minimize the waste of time and effort for deficient submissions (referees’, editors’, publishers’ and readers’ advantage). (p. 296)

Based on the revised manuscript (revised on the basis of ACP RCs and ACP SCs), and in view of the access peer review and interactive public discussion, the editor accepts or rejects the revised manuscript for publication in ACP. For this decision, further external reviewers may be asked to review the revision, if needed.

Manuscript review at AC-IE

A manuscript submitted to AC-IE is usually subject to internal and external review. First, editors at the journal evaluate whether the manuscript contributes to the development of an important area of research (internal review). If the editors find that it does, the submitted manuscript is sent to several independent reviewers (external review), who review it using a fully structured evaluation form together with a separate sheet for comments. The evaluation form contains a set of five questions and associated response categories, e.g., “How important do you consider the results reported?” (response scale: “very important”; “important”; “less important”; “unimportant”). On the separate comment sheet, the reviewers are to provide their answers to the questions on the evaluation form as free text, or they can indicate other relevant aspects for the decision on a manuscript by the AC-IE editors.

On the basis of the reviews of the external reviewers and on their own evaluations, the journal editors make – as a conclusion to the peer review process – the decision to accept or reject a manuscript for publication in AC-IE.

Database for the present study

For the investigation of peer review at ACP we had data for 1,111 manuscripts that went through the complete ACP selection process in the years 2001 to 2006. These manuscripts reached one of the following final statuses: 958 (86%) were published in ACPD and ACP, 74 (7%) were published in ACPD but not in ACP (here, the editor rejected the revised manuscript), and 79 (7%) were not published in either ACPD or ACP (these manuscripts were rejected during the access review). Some of the manuscripts ($n = 38$) submitted to ACP but not published by ACP (because they were rejected during the access review, for example) were submitted by the authors to another journal and published there. The search for the fate of the rejected manuscripts is described in Bornmann, Marx, et al. (2010).

For the text analysis of this study, first we selected the 119 manuscripts that as ACPD papers had been discussed interactively and for which there was at least one ACP SC and one ACP RC in the data set. For these 119 manuscripts a total of 385 comments were available. Six manuscripts were excluded because they had comments shorter than 50 words, which we consider too short for reliable

text analyses. Finally, a total of 362 comments were analyzed (223 ACP RCs and 139 ACP SCs) that had been prepared for 113 manuscripts (there are 3.2 comments on average for each manuscript). Therefore, 11% of the total 1,032 manuscripts that were published in ACPD were included in this study.

For the evaluation of the AC-IE peer reviews, information on a total of 1,899 manuscripts that were reviewed in the year 2000 was used as the database. The information was taken from archive material that was recorded in the Wiley-VCH publishing house. In addition to internal assessment by the publisher, the 1,899 manuscripts received a total of 4,593 reviews on an evaluation form and/or comment sheet. Based on the reviews, 46% ($n = 878$) of the 1,899 manuscripts were accepted for publication in AC-IE, and 54% ($n = 1021$) were rejected. As a search in the Web of Science (Thomson Reuters) and Chemical Abstracts (CA, Chemical Abstracts Service, CAS, Columbus, OH) literature databases revealed, 959 (94%) of the 1,021 rejected manuscripts were published more or less unchanged in another journal. A detailed description of the search for the rejected manuscripts can be found in the supporting information for Bornmann and Daniel (2008a).

For the text analysis of this study, first we selected 1,874 of the total 1,899 submissions for which at least one reviewer comment was available. There was no comment available for 25 manuscripts, mainly because the reviewers filled out the evaluation form only and wrote no comment. To obtain for the text analysis of the AC-IE manuscripts a similar number of papers as for the text analysis of the ACP manuscripts, from the 1,874 manuscripts we randomly selected 119 manuscripts for which 270 comments had been written. Two papers had to be excluded because the reviews were written in a language other than English. After exclusion of comments that were too short for text analysis (< 50 words), 116 manuscripts with 237 comments remained for the analysis.

Text preparation and description of the language categories

All comments were transformed to ANSI text format, spell checked, and edited following the recommendations in the LIWC manual (Pennebaker, et al., 2007). Parts of the documents that did not belong to the actual comment were removed (e.g., titles or references to the paper under review).

Due to the exploratory nature of this study, we selected a large subset of potentially relevant LIWC categories for the analysis of the comments. First, we wanted to give an in-depth description of

the basic linguistic characteristics of the three different types of comments by using the LIWC's linguistic processes categories: word count (i.e., the number of words in a given comment), words per sentence, long words (i.e., number of words longer than six letters), and usage of several kinds of function words (e.g., personal and impersonal pronouns, articles), common verbs and tenses, adverbs, prepositions, conjunctions, negations, quantifiers, and numbers. As basic formal indicators we assessed the use of question marks, exclamation marks, quotation marks, and parentheses in the comments. In sum, these language features provide important information about the cognitive style and complexity of the language used in the comments (Chung & Pennebaker, 2007).

Second and most interestingly, the LIWC has word categories that tap the words in a given text that connote *psychological processes*, i.e., affective processes, cognitive processes, social processes, perceptual processes, and biological processes. Table 1 shows examples of words in each of the categories. These categories have been widely applied in other research domains, such as social, health and personality psychology (for a review, see Pennebaker, et al., 2003; Tausczik & Pennebaker, 2010). They have been proven valid and found useful in the identification of words connoting psychological processes and have been found to be associated with individual differences and the writers' personality characteristics (Fast & Funder, 2008; Pennebaker & King, 1999).

Affective processes refer to the use of emotion words or words that convey emotional tone. Affective processes are further classified into two sub-categories, positive emotion and negative emotion. The categories have been shown to be a valid measure for measuring verbal expression of emotion (Kahn, Tobin, Massey, & Anderson, 2007). *Cognitive process* is a broad category that sums up all kinds of words in the text that indicate cognitive activity. It is based on the sub-categories insight, causation, discrepancy, tentative, certainty, inhibition, inclusive, and exclusive, all of which were used in the present study. *Social processes* include a range of different words in a given text that are social references. The category includes all non-first-person singular pronouns, verbs that are associated with human interaction, and nouns that refer to family, friends, or human beings. *Perceptual processes* include words that refer to various sensory and perceptual dimensions associated with the five senses (seeing, hearing, feeling, etc.). Finally, the category *biological processes* summarizes words that are related to the physical body, health, sexuality, or ingestion.

In addition to these LIWC standard categories, we created a simple custom category called *citations*, which assesses citations or references to other research. We replaced the citations and/or

references in the text by an annotation of the instances in which the category was used. The custom category was assessed using the LIWC program.

Table 1 (category “dictionary words”) shows the proportion of words in the ACP RCs, ACP SCs, and AC-IE RCs that could be assigned to LIWC categories. In the comments analyzed here, the proportion was 70%, which is somewhat less than the percentage in texts from other areas, such as in “emotional writing” or “novels,” where the LIWC captures over 80% of the words used (Pennebaker, et al., 2007). But the percentage found here is still higher than the percentage found for articles in the journal *Science* (54%, see Pennebaker, et al., 2007) and journal articles in educational psychology (60%, see Hartley, et al., 2003). Regarding the mean LIWC hit rate, there are no differences between the three groups, ACP RC, ACP SC, and AC-IE RC.

Statistical procedure

To test the three groups of texts (ACP RCs, ACP SCs, and AC-IE RCs) with regard to statistically significant differences between the mean values in the different LIWC categories, we computed one-way ANOVA (StataCorp., 2011). Using the regress-option in Stata (StataCorp., 2011), in pairwise comparisons we tested statistically significant differences between: (1) ACP RCs and ACP SCs, and (2) ACP RCs and AC-IE RCs. The significance level was set to 5%. In all, the mean values for 46 categories were compared. As two pairwise comparisons were conducted per category, there were 92 tests of significance. Due to the large number of tests conducted, the results section below reports and interprets mainly highly significant p values ($< .01$).

Results

The results of the text analysis using LIWC are presented in Table 1. The table shows mean numbers and proportions of words per LIWC category in ACP RCs and in ACP SCs for manuscripts submitted to ACP and in reviewer comments for manuscripts submitted to AC-IE (AC-IE RCs). Statistically highly significant differences between two mean numbers or proportions are shown in bold typeface (and are marked with two or three asterisks).

A value of 447.31 (mean number) for the category “word count” in column “ACP SC” in Table 1 shows that on average, an ACP SC contained 447 words. A value of 5.63 (mean proportion) for

the category “total pronouns” in column “ACP SC” means that on average, approximately 6% of the words in an ACP SC were pronouns (see here Tausczik & Pennebaker, 2010).

Insert Table 1 about here

Linguistic processes

The upper part of Table 1 shows the mean values for the basic linguistic processes. As the results of the word counts show, ACP RCs (795.26) were longer than ACP SCs (447.31) on average. The AC-IE RCs were comparatively short (200.39) on average; their length is about the length of abstracts found in scientific papers. Hartley and Betts (2009) examined 100 traditional abstracts published in 53 different social science journals and found an average length of 150 words. Table 1 shows further that there were on average 22.42 (AC-IE RC) to 24.51 (ACP RC) words in the sentences in the comments, which is about the same as the sentence length found in journal articles in educational psychology (Hartley, et al., 2003). When comparing ACP SCs, ACP RCs, and AC-IE RCs, sentences written by reviewers (ACP RCs: 24.51 and AC-IE RCs: 22.42) were shorter than sentences written by interested members of the community (ACP SCs: 31.0).

About 5.6% of the words in the comments analyzed in this study were pronouns. In comparison with base rates of word usage (Pennebaker, et al., 2007), this percentage is higher than that found for journal articles in educational psychology: In journal articles in educational psychology less than 2% of the words were pronouns (Hartley, et al., 2003). When comparing pronouns in ACP SCs, ACP RCs, and AC-IE RCs, the first person plural was used statistically significantly more frequently in ACP SCs (0.32) than in ACP RCs (0.06); no statistically significant difference was found between ACP RCs (0.06) and AC-IE RCs (0.04). This finding indicates that the writers of ACP SCs mention their own research more frequently than the writers of reviewer comments (ACP RCs and AC-IE RCs) do. Another statistically significant difference was found for the second person: This personal pronoun was used more frequently by writers of ACP RCs (0.18) than by authors of AC-IE RCs (0.01). With a value of 0.19, second person pronouns were used in the ACP SCs with approximately the same frequency as in the ACP RCs. In PPR the writers of comments evidently address the author of a manuscript directly more frequently than is the case in CPR. This difference can be interpreted to the

effect that in PPR, a discourse develops between authors, reviewers, and interested members of the community – which is precisely what this new form of peer review is intended to stimulate.

Words connoting psychological processes

Fraser and Martin (2009) examined “twelve established biomedical and fundamental clinical and clinical research journals over a twenty year time period for adjectives which modified an otherwise neutral knowledge claim” (p. 2) and found an increasing incidence of adjectives expressing subjective positive judgments in the papers (e.g., new, crucial, innovative, novel, central, unique, powerful) – mainly in papers that were published in reputable journals. The language of science should, in fact, be objective and should place data in the appropriate context. According to Tausczik and Pennebaker (2010), “research suggests that LIWC accurately identifies emotion in language use. For example, positive emotion words (e.g., love, nice, sweet) are used in writing about a positive event and more negative emotion words (e.g., hurt, ugly, nasty) are used in writing about a negative event” (p. 32).

Table 1 shows the percentages for positive and negative emotion words in ACP RCs, ACP SCs, and AC-IE RCs. Compared to texts in other areas, such as “emotional writing” or “novels” (Pennebaker, et al., 2007), emotion words were not particularly more or less frequent in the peer review comments. As the figures in the table show further, the percentage of positive emotions in the ACP RCs (2.16), ACP SCs (2.09), and AC-IE RCs (2.64) was higher than the percentage of negative emotions (ACP RCs: 0.96, ACP SCs: 1.0, and AC-IE RCs: 0.87). As the comments were written mainly for evaluative purposes, consisting in critical examination of manuscripts, we expected to find the opposite result. Content analyses in other peer review research studies showed that reviews contained clearly more negative comments than positive comments (Bakanic, McPhail, & Simon, 1989; Bornmann, Weymuth, et al., 2010). In line with these findings, we found that negations scored higher and assents scored lower in CPR as compared to PPR. However, there were no statistically significant differences – as we expected – for negative emotions. Instead we found a significant difference for positive emotions: More positive emotion words were used in CPR (AC-IE RC: 2.64) than in PPR (ACP RC: 2.16).

Tausczik and Pennebaker (2010) explained regarding the complexity of a text:

Prepositions (e.g., to, with, above), cognitive mechanisms (e.g., cause, know, ought), and words greater than six letters are all also indicative of more complex language. Prepositions,

for example, signal that the speaker is providing more complex and, often, concrete information about a topic. (p. 35)

As more complex language can be assumed for scientific texts, the percentage of prepositions and words longer than six letters in the comments analyzed here and the percentage of words that can be assigned to the category cognitive processes was relatively high. Compared to texts from non-scientific areas, such as “emotional writing” or “novels” (e.g., emotional writing or novels, see here Pennebaker, et al., 2007), it is noticeable that in the comments analyzed here and in *Science* articles (Pennebaker, et al., 2007) the percentage of words longer than six letters was especially high. Moreover, similar to articles in the journal *Science* (Pennebaker, et al., 2007), the proportions of numbers and quantifiers in the comments were high as well. Thus, compared with other text genres, the large proportions of these categories might be viewed as a proxy for the strong natural scientific background of the three journals. Comparing the three groups, we found the largest proportions of numbers and quantifiers in the ACP RCs and the lowest proportion in ACP SCs.

A look at the LIWC categories assigned to the cognitive processes reveals that there were interesting differences mainly for the category “Tentative”: “When people are uncertain or insecure about their topic they use tentative language (e.g., maybe, perhaps, guess)” (Tausczik & Pennebaker, 2010, p. 36). Words belonging in this category were more frequent in PPR (ACP RC: 2.56, ACP SC: 2.31) than in CPR (AC-IE RC: 1.97). Apparently, reviewers in open peer review phrase their points somewhat more cautiously than reviewers in closed peer review do. This view is supported by the pattern of basic syntactical features found in the reviews. Whereas exclamation marks were less frequent, question marks were used more frequently in PPR as compared to CPR, which is an indication of a more tentative writing style in PPR.

Social processes play a very small role in ACP RCs, ACP SCs, and AC-IE RCs, when we compare the numbers in this study with base rates of word usage in other, non-scientific areas, such as “emotional writing” or “novels” (Pennebaker, et al., 2007). As we expected, the writers of comments focused on the content of the research and not on the persons that conducted the research.

Finally, the results indicate that in PPR, statements are more often backed up by references to publications than in CPR. When comparing the proportions of citations in the comments, it became apparent that reviewers in PPR cited other research much more often to prove their points of view, especially the reviewers in ACP SC (but here the citations can be mainly references to the comment

writer's own publications; see above). The differences in the use of quotation marks seem to further support this finding.

Discussion

In this study, using the data of two comprehensive evaluation studies on the CPR process used by AC-IE and the PPR process used by ACP, we examined whether comments written in the traditional peer review process differ from comments written in a modern, open peer review process. More generally, we were interested in discovering the extent to which the LIWC program can be used for automated text analyses of the comments that are written in the peer review process. Using the comprehensive set of LIWC categories, we compared three types of comments (ACP RC, ACP SC, and AC-IE RC) with regard to the language that was used by the reviewers. All in all, the results on the LIWC categories yielded interesting insights on word usage in comments in CPR and PPR that for the most part conformed to expectations. For instance, the results indicate that in PPR, a discourse develops among authors, reviewers, and interested members of the community that is possible only to a limited extent in CPR. The results on the use of emotion words in peer review were very interesting. CPR is not, as one might expect, more prone to negative emotions than PPR. On the contrary, positive emotion words are used in CPR significantly more frequently than in PPR.

To summarize, the language patterns indicate that PPR and ACP SC adds a new quality to the peer review process. In our study, the language use in PPR was characterized by a more open and tentative as well as cognitively elaborated style as compared to CPR. This might be due to the fact that the reviews are open to the interested audience and thus themselves subject to evaluation. This might lead to a more careful and elaborated peer review process. Of course, based on these stylistic analyses no inferences concerning the quality of the reviews can be made. Future research should investigate whether these language patterns are indicative of an improved selection of papers that are finally published.

As in this study the LIWC program was used for the first time for analysis of texts in the area of peer review and as the results of this study are based on only two journals, it would be desirable for future studies to examine review texts written in connection with further journals, and mainly journals in fields other than chemistry (AC-IE) and atmospheric chemistry (ACP). These future studies could

establish the extent to which the present findings can be generalized. Further studies could also focus more strongly on certain language categories and not include some of the categories examined here (such as adverbs, prepositions, or conjunctions). In the run-up to the studies, hypotheses or expectations could be formulated regarding the frequencies of the usage of certain categories, which would then be examined in a well-aimed fashion. This study used the LIWC program for the first time to analyze review texts on journal manuscripts, and being an exploratory study, it did not test hypotheses.

As a limitation we would like to mention that this study included linguistic properties, but the LIWC is a psychology-oriented software program “that counts words in psychologically meaningful categories” (Tausczik & Pennebaker, 2010, p. 24). This study cannot, therefore, replace linguistic studies of reviewers’ comments. A word count strategy such as LIWC is an admittedly crude way to study reviewers’ comments. One of the obvious limitations of this technique is that many words are ambiguous when they are isolated from their contexts. LIWC cannot detect the context or underlying meaning of words. To solve this problem, some researchers are now investigating the frequency of specific combinations of words (e.g., Oberlander & Gill, 2006).

We are left with the question as to whether the language differences are due to the different review processes or differences in the personalities of the reviewers. In PPR, the open review policy process might attract a specific, motivated audience and invite them to contribute their expert views. In contrast, in CPR the selection of reviewers is usually based on a decision by the editors. Future research should take a closer look at potential moderator variables to disentangle the question as to what factors shape the language differences in the reviews.

References

- Anon. (2006). Peer review on trial. *Nature*, 441(7094), 668.
- Bakanic, V., McPhail, C., & Simon, R. J. (1989). Mixed messages: referees' comments on the manuscripts they review. *Sociological Quarterly*, 30(4), 639-654.
- Bingham, C. M., Higgins, G., Coleman, R., & Van Der Weyden, M. B. (1998). *The Medical Journal of Australia* Internet peer-review study. *Lancet*, 352(9126), 441-445.
- Bornmann, L., & Daniel, H.-D. (2008a). The effectiveness of the peer review process: inter-referee agreement and predictive validity of manuscript refereeing at *Angewandte Chemie International Edition*, 47(38), 7173-7178. doi: 10.1002/anie.200800513.
- Bornmann, L., & Daniel, H.-D. (2008b). Selecting manuscripts for a high impact journal through peer review: a citation analysis of Communications that were accepted by *Angewandte Chemie International Edition*, or rejected but published elsewhere. *Journal of the American Society for Information Science and Technology*, 59(11), 1841-1852. doi: 10.1002/asi.20901.
- Bornmann, L., & Daniel, H.-D. (2009a). Extent of type I and type II errors in editorial decisions: a case study on *Angewandte Chemie International Edition*. *Journal of Informetrics*, 3(4), 348-352.
- Bornmann, L., & Daniel, H.-D. (2009b). Reviewer and editor biases in journal peer review: an investigation of manuscript refereeing at *Angewandte Chemie International Edition*. *Research Evaluation*, 18(4), 262-272.
- Bornmann, L., & Daniel, H.-D. (2010a). How long is the peer review process for journal manuscripts? A case study on *Angewandte Chemie International Edition*. *Chimia*, 64(1-2), 72-77.
- Bornmann, L., & Daniel, H.-D. (2010b). Reliability of reviewers' ratings at an interactive open access journal using public peer review: a case study on *Atmospheric Chemistry and Physics*. *Learned Publishing*, 23(2), 124-131.
- Bornmann, L., & Daniel, H.-D. (2010c). The validity of staff editors' initial evaluations of manuscripts: a case study of *Angewandte Chemie International Edition*. *Scientometrics*, 85(3), 681-687.
- Bornmann, L., Marx, W., Schier, H., Rahm, E., Thor, A., & Daniel, H. D. (2009). Convergent validity of bibliometric Google Scholar data in the field of chemistry. Citation counts for papers that were accepted by *Angewandte Chemie International Edition* or rejected but published elsewhere, using Google Scholar, Science Citation Index, Scopus, and Chemical Abstracts. *Journal of Informetrics*, 3(1), 27-35. doi: 10.1016/j.joi.2008.11.001.
- Bornmann, L., Marx, W., Schier, H., Thor, A., & Daniel, H.-D. (2010). From black box to white box at open access journals: predictive validity of manuscript reviewing and editorial decisions at *Atmospheric Chemistry and Physics*. *Research Evaluation*, 19(2), 105-118.
- Bornmann, L., Nast, I., & Daniel, H.-D. (2008). Do editors and referees look for signs of scientific misconduct when reviewing manuscripts? A quantitative content analysis of studies that examined review criteria and reasons for accepting and rejecting manuscripts for publication. *Scientometrics*, 77(3), 415-432.
- Bornmann, L., Neuhaus, C., & Daniel, H.-D. (2011). The effect of a two-stage publication process on the Journal Impact Factor: a case study on the interactive open access journal *Atmospheric Chemistry and Physics*. *Scientometrics*, 86, 93-97. doi: 10.1007/s11192-010-0250-4.
- Bornmann, L., Schier, H., Marx, W., & Daniel, H.-D. (2011). Is interactive open access publishing able to identify high-impact submissions? A study on the predictive validity of *Atmospheric Chemistry and Physics* by using percentile rank classes. *Journal of the American Society for Information Science and Technology*, 62(1), 61-71.
- Bornmann, L., Weymuth, C., & Daniel, H.-D. (2010). A content analysis of referees' comments: how do comments on manuscripts rejected by a high-impact journal and later published in either a low- or high-impact journal differ? *Scientometrics*, 83(2), 493-506.
- Chung, C. K., & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedler (Ed.), *Social Communication* (pp. 343-359). New York, NY, USA: Psychology Press.
- Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology*, 94(2), 334-346.
- Fraser, V. J., & Martin, J. G. (2009). Marketing data: has the rise of impact factor led to the fall of objective language in the scientific article? *Respiratory Research*, 10. doi: 10.1186/1465-9921-10-35.
- Harnad, S. (2000). The invisible hand of peer review. *Exploit Interactive*(5).
- Hartley, J., & Betts, L. (2009). Common weaknesses in traditional abstracts in the social sciences. *Journal of the American Society for Information Science and Technology*, 60(10), 2010-2018.

- Hartley, J., Pennebaker, J. W., & Fox, C. (2003). Abstracts, introductions and discussions: how far do they differ in style? *Scientometrics*, 57(3), 389-398.
- Kahn, J. H., Tobin, R. M., Massey, A. E., & Anderson, J. A. (2007). Measuring emotional expression with the Linguistic Inquiry and Word Count. *American Journal of Psychology*, 120(2), 263-286.
- Koonin, E., & Lipman, D. (2006). Systems: reviving a culture of scientific debate. Can 'open peer review' work for biologists? Biology Direct is hopeful. Retrieved 21 June 2006, from <http://www.nature.com/nature/peerreview/debate/nature05005.html>
- Mungra, P., & Webber, P. (2010). Peer review process in medical research publications: language and content comments. *English for Specific Purposes*, 29(1), 43-53. doi: 10.1016/j.esp.2009.07.002.
- Oberlander, J., & Gill, A. J. (2006). Language with character: a stratified corpus comparison of individual differences in e-mail communication. *Discourse Processes*, 42(3), 239-270.
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A., & Booth, R. J. (2007). *The development and psychometric properties of LIWC 2007*. Austin, TX, USA: University of Texas, University of Auckland, New Zealand.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6), 1296-1312.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: our words, our selves. *Annual Review of Psychology*, 54, 547-577. doi: 10.1146/annurev.psych.54.101601.145041.
- Popping, R. (2000). *Computer-assisted text analysis*. London, UK: Sage Publications.
- Pöschl, U. (2004). Interactive journal concept for improved scientific publishing and quality assurance. *Learned Publishing*, 17(2), 105-113. doi: 10.1087/095315104322958481.
- Pöschl, U. (2010). Interactive open access publishing and peer review: the effectiveness and perspectives of transparency and self-regulation in scientific communication and evaluation. *Liber Quarterly*, 19(3/4), 293-314.
- StataCorp. (2011). *Stata statistical software: release 12*. College Station, TX, USA: Stata Corporation.
- Tausczik, Y. R., & Pennebaker, J. W. (2010). The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1), 24-54. doi: 10.1177/0261927x09351676.

Table 1

Mean numbers and proportions of words per LIWC category in reviewer comments (ACP RC) and in short comments (ACP SC) for manuscripts submitted to ACP and in reviewer comments for manuscripts submitted to AC-IE (AC-IE RC). Except for the categories “Word count” and “Words per sentence,” all means are expressed as percentage of total words in any given comment. Statistically significant differences between two mean proportions are printed with asterisks; statistically highly significant differences are shown in bold typeface.

Category	Abbreviation	Examples	ACP RC (<i>n</i> = 223)		ACP SC (<i>n</i> = 139)		Absolute difference between ACP RC and ACP SC	AC-IE RC (<i>n</i> = 237)		Absolute difference between ACP RC and AC-IE RC
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		<i>M</i>	<i>SD</i>	
General descriptor and linguistic property (number and proportion)										
Word count (number)	wc		795.26	520.64	447.31	348.67	347.95***	200.39	157.36	594.87***
Words per sentence (number)	wps		24.51	6.60	31.00	11.68	6.49***	22.42	7.62	2.1**
Dictionary words (proportion)	Dic		68.99	5.16	68.39	4.98	0.59	68.98	5.38	0.01
Words>6 letters (proportion)	Sixltr		28.25	4.01	29.89	5.07	1.64**	31.91	5.33	3.65***
Total function words (proportion)	funct		46.77	3.81	46.41	4.31	0.36	48.43	4.63	1.65***
Question marks	Qmark	?	0.73	0.76	0.33	0.66	0.4	0.26	0.53	0.46***
Exclamation marks	Exclam	!	0.03	0.11	0.02	0.08	0.01	0.12	0.44	0.09**
Quotation marks	Quote	“	1.14	1.64	0.54	1.16	0.6***	0.00	0.00	1.14***
Parentheses	Parenth	()	1.11	0.73	1.28	0.97	0.17	1.31	1.41	0.21
Total pronouns	pronoun	I, them, itself	5.61	1.74	5.63	2.17	0.02	5.61	2.65	0.0
Personal pronouns	ppron	I, them, her	1.41	0.93	1.60	1.34	0.19	1.36	1.30	0.05
1st person singular	i	I, me, mine	0.82	0.68	0.69	0.86	0.13	0.93	1.03	0.12
1st person plural	we	we, us, our	0.06	0.12	0.32	0.59	0.26***	0.04	0.18	0.03
2nd person	you	you, your, thou	0.18	0.46	0.19	0.64	0.01	0.01	0.06	0.17***
3rd person singular	shehe	she, her, him	0.01	0.04	0.03	0.23	0.02	0.01	0.06	0.0
3rd person plural	they	they, their, they'd	0.34	0.37	0.37	0.56	0.03	0.38	0.71	0.04
Impersonal pronouns	ipron	it, it's, those	4.19	1.26	4.02	1.62	0.17	4.25	2.05	0.05
Articles	article	a, an, the	9.93	2.01	9.87	2.23	0.06	10.42	2.47	0.49*
Common verbs	verb	walk, went, see	9.66	1.89	8.82	2.26	0.84***	10.35	2.89	0.69**
Auxiliary verbs	auxverb	am, will, have	7.34	1.55	6.73	1.87	0.61**	8.20	2.29	0.86***
Past tense	past	went, ran, had	1.45	0.78	1.49	1.15	0.03	1.50	1.21	0.05
Present tense	present	is, does, hear	5.53	1.24	4.97	1.90	0.56**	6.20	2.03	0.67***
Future tense	future	will, is gonna	1.25	0.61	1.10	0.68	0.15*	1.23	0.94	0.03

Adverbs	adverb	very, really, quickly	2.83	0.94	2.81	1.18	0.02	2.85	1.72	0.02
Prepositions	preps	to, with, above	13.90	1.76	14.87	1.84	0.97***	13.80	2.63	0.1
Conjunctions	conj	and, but, whereas	4.73	1.11	4.65	1.43	0.07	4.64	1.82	0.08
Negations	negate	no, not, never	1.00	0.56	0.86	0.70	0.13*	1.54	1.19	0.54***
Assent	assent	agree, OK, yes	0.12	0.29	0.11	0.35	0.01	0.06	0.22	0.06*
Quantifiers	quant	few, many, much	3.47	1.26	2.89	1.30	0.58***	3.31	1.55	0.16
Numbers	number	second, thousand	4.50	2.43	3.03	2.09	1.47***	3.11	2.75	1.39***
Words connoting psychological processes (proportion)										
Social processes	social	mate, talk, they, child	2.46	1.01	2.40	1.23		2.52	1.60	0.07
Affective processes	affect	happy, cried, abandon	3.11	0.95	3.04	1.41	0.06	3.48	1.69	0.38**
Positive emotion	posemo	love, nice, sweet	2.16	0.80	2.09	1.14	0.07	2.64	1.55	0.48***
Negative emotion	negemo	hurt, ugly, nasty	0.96	0.58	1.00	0.82	0.06	0.87	0.84	0.09
Cognitive processes	cogmech	cause, know, ought	16.56	2.57	15.79	3.07	0.03*	15.58	3.14	0.98***
Insight	insight	think, know, consider	3.49	1.08	3.37	1.48	0.76	3.04	1.58	0.44**
Causation	cause	because, effect, hence	3.01	1.01	3.11	1.20	0.12	2.89	1.63	0.11
Discrepancy	discrep	should, would, could	1.74	0.74	1.52	0.96	0.1*	1.70	1.13	0.04
Tentative	tentat	maybe, perhaps, guess	2.56	0.99	2.31	1.19	0.22	1.97	1.27	0.59***
Certainty	certain	always, never	1.15	0.71	1.06	0.77	0.25	1.04	0.96	0.11
Inhibition	inhib	block, constrain, stop	0.44	0.39	0.45	0.52	0.09	0.62	0.88	0.18**
Inclusive	incl	and, with, include	3.49	1.08	3.64	1.61	0.01	3.29	1.45	0.21
Exclusive	excl	but, without, exclude	1.87	0.89	1.61	0.95	0.15*	2.10	1.41	0.23*
Perceptual processes (proportion)	percept	observing, heard, feeling	0.86	0.61	1.00	1.06	0.26	0.66	0.70	0.21**
Biological processes (proportion)	bio	eat, blood, pain	0.45	0.44	0.47	0.50	0.02	0.34	0.56	0.11*
Custom category (proportion)										
Citations			0.51	0.71	1.31	1.49	0.8***	0.28	0.56	0.23***

Note. * p < .05. ** p < .01. *** p < .001.

ACP = *Atmospheric Chemistry and Physics*; ACP RC = reviewer comments on manuscripts submitted to ACP; ACP SC = short comments on manuscripts submitted to ACP; AC-IE = *Angewandte Chemie International Edition*; AC-IE RC = reviewer comments submitted to AC-IE