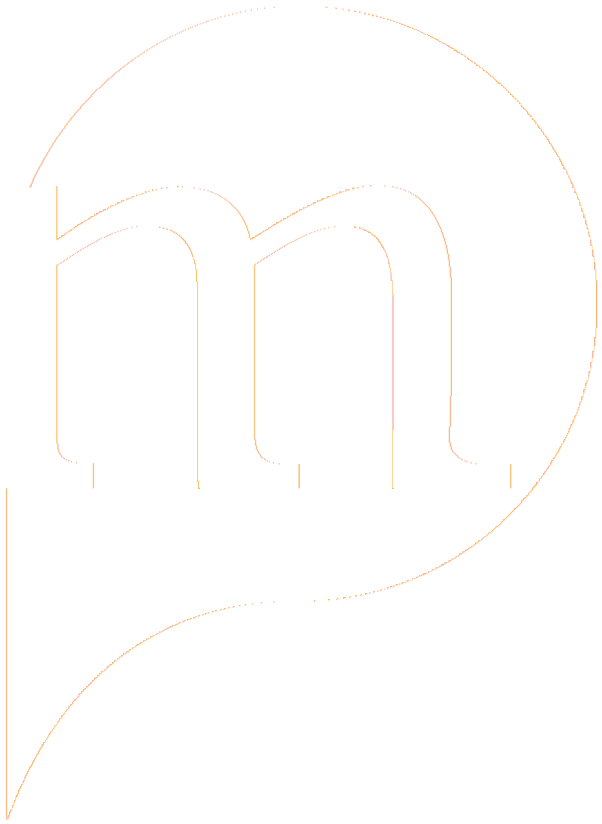

What We Still Don't Know About Peer Review

OMAR SABAJ MERUANE, CARLOS GONZÁLEZ VERGARA, and ÁLVARO PINA-STRANGER

Despite criticisms, the peer review process (PRP) is undoubtedly well established as an official and legitimated mechanism for evaluating and controlling scientific production. Although PRP has been a prominent object of study, we argue in this article that empirical research on PRP has not been addressed in a comprehensive way. Nine categories were applied to 150 empirical research articles on the topic with results revealing various gaps in empirical PRP research: (1) the research has been dedicated to the evaluation of the system rather than to the actual description of PRP as a concrete socio-discursive practice; (2) the most productive group of studies considers the multiple relationships between the sociological attributes (socio-demographic or scientometrical) of the actors (authors, reviewers, and editors) and the results of the process but does not take into account the texts exchanged by those actors; and (3) the few studies that do analyze the texts interchanged in the process do not take into account any of the variables included (such as scientometrical data, agreement, and rejection rates) in the more productive areas of the field. This lack of integration among the methodological approaches to PRP results in a partial comprehension of this important process, which determines the production and dissemination of an important part of scientific knowledge.

Keywords: peer review; research articles; production and evaluation of scientific knowledge

> Sadly, the majority of peer review research is dilettante science research, practised by scientists who are not trained to observe their own practices.[1]

In the period 1969–2006, there are 3720 publications under the topic search of the term 'peer review.'[2] Using the 'advanced search' function (TS = 'Peer review' and Document types = Article) of the Web of Science

(WOS), version 5.10 interface, one can obtain 1241 records, coded as research articles, in the period 2007–13. A similar search on SpringerLink for the period 1969–2013 provides 22,926 publications. In addition, Lutz Bornmann suggests that in order to have a complete picture of the research done on the peer review process (PRP), one should also consider the great amount of grey literature on the topic.[3] In fact, ever since the seminal study of Harriet Zuckerman and Robert Merton,[4] the interest in PRP has never decreased.

There are several reasons for the proliferation of research in this field. First, the importance of PRP to the production, evaluation, and consumption of science is evident. Second, as Bornmann shows, the results of the empirical evidence of the process are inconclusive.[5] Third, PRP has been the subject of multiple criticisms that have encouraged researchers to debate whether the process is biased or unbiased, reliable or unreliable, predictively valid or invalid (or if it is able, or unable, to predict the impact of a paper as measured by the number of citations it generates).[6]

Considering the enormous amount of research on the subject, it would be risky to state that PRP is an under-researched topic. Nevertheless, as a result of the analysis of some of the most-cited reviews on the subject,[7] it can be inferred that the evidence does not tend to be conclusive regarding the multiple types of relationships involved in this process. This is a clear indicator that, despite the great amount of research accumulated, PRP is a field that has yet to be fully defined. In particular, Stefan Hirschauer has pointed out some important gaps, or under-researched areas, in empirical studies of PRP.[8] In this article, we show evidence of these and other gaps in order to promote, as Reinhart does, more interdisciplinary research on the topic.[9]

This article is organized as follows. In the first section, various definitions for PRP are critically revised before a general delineation of the term is given; then, a general review of the topic is provided, describing the main disciplines that have studied PRP and their most frequent objects of study. In the second section, some gaps in the field will be critically described based on the analysis of 150 empirical studies on PRP, using the guidelines of grounded theory.[10] As a result, it will be argued that, despite the great amount of research on the topic, empirical research in the field has not been comprehensively done.

PRP: SOME PRELIMINARY DEFINITIONS

PRP is considered to be the main quality control mechanism of science[11] and can be described in the following steps: (1) an author, or a group or authors, sends an article to the editor of a journal; (2) the editor selects a number of reviewers (usually two); (3) the reviewers write a report with evaluative comments and a recommendation (accepted, minor or major revisions, or rejected); and (4) considering the recommendations of the reviewers, the editor makes a decision that is communicated to the author(s).[12] More steps can be added, such as when the reviewers are asked to evaluate the changes in the revised manuscript and to state their conformity or disconformity with them. Also, when the referees do not agree in their recommendations, the editor usually selects another reviewer. Several studies report that PRP is implemented with a high degree of heterogeneity among disciplines and journals.[13] These differences can be identified in the decision process, the selection of reviewers, or in the criteria used in evaluations. Commonly, PRP is also classified based on the degree of secrecy in the process: double blind, single blind, open, or public.[14]

PRP is not easy to define because its characterization strongly depends on the scope used to delineate it. From the point of view of public policies for research, PRP is a scientific quality control mechanism that determines the allocation of resources to finance the scientific industry. For editors, PRP is a system that allows them to 'separate the wheat from the chaff,' an input helping them to decide what to publish. For authors, PRP is an obligatory step to have their work published as well as an opportunity to receive feedback from the members of the community, which permits them to improve the quality of their manuscripts. For reviewers, PRP is a way to help maintain the high standards of, and participate in, the generation of knowledge in their fields.

From a socio-economic framework, Flaminio Squazzoni[15] considers PRP (and science in general) to be an imperfect economic exchange system because the same actors involved in the system pursue different interests as they assume different roles (authors, reviewers, and editors). Besides, these actors do not share the same information when they take on specific roles due to secrecy requirements. In this article, we use the following general definitions of the term: Conceptually, PRP is a collective action in which actors epistemically coordinate. In particular, it is a socio-technical judgmental discursive practice that determines the

production, dissemination, and consumption of scientific knowledge. Operationally, a single PRP is an interaction among actors who have specific social attributes and exchange texts during the process. Furthermore, the same actor can fulfil different roles (author, reviewer, or editor) in various single PRPs (that is, as an author in one case, as a reviewer in another, and maybe even as an editor in a third case).

THE MAIN DISCIPLINES THAT HAVE STUDIED PRP

The seminal work of Zuckerman and Merton[16] shows that the relative status between authors and referees does not influence the rejection or acceptance rates, although status by itself does affect the duration of the process and is correlated with specific decision paths (resulting in verdicts of accepted, rejected, or requiring revision, among others). During the 1960s and 1970s of the last century, PRP was researched, directly or indirectly, by some sociologists of science. The work of the Cole brothers,[17] heirs of the Mertonian tradition, is a good example of this first stage in the research on PRP. Although the sociology of science was still generating knowledge about PRP,[18] two disciplines started concentrating the empirical research on PRP in the 1980s: medicine (particularly medical sub-specialties) and psychology. These studies have mainly focused on the evaluation of the system in terms of its reliability, predictive validity, and fairness or the absence (presence) of bias.[19] All of this research has been conducted mainly by the editors of medical journals who were worried about the criticism of the system. They looked for ways to improve its deficiencies and propose new alternatives.[20] A summary of these disciplines, their data, and the objects of study is provided in Table 1.

METHODS

To show the existence of gaps in the empirical research on peer review, three stages were followed: searching and sampling a corpus of empirical research on PRP, cleaning the database, and analyzing the data set. In the first stage, keyword searches were conducted in the most-used scientific databases (such as WOS, Scopus, ScienceDirect, Elsevier, SpringerLink, and JSTOR), using variations of the following terms: scientific communication, peer review, evaluation of science, scientific quality, journals. The references cited by Eugene Garfield,[21] Bornmann,[22] and Juan Miguel Campanario[23] were also used as a point of comparison with our corpus of analysis. The intention was not to produce an exhaustive

TABLE 1. The Main Disciplines That Have Studied PRP

Discipline	Objects of study
Sociology of Science	• Rates of acceptance / rejection / revision • Socio-demographic data: sex, age, nationality, affiliation, discipline, and journal • Relational data = sociometric, bibliometric, and scientometric data:[24] citation analysis, impact factor, H-index, number of papers published, rankings, prestige, funds obtained, and other indicators. • Reward systems (involving prestige and funding) • Science production (involving the number of papers published, their quality, and impact) • Evaluation patterns
Medicine / Psychology	• Rates of acceptance / rejection / revision • Bias / fairness • Agreement of reviewers / reliability • Predictive validity • Evaluation criteria

corpus but, rather, to have a representative sample of the most prominent areas, journals, and authors that appear in the recognized literature.[25] Concordantly, all articles that did not have full-text availability online were excluded. The great majority of analyzed articles correspond to the study of peer review for journals, although research on the PRP of grant applications was not excluded. Two independent researchers cleaned the database of false positives, which were studies on peer review in other settings (educational and psychological).

The analysis consisted of two steps. First, all analysis was done using the guidelines of grounded theory, which essentially include the general exploration of the data, the generation of relevant categories, the definition and use of those categories to describe all data, and a mechanism for assuring the consistency of the categories in terms of agreement or consensus.[26] After a general exploration of the data, the following categories were generated to classify the corpus in terms of types of studies: (1) empirical studies; (2) essays, editorials, and general information; (3) improvement proposals; (4) reviews (see Table 2). In addition, the doctoral background of each main author was registered (see Table 3).

TABLE 2. The Data Set

Type of study	N
Empirical studies[27]	150
Essays, editorials or general information	114
Improvement proposals	53
Reviews	15
Total	332

TABLE 3. The Doctoral Backgrounds of the Main Authors of Empirical Research on PRP

N	Disciplines	N	%
1	Medicine/Psychology	58	38
2	Bibliometrics/Scientometrics	36	24
3	Sociology of science/Science and technology studies	19	13
4	Discourse analysis	11	7
5	Economics/Administration	9	6
6	Education/Public policy	4	3
7	Other (anthropology, biology, computer science, environmental science, philosophy, physics, engineering, chemistry)	13	9
	Totals	150	100

The second step of the analysis only considered the articles included in the empirical studies category. Following the same procedure, nine categories were generated to identify the variables, or topics, that were relevant in characterizing empirical research on PRP: (1) rates of acceptance/rejection/revision (minor or major); (2) bias; (3) agreement among reviewers; (4) predictive validity; (5) evaluation criteria; (6) socio-demographic data; (7) sociometric/scientometric data; (8) discourse analysis of isolated texts; and (9) discourse analysis of related texts (see Table 4).

For each category, a question and a description defining that category were proposed along with an example and a counter-example. The three researchers agreed on the pertinence of the examples and the counter-examples, and only the questions and descriptions were edited by consensus.[28] A pair of assistants independently applied, without further

TABLE 4. Categories Studied in Empirical Research on PRP

Code	Categories	N	%
A	Rates of acceptance/rejection/revision (minor or major)	38	25
B	Bias	46	31
C	Agreement among reviewers	26	17
D	Predictive validity	36	24
E	Evaluation criteria	54	36
F	Socio-demographic data	124	83
G	Sociometric/Scientometric data	71	47
H	Discourse analysis of isolated texts	13	9
H1	Discourse analysis of related texts	7	5

TABLE 5. Interaction between Categories

Code	Categories	A	B	C	D	E	F	G	H	H1
A	Rates of acceptance/rejection/revision (minor or major)		10	5	10	10	21	9	2	2
B	Bias			4	9	17	26	15	2	2
C	Agreement among reviewers				6	7	14	7	0	0
D	Predictive validity					10	17	16	1	0
E	Evaluation criteria						31	17	4	4
F	Socio-demographic data							36	7	4
G	Sociometric/Scientometric data								1	1
H	Discourse analysis of isolated texts									2
H1	Discourse analysis of related texts									

instruction, these nine categories to all of the articles identified as empirical studies, keeping in mind that, as the categories are not exclusive, the same article could respond positively to more than one category. As a means of checking the consistency of the categories, the coders shared one third of the articles to be classified; the resulting percentages of agreement ranged from 60 per cent to 92 per cent. To obtain the final results, all cases in which there was disagreement were resolved by consensus by the three main researchers. Simple frequency and percentages were used to analyze interactions between categories (see Table 5), and

TABLE 6. The Number of Categories included in Empirical Studies of PRP

Number of categories included	Number of studies	%
1–3	111	74
4	23	15
5–7	16	11
Totals	150	100

the number of categories considered in these studies was also computed as an indicator of comprehensiveness in the empirical research on PRP (see Table 6).

RESULTS

The Final Data Set

In the first search, 389 articles were collected and analyzed, and fifty-seven were discarded as they were considered to be off-topic. Generally, these corresponded to studies in the field of peer interaction in educational settings. The resulting general data set was then further classified according to the 'types of study' categories mentioned earlier. The relevant classes selected are shown in Table 2.

The information in Table 2 partially confirms the disadvantage of empirical research on PRP pointed out by Hirschauer.[29] Although the number of studies analyzed in this article is less than the combined number of references taken from the most-cited reviews in the field (181 references in Campanario;[30] 150 references in Reinhart;[31] and 259 references in Bornmann[32]), it is a significant corpus of references that differs from those in the following aspects: it contains only empirical works, its analysis is made using content categories, and it contains more up-to-date references (thirty-one references from 2011 to 2013).

The Doctoral Backgrounds of the Main Authors of Empirical Research on PRP

After the revision of the 150 articles on empirical PRP, it was found that the main authors doing empirical research on PRP hold doctorates in different disciplines (see Table 3). As shown in Table 3, our corpus

indicates that investigators doing empirical research on PRP are still coming mainly from the fields of medicine/psychology, bibliometrics/scientometrics, and the sociology of science (see Table 1), which support Hirschauer's[33] and Reinhart's[34] observations on the lack of integration of other fields in the study of PRP.

Categories of Empirical Research on PRP

Table 4 presents the frequencies and percentages for the categories (objects of study) considered in the empirical research on peer review. As Table 4 shows, 83 per cent of the empirical research on PRP includes socio-demographic variables (Category F), which account for the disciplines, journals, countries, age, and/or sex of the actors (authors, reviewers, or editors) under study. Besides this single category, the most prominent group of variables considered in PRP research includes scientometrical variables (Category G), such as prestige, citations, and ranking; the criteria or arguments used in evaluations (Category E); and the presence of bias (Category B). The second most prominent group consists of acceptance/rejection rates (Category A), predictive validity (Category D), and agreement among reviewers (Category C). The last group includes the discourse analysis of isolated texts (Category H) as well as those that are connected (Category H1), both of which are the least frequent types of work in the empirical field of PRP.

Table 5 shows the percentages of interaction between categories (that is, the percentage of studies that share two variables). From Table 5, we can see that the variables that most interact are Categories F and G. In general, Categories H and H1 are the ones that have the least interaction with the rest of the categories. The highest percentage of 36 per cent (Categories F and G) shows that more than a third of empirical studies include both socio-demographic data along with scientometrical variables. In general, what Table 5 indicates is that the empirical study of PRP only addresses particular aspects of the process and that the texts interchanged by the actors in the process are not prominent objects of study in the field.

Table 6 shows the number of categories included in the data of analysis. The majority of the studies (74%) consider one to three categories. Only 15 per cent of the articles analyzed includes four categories, and a minority (11%) considers more than five categories. This is an indirect measure of the comprehensiveness of the empirical research on PRP.

TABLE 7. Categories Grouped into Classes

Category/Classes	Categories
Output (systemic)	A. Rates of acceptance/rejection/modification (major or minor) B. Bias C. Agreement among reviewers (reliability) D. Predictive validity regarding scientific impact/quality E. Evaluative criteria/patterns
Socio-demographic	F. Socio-demographic data: sex, age, nationality, affiliation, discipline, journal
Relational	G. Sociometric/Scientometric/Bibliometric data: citation analysis, impact factor, H-Index, number of papers published, ranking, reward related (prestige, funding), and so on.
Content functional	H. Discursive analysis of the content and function (purpose, polarity) of one PRP text, namely, the reviewers' report.
Process (and content functional)	H1. Discursive analysis of the content and function (purpose, polarity) of part, or all, of the texts involved in a PRP.
Multiple role (a non-category class)	The actors in a PRP have had the experience of multiple roles in various PRP's (in some cases, that of author; in other cases, that of reviewer; and even that of editor)

The Gaps Found in Empirical PRP Research

In Table 7, categories of the same type are placed into classes that will be used to denote the gaps found in the overview of PRP made in this investigation.

In Table 8, the disciplines involved in the study of PRP are also placed into classes based on the objects of study they hold in common. Each group of objects of study is then paired up with the corresponding category classes from Table 7. From there, it can be easily deduced what the gaps are for each class of discipline since they are precisely those category classes that are not covered in their objects of study.

DISCUSSION

The output class of studies (medicine and psychology) has a systemic, not a procedural, conception of peer review. Analyzing typical objects of study in these disciplines (such as agreement, validity, and bias),

TABLE 8. The Relationships between Disciplines, Categories, and Gaps

Discipline classes		Disciplines	Objects of study according to categories	Objects of study according to category classes	Gaps in the objects of study according to category classes
Output (systemic)		Medicine psychology	A, B, C, D, E, F, G	Output relational	Content-functional process, multiple role
Scientific indicator	1	Sociology of science	A, D, E, F, G	Output relational	Content-functional process, multiple role
	2	Bibliometrics	G	Relational	
		Library and information science			
		Social network analysis			
		Sociometrics			
		Technology and science studies			
Content		Discourse Analysis	F, H	Content functional	Output-relational process, multiple role

one can argue that, in fact, all these topics depend more on the output (recommendation, editorial decision) of the process than on the real, socio-textual, interactions involved in it, resulting in content-functional and process gaps. Georg Steinhauser and his colleagues have indicated the inconvenience of measuring peer review with scientific standards.[35] We concur since, in our view, this field of research (with a focus on the system and not on the process) has been looking for something that is potentially dangerous to scientific development because, as some researchers argue, high agreement among reviewers could be a problem as this might result in a lack of diversity, and even redundancy, in the reviews, whereas disagreement allows for evaluation from a number of different perspectives.[36] Furthermore, not all bias is necessarily negative as there has been a documented case that favours research that is important, original, well designed, and well reported.[37]

Another limitation applies specifically to bibliometrics, scientometrics, and science and technology studies exploring peer review (Scientific Indicator 2 in Table 8). These studies are the ones focusing on Category G.

Citation analysis, impact factor, H-index, and other indicators constitute an independent, established field of empirical research on peer review. Researchers representative of this field not only come from the discipline of library and information science[38] but also from social network analysis.[39] Although this approach has provided multiple, and varied, data that enhances our understanding of the scientific communication process, the debate focuses on the relevance of the indicators as a proxy of scientific quality, or impact, without considering the process itself or the functions (content plus action) that an actor fulfils when interacting with others through a citation. This results in content-functional and process gaps. In fact, only a few works,[40] which are not necessarily in the field of peer review, combine some of the indicators with content, or with functional, analysis categories. The tendency seems to be that if an investigation has to do with indicators, it will not include content, or functional, categories—that is, we can know how many times a paper was cited, but we do not know the purpose, the polarity (positive or negative), or the content of that citation.

Studies dealing with Categories A, D, E, F, and G correspond mainly to the field of the sociology of science (Scientific Indicator 1 in Table 8), which has a long-standing tradition of research on, and concern with, PRP. These studies usually analyze sociological phenomena such as science production, reward systems, and evaluation patterns in PRP.[41] Although more dynamic (procedural) than the medicine and psychology approach, which mainly uses the result of the process as the main variable (basically just Category A), this group rarely includes actual texts (such as articles submitted, reviewers' reports, and articles published) as data involved in PRP.[42] Instead, surveys and in-depth interviews are the main instruments of data recollection.[43]

To describe the phenomena under study (science production, reward systems, and evaluation patterns), scientists in this area add the sociological attributes of the actors (authors, reviewers, and editors) to the main variable (Category A) used in the fields of medicine and psychology (the result of the process and the editorial decision). These attributes may be socio-demographic (Category F) or scientometrical (Category G). In the first case, actors are classified in terms of a general category, such as sex, age, discipline, or affiliation, among others. Scientometrical attributes are relational in the sense that they represent the relative status, or position, of an actor with respect to the rest of the members of the

community. Papers published, rankings, citations received, prestige, and funds obtained, among others, are typical examples of scientometrical data. In summary, the sociology-of-science approach relates sociological attributes to PRP to describe several phenomena, such as reward systems and editorial decisions, among others, but it does not analyze the texts that are interchanged by the actors throughout the process. All of this results in content-functional and process gaps.

Discourse analysis studies (the content class in Table 8) are relatively scarce in comparison to other fields. Excuses for this lack of research have been given in terms of the confidentiality, or the occluded nature, of the texts involved.[44] As is the case in the empirical research of peer review, ethical issues regarding sensitive data are far from being a simple issue. Discourse analysis studies have two other problems that limit our understanding of PRP. First, and most important, is the fact that scientific discourse researchers tend to ignore advances in other fields, specifically concerning bibliometrical and scientometrical indicators (Category G) as well as the social networks, or mechanisms, underlying scientific knowledge as described by classic works in the sociology of science.[45]

Much of the research using discourse analysis to investigate PRP contains only the socio-demographic data (Category F) of the actors, such as sex, age, discipline, or the journal where they have published, but none of the cases include any relational information, such as papers published and citations received, among other bibliometrical data (Category G).[46] In addition, the editorial decision (Category A), or the result of the process, which is a central variable for the rest of the fields exploring PRP, is commonly excluded in discourse analysis studies. A second restriction of discourse studies on peer review is that they tend to focus on only one specific text of the process, namely the reviewers' report. This emphasis on only one of the texts impedes analyzing peer review as a process and implies a static, partial account of it. The exception to this last critique is the classic, detailed, and rigorous work of Greg Myers.[47] All of this results in output gaps, relational gaps, and process gaps.

Finally, one additional limitation common to all of the fields described is what Campanario[48] and Cassidy Sugimoto and Blaise Cronin[49] have not sufficiently emphasized: a major flaw of most productive disciplines studying peer review is the independent data treatment for both actors

and the roles they fulfil. A basic issue to be considered when studying peer review is that the same actors can fulfil different roles in various single PRPs.[50] Thus, for example, when analyzing all of the actors (authors, reviewers, and editors) participating in a specific journal, a primary and relevant piece of information to have is which actors are more participative and in which roles, revealing the functional overlaps among actors. This results in a multiple role gap.

As a result, there are various important unresolved questions regarding PRP. Although these questions have to do with the interaction of all of the categories described before, in our view, the most urgent ones are those from the less developed areas—for example, discourse analysis (Categories H and H1). For instance, some of these questions are: Do the discursive characteristics of the revision vary according to the sociometric attributes of the reviewers (Category G-related question)? Is there a discursive structural, or linguistic, difference between articles rejected and accepted (Category A-related question)? Do actors practice what they preach—that is, if you, as a reviewer, usually focus on methods, are your methods sound when acting as an author (multiple role-related question)?

FINAL REMARKS

As we have seen, the main areas in PRP research contain several gaps. Medicine and psychology, the most productive empirical areas in the study of PRP, conceive of it as a system and not as a process, focusing more on the evaluation than on the actual description. The sociology of science and, more recently, library and information science are the second most productive fields. This group, however, pays little attention to the content and functions of the texts interchanged during PRP. Instead, they describe the relationships (using several measures, such as co-authorship, citations, affiliation networks, and so on) among the actors involved in the process. Due to restricted access to data and confidentiality matters, discourse studies are less frequent in the field. Although these studies contain socio-demographic information about the actors, they do not take into account the relationships between the scientometrical attributes of the actors and the discursive actions they use in the texts involved in PRP. In addition, discourse studies describe texts in isolation, mainly the reviewer's report, without considering that,

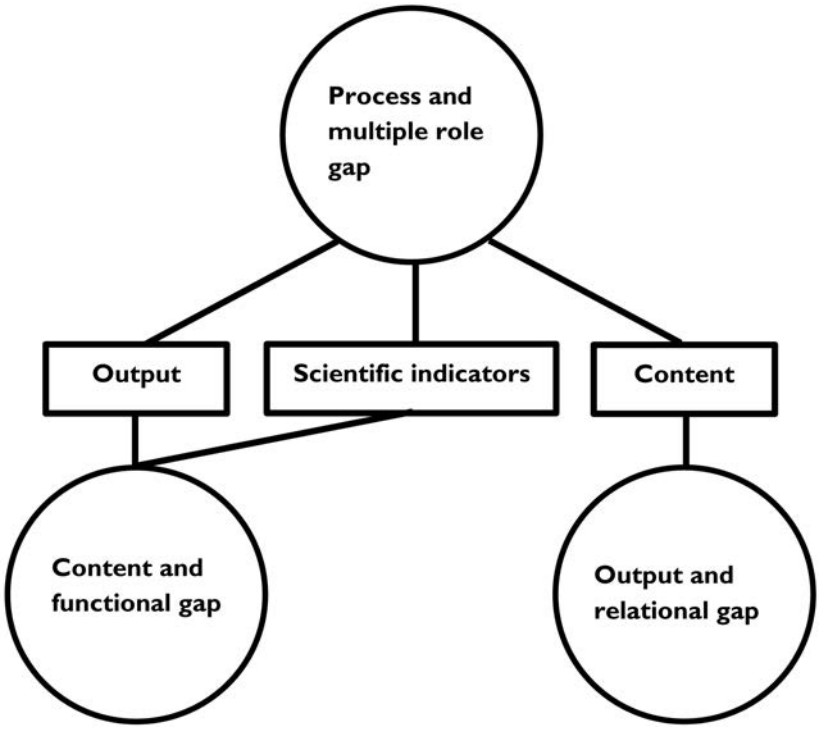

FIGURE 1. Gaps in Empirical Research on Peer Review (per discipline classes)

in PRP, texts are the products of the interactions of all of the actors who participate in it. Finally, we have noted that, when investigating the actors in PRP, all of these fields treat the sociological attributes separately from the roles the actors fulfil, ignoring the fact that there is usually an overlap—that is, the same actors can participate with different roles in various single PRPs. Figure 1 presents a diagram that shows the gaps in empirical research on PRP in relation to the classes of discipline they apply to.

As these results suggest, the study of PRP requires not only further, but more comprehensive, approaches. We have often found that the 'black box' metaphor is used to describe PRP, and as we have argued in this article, despite the massive amount of literature, this box is far from being open.[51]

OMAR SABAJ is professor of linguistics at Universidad de La Serena, Chile. His main research area is scientific discourse analysis and scientific literacy.

CARLOS GONZÁLEZ VERGARA is professor of Spanish grammar at Pontificia Universidad Católica de Chile. He is also editor-in-chief of *Onomázein: Journal of Linguistics, Philology and Translation.*

ÁLVARO PINA-STRANGER is professor of socio-economics at Université de Rennes 1 in France. His area of expertise is social network analysis in innovation systems.

ACKNOWLEDGEMENTS

We are especially grateful to Christopher F.S. Maligec, José O. Valderrama, and Paulina Meza for valuable comments on the drafts of this article. The authors thank the support of the Chilean National Fund for Scientific and Technological Development (Fondecyt No. 1130290) and the European Research Council (7th Framework Programme, ERC-2010-StG 263529).

NOTES

1. Stefan Hirschauer, 'Editorial Judgments: A Praxeology of "Voting" in Peer Review,' *Social Studies of Science* 40, 71 (2010): 71–104
2. Eugene Garfield, 'Index of Peer Review HistCite Collections' (2006), available at http://garfield.library.upenn.edu/histcomp/index-peer-rev.html (accessed 13 June 2014)
3. Lutz Bornmann, 'Scientific Peer Review,' *Annual Review of Information Science and Technology* 45, 1 (2011): 197–245
4. Harriet Zuckerman and Robert Merton, 'Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System,' *Minerva* 9, 1 (1971): 66–100
5. Bornmann, 'Scientific Peer Review'
6. Ibid.
7. Garfield, 'Index of Peer Review HistCite Collections'; Bornmann, 'Scientific Peer Review'; Eugene Garfield, 'Refereeing and Peer Review, Part 1: Opinion and Conjecture on the Effectiveness of Refereeing,' *Essays of an Information Scientist* 9, 31 (1986): 3–11; Eugene Garfield, 'Refereeing and Peer Review, Part 2: The Research on Refereeing and Alternatives to the Present System,' *Essays of an Information Scientist* 9, 32 (1986): 3–12; Eugene Garfield, 'Refereeing and Peer Review, Part 4: Research on the Peer Review of Grant Proposals and Suggestions for Improvement,' *Essays of an Information Scientist* 10, 5 (1987): 27–33; Juan Miguel Campanario, 'Peer Review for Journals As It Stands Today—Part 1,' *Science Communication* 19, 3 (1998): 181–211; Juan Miguel Campanario, 'Peer

Review for Journals As It Stands Today—Part 2,' *Science Communication* 19, 4 (1998): 277–306

8. Hirschauer, 'Editorial Judgments: A Praxeology of "Voting" in Peer Review'
9. Martin Reinhart, 'Peer Review of Grant Applications in Biology and Medicine: Reliability, Fairness, and Validity,' *Scientometrics* 81, 3 (2009): 789–909
10. Anselm Strauss and Juliet Corbin, eds., *Grounded Theory in Practice* (London: Sage 1997).
11. Bornmann, 'Scientific Peer Review'
12. Bornmann, 'Scientific Peer Review'; Lutz Bornmann and Hans-Dieter Daniel, 'Reviewer and Editor Biases in Journal Peer Review: An Investigation of Manuscript Refereeing,' *Research Evaluation* 18, 4 (2009): 262–72; Lowell Hargens and Jerald Hertings, 'Analyzing the Association between Referees' Recommendations and Editors' Decisions,' *Scientometrics* 67, 1 (2006): 15–26
13. Campanario, 'Peer Review for Journals As It Stands Today, Part 1'; Lowell Hargens, 'Scholarly Consensus and Journal Rejection Rates,' *American Sociological Review* 53 (1988): 139–51; Stephen Donovan, 'Big Journals, Small Journals, and the Two Peer Reviews,' *Journal of Scholarly Publishing* 42, 4 (2011): 534–8
14. Lutz Bornmann et al., 'In Public Peer Review of Submitted Manuscripts, How Do Reviewer Comments Differ from Comments Written by Interested Members of the Scientific Community? A Content Analysis of Comments Written for Atmospheric Chemistry and Physics,' *Scientometrics* 93 (2012): 915–29; Benjamin Baez, 'Confidentiality and Peer Review: The Paradox of Secrecy in Academe,' *Review of Higher Education* 25, 2 (2002): 163–83
15. Flaminio Squazzoni, 'Peering into Peer Review,' *Sociologica* 3 (2010): 112–15
16. Zuckerman and Merton, 'Patterns of Evaluation in Science'
17. Stephen Cole and Jonathan Cole, 'Scientific Output and Recognition: A Study in the Operation of the Reward System in Science,' *American Sociological Review* 32, 3 (1967): 377–90; Stephen Cole, Jonathan Cole, and Leonard Rubin, 'Peer Review and the Support of Science,' Scientific *American* 237, 4 (1977): 1–9; Stephen Cole, Jonathan Cole, and Simon Gary, 'Chance and Consensus in Peer Review,' *Science*, New Series 214, 4523 (1981): 881–6
18. Daryl Chubin, 'Beyond Invisible Colleges: Inspirations and Aspirations of Post: 1972 Social Studies of Science,' *Scientometrics* 7, 3–6 (1985): 221–54
19. Bornmann, 'Scientific Peer Review'
20. Ibid.
21. Garfield, 'Index of Peer Review HistCite Collections'
22. Bornmann, 'Scientific Peer Review'
23. Campanario, 'Peer Review for Journals as It Stands Today, Part 1'; Campanario, 'Peer Review for Journals as It Stands Today, Part 2'

24. Scientometrical attributes are relational in the sense that they represent the relative status or position of an actor with respect to the rest of the members of the community.
25. Garfield, 'Index of Peer Review HistCite Collections'; Bornmann, 'Scientific Peer Review'; Garfield, 'Refereeing and Peer Review, Part 1'; Eugene Garfield, 'Refereeing and Peer Review, Part 2'; Garfield, 'Refereeing and Peer Review, Part 4'; Campanario, 'Peer Review for Journals As It Stands Today, Part 1'; Campanario, 'Peer Review for Journals As It Stands Today, Part 2'
26. Hargens, 'Scholarly Consensus and Journal Rejection Rates'
27. The list of the 150 empirical studies that were analyzed appears in Appendix 1.
28. The final categories, questions, and descriptions used are available from the authors.
29. Hirschauer, 'Editorial Judgments: A Praxeology of "Voting" in Peer Review'
30. Campanario, 'Peer Review for Journals As It Stands Today, Part 2'
31. Reinhart, 'Peer Review of Grant Applications in Biology and Medicine'
32. Bornmann, 'Scientific Peer Review'
33. Hirschauer, 'Editorial Judgments: A Praxeology of "Voting" in Peer Review'
34. Reinhart, 'Peer Review of Grant Applications in Biology and Medicine'
35. Georg Steinhauser et al., 'Peer Review versus Editorial Review and Their Role in Innovative Science,' *Theoretical Medicine and Bioethics* 33, 5 (2012): 359–76
36. Campanario, 'Peer Review for Journals As It Stands Today, Part 1'
37. Bornmann, 'Scientific Peer Review'
38. Garfield, 'Index of Peer Review HistCite Collections'; Cassidy Sugimoto and Blaise Cronin, 'Citation Gamesmanship: Testing for Evidence of Ego Bias in Peer Review,' *Scientometrics* 95 (2013): 851–62
39. Loet Leydesdorff, 'Alternatives to the Journal Impact Factor: I3 and the Top-10% (or top 25%?) of the Most Highly Cited Papers,' *Scientometrics* 92 (2012): 355–65; Loet Leydesdorff and Tobias Opthof, 'A Rejoinder on Energy versus Impact Indicators,' *Scientometrics* 90 (2012): 745–8
40. Daryl Chubin and Soumyo Moitra, 'Content Analysis of References: Adjunct or Alternative to Citation,' *Social Studies of Science* 5, 4 (1975): 423–41; Álvaro Pina-Stranger et al.,'Estrategias académicas de inserción científica: Una propuesta metodológica para el estudio de las reivindicaciones epistémicas en los artículos de investigación,' *INNOVAR Journal of Administrative and Social Sciences* 23, 48 (2013): 67–82
41. Zuckerman and Merton, 'Patterns of Evaluation in Science'; Cole and Cole, 'Scientific Output and Recognition'
42. Chubin, 'Beyond Invisible Colleges'
43. Michèle Lamont, *How Professors Think: Inside the Curious World of Academic Judgment* (Harvard: Harvard University Press 2009)

44. John Swales, 'Occluded Genres in the Academy: The Case of the Submission Letter,' in E. Ventola and A. Mauranen, eds., *Academic Writing: Intercultural and Textual Issues* (Amsterdam: John Benjamins 1996): 45–58; Hugh Gosden, 'Why Not Give Us the Full Story?: Functions of Referees' Comments in Peer Reviews of Scientific Research Papers,' *Journal of English for Academic Purposes* 2 (2003): 87–10; Inmaculada Fortanet, 'Evaluative Language in Peer Review Referee Reports,' *Journal of English for Academic Purposes* 7 (2008): 27–37
45. Cole and Cole, 'Scientific Output and Recognition'; Robert Merton,'The Matthew Effect in Science,' *Science* 159, 3810 (1968): 59–63; Ian Mitroff and Daryl Chubin, 'Peer Review at the NSF: A Dialectical Policy Analysis,' *Social Studies of Science* 9, 2 (1979): 199–232
46. Adriana Bolívar, 'El informe de arbitraje como género discursivo en la dinámica de la investigación,' *Revista Latinoamericana de Estudios del Discurso* 8, 1 (2008): 41–64; Adriana Bolívar, 'Funciones discursivas de la evaluación negativa en informes de arbitraje de artículos de investigación en educación,' *Núcleo* 28 (2011): 59–89
47. Greg Myers 'Texts As Knowledge Claims: The Social Construction of Two Biology Articles,' *Social Studies of Science* 15, 4 (1985): 593–630
48. Juan Miguel Campanario, 'Have Referees Rejected Some of the Most-Cited Articles of All Times?' *Journal of the American Society for Information Science* 47, 4 (1996): 302–10
49. Sugimoto and Cronin, 'Citation Gamesmanship: Testing for Evidence of Ego Bias in Peer Review'
50. Squazzoni, 'Peering into Peer Review'
51. Garfield, 'Index of Peer Review HistCite Collections'

APPENDIX 1: LIST OF REFERENCES ANALYZED

1. Hendy Abdoul, Christophe Perrey, Philippe Amiel, Florence Tubach, Serge Gottot, Isabelle Durand-Zaleski and Corinne Alberti, 'Peer Review of Grant Applications: Criteria Used and Qualitative Study of Reviewer Practices,' *PlosOne* 7, 9 (2012), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0046054 (accessed 25 June 2014)
2. Giovanni Abramo and Ciriaco andrea D'Angelo, 'Evaluating Research: From Informed Peer Review to Bibliometrics,' *Scientometrics* 87 (2011): 499–514
3. Fernando Alonso. 'The 'Peer-Review' Process in Biomedical Journals: Characteristics of 'Elite' Reviewers,' *Neurología* 25, 9 (2010): 521–9
4. Mohammad Abooyee Ardakan, Syyed Ayatollah Mirzaie, and Fatameh Sheikhshoaei, 'The Peer-Review Process for Articles in Iran's Scientific Journals,' *Journal of Scholarly Publishing* 42, 2 (2011): 243–61
5. J. Scott Armstrong, 'Peer Review for Journals: Evidence on Quality Control, Fairness, and Innovation. *Science and Engineering Ethics* 3, 1 (1997): 63–84
6. J. Scott Armstrong, 'Barriers to Scientific Contributions: The Author's Formula,' *Behavioural and Brain Sciences* 5, 2 (1982): 197–9
7. Christopher Baethge, Jeremy Franklin, and Stephan Mertens, 'Substantial Agreement of Referee Recommendations at a General Medical Journal: A Peer Review Evaluation at Deutsches Ärzteblatt International,' *PlosOne* 8, 5 (2013), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0061401 (accessed 2 June 2013)
8. Cark. M. von Bakanic and Rita Simon, 'The Manuscript Review and Decision-Making Process,' *American Sociological Review* 52, 5 (1987): 631–42
9. Dale Baker, 'The Peer Review Process in Science Education Journals,' *Research in Science Education* 32, 2 (2002): 171–80
10. Wim G. Bend and Tim Engels, 'The Predictive Validity of Peer Review: A Selective Review of the Judgmental Forecasting Qualities of Peers, and Implications for Innovation in Science,' *International Journal of Forecasting* 27, 1 (2011): 166–82.

11. Leila Bender, Mary Klingensmith, Bradley Freeman, William Chapman, William C. Dunagan, Jonathan Gottlieb, and Bruce Hall, 'Anonymous Group Peer Review in Surgery Morbidity and Mortality Conference,' *American Journal of Surgery* 198 (2009): 270–6
12. Janice Beyer, Roland Chanove, and William Fox 'The Review Process and the Fates of Manuscripts Submitted to AMJ,'*Academy of Management Journal* 38, 5 (1995) 1219–60
13. Craig Bingham, Gail Higgins, Rose Coleman, and Martin van Der Weyden, 'The Medical Journal of Australia Internet Peer-Review Study,' *Lancet* 352, 9126 (1998): 1–10
14. Nick Black, Susan van Rooyen, Fiona Godlee, Richard Smith, and Stephen Evans, 'What Makes a Good Reviewer and a Good Review for a General Medical Journal?' *Journal of the American Medical Association* 280, 3 (1998): 231–3
15. Adriana Bolívar, 'El informe de arbitraje como género discursivo en la dinámica de la investigación,' *Revista Latinoamericana de Estudios del Discurso* 8, 1 (2008): 41–64
16. Adriana Bolívar, 'Funciones discursivas de la evaluación negativa en informes de arbitraje de artículos de investigación en educación,' *Núcleo* 28 (2011): 59–89
17. Lutz Bormann, Ruediger Mutz, and Hans-Dieter Daniel, 'A Reliability-Generalization Study of Journal Peer Reviews: A Multilevel Meta-Analysis of Inter-Rater Reliability and Its Determinants,' *PlosOne* 5, 12 (2010), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0014331 (accessed 8 August 2014)
18. Lutz Bornmann and Hans-Dieter Daniel, 'Reviewer and Editor Biases in Journal Peer Review: An Investigation of Manuscript Refereeing at Angewandte Chemie International Edition,' *Research Evaluation* 18, 4 (2009): 262–72
19. Lutz Bornmann and Hans-Dieter Daniel, 'Selection of Research Fellowship Recipients by Committee Peer Review: Reliability, Fairness and Predictive Validity of Board of Trustees' Decisions,' *Scientometrics* 63, 2 (2005): 297–320
20. Lutz Bornmann and Hans-Dieter Daniel, 'The Validity of Staff Editors' Initial Evaluations of Manuscripts: A Case Study of Angewandte Chemie International Edition,' *Scientometrics* 85 (2010): 681–7

21. Lutz Bornmann and Hans-Dieter Daniel, 'The Usefulness of Peer Review for Selecting Manuscripts for Publication: A Utility Analysis Taking As an Example a High-Impact Journal,' *PlosOne* 5, 6 (2010), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0011344 (accessed 3 November 2014)
22. Lutz Bornmann and Loet Leydesdorff, 'Macro-Indicators of Citation Impacts of Six Prolific Countries: InCites Data and the Statistical Significance of Trends,' *PlosOne* 8, 2 (2013), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0056768 (accessed 6 December 2014)
23. Lutz Bornmann, 'The Hawthorne Effect in Journal Peer Review,' *Scientometrics* 91, 3 (2011): 857–62
24. Lutz Bornmann and Hans-Dieter Daniel, 'Selecting Scientific Excellence through Committee Peer Review: A Citation Analysis of Publications Previously Published to Approval or Rejection of Post-Doctoral Research Fellowship Applicants,' *Scientometrics* 68, 3 (2006): 427–40
25. Lutz Bornmann, 'Does the Journal Peer Review Select the "Best" from the Work Submitted? The State of Empirical Research,' *IETE Tech Review* 27, 2 (2010): 93–6
26. Lutz Bornmann and Hans-Dieter Daniel, 'The Manuscript Reviewing Process: Empirical Research on Review Requests, Review Sequences, and Decision Rules in Peer Review,' *Library and Information Science Research* 32 (2010): 5–12
27. Lutz Bornmann, Hanna Herich, Hanna Joos, and Hans-Dieter Daniel, 'In Public Peer Review of Submitted Manuscripts, How Do Reviewer Comments Differ from Comments Written by Interested Members of the Scientific Community? A Content Analysis of Comments Written for Atmospheric Chemistry and Physics,' *Scientometrics* (2012), available at http://www.iac.ethz.ch/people/hjoos/bornmann_et_al_2012 (accessed 24 November 2014)
28. Lutz Bornmann, Ruediger Mutz, and Hans-Dieter Daniel, 'Row-Column (RC) Association Model Applied to Grant Peer Review,' *Scientometrics* 73, 2 (2007): 139–47
29. Lutz Bornmann, Ruediger Mutz, and Hans-Dieter Daniel, 'How to Detect Indications of Potential Sources of Bias in Peer Review: A Generalized Latent Variable Modeling Approach Exemplified by a Gender Study,' *Journal of Informetrics* 2 (2008): 280–7

30. Lutz Bornmann, Ruediger Mutz, and Hans-Dieter Daniel, 'The Influence of the Applicants' Gender on the Modelling of a Peer Review Process by Using Latent Markov Models,' *Scientometrics* 81, 2 (2009): 407–11
31. Lutz Bornmann, Irina Nast, and Hans-Dieter Daniel, '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 (2008): 415–32
32. Lutz Bornmann, Christophe Weymuth, and Hans-Dieter Daniel, 'A Content Analysis of Referees' Comments: How Do Comments on Manuscripts Rejected by a High-Impact Journal and Later Published in Either a Lower High-Impact Journal Differ?' *Scientometrics* 83 (2010): 493–506
33. Lutz Bornmann, Markus Wolf, and Hans-Dieter Daniel, 'Closed Versus Open Reviewing of Journal Manuscripts: How Far Do Comments Differ in Language Use?' *Scientometrics* 91 (2012): 843–56
34. Christina Bunner and Elaine Larson, 'Assessing the Quality of the Peer Review Process: Author and Editorial Board Member Perspectives,' *American Journal of Infection Control* 40 (2012): 701–4
35. Michael Callaham, Robert Knoop, and John Gallagher, 'Effect of Written Feedback by Editors on Quality of Reviews,' *Journal of the American Medical Association* 287, 21 (2002): 2781–3
36. Juan Campanario, 'Consolation for the Scientist: Sometimes It Is Hard to Publish Papers That Are Later Highly Cited,' *Social Studies of Science* 23, 2 (1993): 342–62
37. Juan Miguel Campanario, 'Have Referees Rejected Some of the Most-Cited Articles of All Times?' *Journal of the American Society for Information Science* 47, 4 (1996): 302–10
38. Kathleen Carson and Rainer Glaser, 'Chemistry Is in the News: Assessing Intra-Group Peer Review,' *Assessment and Evaluation in Higher Education* 35, 4 (2010): 381–402
39. Yin Lin Cheung, 'Critical Feedback on Peer Review Research,' *Procedia Social and Behavioural Sciences* 15 (2011): 535–8
40. Daryl Chubin, 'Beyond Invisible Colleges: Inspirations and Aspirations of Post 1972 Social Studies of Science,' *Scientometrics* 7, 3–6 (1985): 221–54

41. Domenic Cicchetti and Harold Conn, 'A Statistical Analysis of Reviewer Agreement and Bias in Evaluating Medical Abstracts,' *Yale Journal of Biology and Medicine* 49 (1976): 373–83
42. Stephen Cole and Jonathan Cole, 'Scientific Output and Recognition: A Study in the Operation of the Reward System in Science,' *American Sociological Review* 32, 3 (1967): 377–90
43. Stephen Cole, Jonathan Cole, and Leonard Rubin, 'Peer Review and the Support of Science,' *Scientific American* 237, 4 (1977): 1–9
44. Stephen Cole, Jonathan Cole, and Gary Simon, 'Chance and Consensus in Peer Review,' *Science, New Series* 214, 4523 (1981): 881–6
45. Tom Coupé, 'Peer Review Versus Citations: An Analysis of Best Paper Prizes,' *Research Policy* 42 (2013): 295–301
46. Diana Crane, 'The Gatekeepers of Science: Some Factors Affecting the Selection of Articles for Scientific Journals,' *American Sociologist* 2, 4 (1967): 195–201
47. Susan Eberley and Keith Warner, 'Fields or Subfields of Knowledge: Rejection Rates and Agreement in Peer Review,' *American Sociologist* 21, 3 (1990): 217–31
48. Nancy Eisenberg, Marylin Thompson, Susan Augir, and Elizabeth Harris, '"Getting In" Revisited: An Analysis of Manuscript Characteristics, Reviewers' Ratings, and Acceptance of Manuscripts in Psychological Bulletin,' *Psychological Bulletin* 128, 6 (2002): 997–1004
49. Rune Elvik, 'Are Road Safety Evaluation Studies Published in Peer Reviewed Journals More Valid Than Similar Studies Not Published in Peer Reviewed Journals?' *Accident Analysis and Prevention* 30, 1 (1998): 101–18
50. Edzard Ernst and Klaus Resch, 'Reviewer Bias against the Unconventional? A Randomized Double-Blind Study of Peer Review,' *Complementary Therapies in Medicine* 7 (1999): 19–23
51. Arthur Evans, Robert Mcnutt, Suzanne Fletcher, and Robert Fletcher, 'The Characteristics of Peer Reviewers Who Produce Good-Quality Reviews,' *Journal of General Internal Medicine* 8, 8 (1993): 422–8
52. Robert Felt, "Readers" Rights: Peer Review in Chinese Medical Publication,' *Clinical Acupuncture and Oriental Medicine* 2, 1 (2001): 9–16

53. Donald Fiske and Louis Fogg, 'But the Reviewers Are Making Different Criticisms of My Paper! Diversity and Uniqueness in Reviewer Comments,' *American Psychologist* 45, 5 (1990): 591–8
54. Alan Forster, Monica Taljaard, Carol Benett, and Carl Walraven, 'Reliability of the Peer-Review Process for Adverse Event Rating,' *Plosone* 7, 7 (2012), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0041239 (accessed 3 January 2013)
55. Inmaculada Fortanet, 'Evaluative Language in Peer Review Referee Reports,' *Journal of English for Academic Purposes* 7 (2008): 27–37
56. Jean- Michel Fortin and David Currie, 'Big Science vs. Little Science: How Scientific Impact Scales with Funding,' *PlosOne* 8, 6 (2013), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0065263 (accessed 5 March 2014)
57. Robert Frodeman and Adam Briggle, 'The Dedisciplining of Peer Review,' *Minerva* 50 (2012): 3–19
58. Adela García-Aracil, Antonio Gutiérrez, and Marián Pérez-Marín, 'Analysis of the Evaluation Process of the Research Performance: An Empirical Case,' *Scientometrics* 67, 2 (2006): 213–30
59. Armen Gasparyan and George Kitas, 'Best Peer Reviewers and the Quality of Peer Review in Biomedical Journals,' *Croat Med Journal* 53, 4 (2012): 386–9
60. Fiona Godlee, Catharine Gale, and Christopher Martyn, 'Effect on the Quality of Peer Review of Blinding Reviewers and Asking Them to Sign Their Reports: A Randomized Controlled Trial,' *Journal of the American Medical Association* 280, 3 (1998): 237–40
61. Steven Goodman, Jesse Berlin, Suzanne Fletcher, and Robert Fletcher, 'Manuscript Quality before and after Peer Review and Editing at Annals of Internal Medicine,' *Annals of Internal Medicine* 121, 1 (1994): 11–21
62. Hugh Gosden, 'Why Not Give Us the Full Story?': Functions of Referees' Comments in Peer Reviews of Scientific Research Papers,' *Journal of English for Academic Purposes* 2, 2 (2003): 87–101
63. Sarah Greaves, Joanna Scott, Maxine Clarke, Linda Miller, Timo Hannay, Annette Thomas, and Philip Campbell, 'Nature's Trial of Open Peer Review,' *Nature* (2006), available at http://www.nature.com/nature/peerreview/debate/nature05535.html (accessed 26 April 2013)

64. Alicia Grubb and Steve Easterbrook, 'On the Lack of Consensus over the Meaning of Openness: An Empirical Study,' *PlosOne* 6, 8, (2011) available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0023420 (accessed 26 May 2014)
65. Lowell Hargens and Jerald Herting, 'Neglected Considerations in the Analysis of Agreement among Journal Referees,' *Scientometrics* 19, 1–2 (1990): 91–106
66. Lowell Hargens and Jerald Herting, 'Analyzing the Association between Referees' Recommendations and Editors' Decisions,' *Scientometrics* 67, 1 (2006): 15–26
67. Glen Hastings, Robert Sonneborn, Gail Lee, Linda Vick, and Louis Sasmor, 'Peer Review Checklist: Reproducibility and Validity of a Method for Evaluating the Quality of Ambulatory Care,' *American Journal of Public Health* 70, 3 (1980): 222–8
68. Bryan Haynes, Chris Cotoi, Jennifer Holland, Leslie Walters, Nancy Wilczynski, Dawn Jedraszewsky, James McKinlay, Richard Parrish, and Ann McKibbon, 'Second-Order Peer Review of the Medical Literature for Clinical Practitioners,' *Journal of the American Medical Association* 295, 15 (2006): 1801–8
69. Agnes He, 'Language Use in Peer Review Texts,' *Language in Society* 22, 3 (1993): 403–20
70. Monica Helton and William Balistreri, 'Peering into Peer Review,' *Journal of Pediatrics* 159, 1 (2011): 150–1
71. Daniel Herron, 'Is Expert Peer Review Obsolete? A Model Suggest That Post Publication Reader Review May Exceed the Accuracy of Traditional Peer Review,' *Surgical Endoscopy and Other Interventional Techniques* 26 (2012): 2275–80
72. Stefan Hirschauer, 'Editorial Judgments: A Praxeology of "Voting" in Peer Review,' *Social Studies of Science* 40, 71 (2010): 71–104
73. Timothy Hofer, Steven Bernstein, Sonya DeMonner, and Rodnay Hayward, 'Discussion between Reviewers Does Not Improve Reliability or Peer Review of Hospital Quality,' *Medical Care* 38, 2 (2000): 152–61
74. Richard Horton, 'The Hidden Research Paper,' *Journal of the American Medical Association* 287, 21 (2002): 2775–8
75. Louise Howard and Greg Wilkinson, 'Peer Review and Editorial Decision- Making,' *British Journal of Psychiatry* 173 (1998): 110–13

76. Sarwat Hussain, Jawad Hussain, Adib Karam, and Gopal Vijayaraghavan, 'Focused Peer Review: The End Game of Peer Review,' *American College of Radiology* 9 (2012): 430–3
77. Jeffrey Jackson, Malathi Srinivasan, Joanna Rea, Kathlyn Fletcher, and Richard Kravitz, 'The Validity of Peer Review in a General Medicine Journal,' *PlosOne* 6, 7 (2011), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0022475 (accessed 25 October 2014)
78. Tom Jefferson, 'Quality and Value: Models of Quality Control for Scientific Research,' *Nature International Weekly Journal of Science* (2006), available at http://www.nature.com/nature/peerreview/debate/nature05031.html (accessed 5 June 2014)
79. Tom Jefferson, Philip Alderson, Elizabeth Wagner, and Frank Davidoff, 'Effects of Editorial Peer Review: A Systematic Review,' *Journal of the American Medical Association* 287, 21 (2002): 2784–6
80. Tom Jefferson, Elizabeth Wager, and Frank Davidoff, 'Measuring the Quality of Editorial Peer Review,' *Journal of the American Medical Association* 287, 21 (2002): 2786–90
81. Claire Johnson, 'Peer Review and Manuscript Processes for the Journal of Manipulative and Physiological Therapeutics,' *Journal of Manipulative and Physiological Therapeutics* 31, 4 (2008): 255–6
82. Liu Jun and Sadler Randall, 'The Effect and Affect of Peer Review in Electronic Versus Traditional Modes on L2 Writing,' *Journal of English for Academic Purposes* 2 (2003): 193–227
83. Nicholas Kadar, 'Systemic Bias in Peer Review: Suggested Causes, Potential Remedies,' *Journal of Laparoendoscopic & Advanced Surgical Techniques* 20, 2 (2010): 123–8
84. Davis Kaplan, Nicola Lacetera, Celia Kaplan, 'Sample Size and Precision in NIH Peer Review,' *PlosOne* 3, 7 (2008), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002761 (accessed 4 July 2014)
85. Christopher Kelty,Sidney Burrus and Richard Baraniuk, 'Peer Review Anew: Three Principles and a Case Study in Postpublication Quality Assurance,' *Proceedings of the IEEE* 96, 6 (2008): 1000–11
86. Muhammad Khan, 'Can Citations Predict Socio-Cognitive Relationships in Peer Review System?' *IADIS European Conference* 23, 6 (2010): 19–26

87. Walter Klopffer, Gustav Sudstrom and Griefhammer, 'The Peer Reviewing Process - A Case Study European Life Cycle Inventory for Surfactant Production,' *LCA Peer Reviews* 1, 2 (1996): 113–15
88. Karin Knorr-Cetina, 'The Ethnographic Study of Scientific Work: Towards a Constructivist Interpretation of Science,' *Science observed: Perspectives on the Social Study of Science* (1983): 115–40
89. Marcin Kozak, and Lutz Bormann, 'A New Family of Cumulative Indexes for Measuring Scientific Performance,' *PlosOne* 7, 10 (2012), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0047679 (accessed 4 May 2014)
90. Richard Kravitz, Peter Franks, Mitchel Feldman, Martha Gerrity, Cindy Byrne and William Tierney, 'Editorial Peer Reviewers' Recommendations at a General Medical Journal: Are They Reliable and Do Editors Care?,' *PlosOne* 5, 4 (2010), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0010072 (accessed 27 February 2014)
91. Abhaya Kulkarni,Brittany Aziz, Iffat Shams and Jason Busse, 'Comparisons of Citations in Web of Science, Scopus, and Google Scholar for Articles Published in General Medical Journals,' *Journal of the American Medical Association* 302, 10 (2009): 1092–6
92. Charles Leslie, 'Scientific Racism: Reflections on Peer Review, Science and Ideology,' *Social Science and Medicine* 31, 8 (1990): 891–912
93. Loet Leydesdorff and Tobias Opthof, 'A Rejoinder on Energy Versus Impact Indicators' *Scientometrics* 90 (2012): 745–8
94. Loet Leydesdorff, 'Alternatives to the Journal Impact Factor: I3 and the Top-10% (Or Top-25%?) of the Most-Highly Cited Papers,' *Scientometrics* 92 (2012): 355–65
95. Ann Link, 'US and Non-US Submissionsan Analysis of Reviewer Bias,' *Journal of the American Medical Association* 280, 3 (1998): 246–7
96. Faina Linkov, Mita Lovalekar,and Ronald Laporte, 'Quality Control of Epidemiological Lectures Online: Scientific Evaluation of Peer Review,' *Croatian Medical Journal* 48, 2 (2007): 249–55
97. Yu Liping, Pan Yuntao, Yang Chun, and Wu Yishan, 'Study on Peer Review and Multi-Indicators Evaluation in Scientific and Technological Assessment,' *KAM '08 Proceedings of the 2008 International Symposium on Knowledge Acquisition and Modeling* (2008),

available at http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=4732938&url=http%3A%2F%2Fieeexplore.ieee.org%2Fstamp%2Fstamp.jsp%3Ftp%3D%26arnumber%3D4732938 (accessed 5 August 2014)

98. Suzanne Lippert, Michael Callaham, and Bernard Lo, 'Perceptions of Conflict of Interest Disclosures among Peer Reviewers,' *PlosOne* 6, 11 (2011), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0026900 (accessed 8 May 2014)
99. Wendy Lipworth, Ian Kerridge, Stacy Carter, and Miles Little, 'Journal Peer Review in Context: A Qualitative Study of the Social and Subjective Dimensions of Manuscript Review in Biomedical Publishing,' *Social Science and Medicine* 72, 7 (2011): 1056–63
100. John Loannidis, 'Effect of the Statistical Significance of Results on the Time to Completion and Publication of Randomized Efficacy Trials,' *Journal of the American Medical Association* 279, 4 (1998): 281–6
101. Lee Lyman, 'A Three-Decade History of the Duration of Peer Review,' *Journal of Scholarly Publishing* 44, 3 (2013): 211–20
102. Bruce Macfarlane and Ming Cheng, 'Communism, Universalism and Disinterestedness: Re-Examining Contemporary Support among Academics for Merton's Scientific Norms,' *Journal of Academic Ethics* 6 (2008): 67–68
103. Gunther Maier, 'Impact Factors and Peer Judgment: The Case of Regional Science Journals,' *Scientometrics* 69, 3 (2006): 651–67
104. Herbert Marsh, Nigel Bond, and Upali Jayasinghe, 'Peer Review Process: Assessments by Applicant-Nominated Referees Are Biased, Inflated, Unreliable and Invalid,' *Australian Psychologist* 42, 1 (2007): 33–8
105. Michael Martin, andrea Kopstein and Joy Janice, 'An Analysis of Preliminary and Post-Discussion Priority Scores for Grant Applications Peer Reviewed by the Center for Scientific Review at the NIH,' *Plosone* 5, 11 (2010), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0013526 (accessed 5 November 2014)
106. Rogerio Meneghini, Abel Packer, and Lillian Nassi-Calò, 'Articles by Latin American Authors in Prestigious Journals Have Fewer Citations,' *PlosOne* 3, 11 (2008), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0003804 (accessed 20 December 2014)

107. Robert Merton, 'The Matthew Effect in Science,' *Science* 159, 3810 (1968): 59–63
108. Ian Mitroff and Daryl Chubin, 'Peer Review at the NSF: A Dialectical Policy Analysis,' *Social Studies of Science* 9, 2 (1979): 199–232
109. Ian Mitroff, 'Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists,' *American Sociological Review* 39 (1974): 579–95
110. Fang Qing, Xu Lifang and Lian Xiaochuan, 'Peer-Review Practice and Research for Academic Journals in China,' *Journal of Scholarly Publishing* 39, 4 (2008): 417–27
111. Adrian Mulligan, Louise Hall, and Ellen Raphael, 'Peer Review in a Changing World: An International Study Measuring the Attitudes of Researchers,' *Journal of American Society for Information Science and Technology* 64, 1 (2013): 132–61
112. Rüdiger Mutz, Lutz Bornmann, and Daniel Hans-Dieter, 'Heterogeneity of Inter-Rater Reliabilities of Grant Peer Reviews and Its Determinants: A General Estimating Equations Approach,' *PlosOne* 7, 10 (2012), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0048509 (accessed 3 April 2014)
113. Greg Myers, 'Texts as Knowledge Claims: The Social Construction of Two Biology Articles,' *Social Studies of Science* 15, 4 (1985): 593–630
114. Jörg Neufeld and Markus von Ins, 'Informed Peer Review and Uninformed Bibliometrics?' *Research Evaluation* 20, 1 (2011): 31–46
115. Larry Osborne and William Purkey, 'A Model Faculty Peer Review Process for Counselor Education Programs,' *Journal of Counseling and Development* 73, 6 (1995): 654–8
116. Jeff Payne, 'An Experimental Examination of the Peer Review Process,' *Research in Accounting Regulation* 16 (2003): 209–25
117. Jean-Pierre Pierie, Henk Walvoort and A. John Overbeke, 'Readers' Evaluation of Effect of Peer Review and Editing on Quality of Articles in the Nederlands Tijdschrift Voor Geneeskunde,' *Lancet* 348 (1996): 1480–3
118. David Pontille and Didier Torny, 'The Controversial Policies of Journal Ratings: Evaluating Social Sciences and Humanities,' *Research Evaluation* 19, 5 (2010): 347–60
119. Al Porter and Eugene Chubin, 'An Indicator of Cross-Disciplinary Research,' *Scientometrics* 8, 3–4 (1984): 161–76

120. Fang Qing, Xu Lifang, and Lian Xiaochuan, 'Peer-Review Practice and Research for Academic Journals in China,' *Journal of Scholarly Publishing* 39, 4 (2008): 417–27
121. Azzura Ragone, Katsiaryna Mirylenka, Fabio Casati, and Maurizio Marchese, 'On Peer Review in Computer Science: Analysis of Its Effectiveness and Suggestions for Improvement,' *Scientometrics* 97 (2013): 317–56
122. Martin Reinhart, 'Peer Review of Grant Applications in Biology and Medicine, Reliability, Fairness, and Validity,' *Scientometrics* 81, 3 (2009): 789–809
123. Fanny Rinck, 'L'analyse linguistique des enjeux de connaissance dans le discours scientifique,' *Revue d'anthropologie des Connaissances* 4, 3 (2010): 427–50
124. Ed Rinia, Thed van Leeuwen, Hendrik van Vuren, and Anthony van Raan, 'Comparative Analysis of a Set of Bibliometric Indicators and Central Peer Review Criteria: Evaluation of Condensed Matter Physics in The Netherlands,' *Research Policy* 27 (1998): 95–107
125. Paul Roebber and David Schultz, 'Peer Review, Program Officers and Science Funding,' *Plosone* 16, 4 (2011), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0018680 (accessed 24 November 2014)
126. Scott Rosas, Jonathan Kagan, Jeffrey Schouten, Perry Slack, and William Trochim, 'Evaluating Research and Impact: A Bibliometric Analysis of Research by the NIH/NIAID HIV/AIDS Clinical Trials Networks,' *PlosOne* 6, 3 (2011), available at http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0017428 (accessed 2 March 2013)
127. Peter Rothwell and Christopher Martyn, 'Reproducibility of Peer Review in Clinical Neuroscience: Is Agreement between Reviewers Any Greater Than Would Be Expected by Chance Alone?,' *Brain* 123 (2000): 1964–9
128. Terrence Shaneyfelt, Michel Mayo-Smith, and Johann Rothwangl, 'Are Guidelines Following Guidelines? The Methodological Quality of Clinical Practice Guidelines in the Peer-Reviewed Medical Literature,' *Journal of the American Medical Association* 281, 20 (2009): 1901–5
129. Mona Shattell, Peggy Chinn, Sandra Thomas, and Richard Cowling, 'Authors' and Editors' Perspectives on Peer Review Quality in Three

Scholarly Nursing Journals,' *Journal of Nursing Scholarship* 42, 1 (2010): 58–65
130. Stnaley Siegelman, 'Assassins and Zealots: Variations in Peer Review,' *Radiology* (1991): 637–42
131. Lee Sigelman and Marcia Whicker, 'Some Implications of Bias in Peer Review: A Simulation-Based Analysis,' *Social Science Quarterly* 68, 3 (1986): 494–509
132. Christopher Sigouin and Alejandro Jadad, 'Awareness of Sources of Peer- Reviewed, Research Evidence on the Internet,' *Journal of the American Medical Association* 287 (2002): 2867–9
133. Ina Spiegel-Rösing, 'Science Studies: Bibliometric and Content Analysis,' *Social Studies of Science* 7, 1 (1977): 97–113
134. Arthur Stamps, 'Advances in Peer Review Research: An Introduction,' *Science and Engineering Ethics* 3 (1997): 3–10
135. William Starbuck, 'How Much Better Are the Most-Prestigious Journals?,' *Organization Science* 16, 2 (2005): 180–200
136. John Swales, 'Occluded Genres in the Academy: The Case of the Submission Letter,' in E. Ventola and A. Mauranen, eds., *Academic Writing: Intercultural and Textual Issues* (Amsterdam: John Benjamins 1996): 45–58
137. Cassidy Sugimoto and Blaise Cronin, 'Citation Gamesmanship: Testing for Evidence of Ego Bias in Peer Review,' *Scientometrics* 95 (2013): 851–62
138. María Tavares, 'El Peer Review de las revistas científicas en humanidades y ciencias sociales: políticas y prácticas editoriales declaradas,' *Revista Española de Documentación Científica* 34 (2011): 141–64
139. Sandra Thomas, 'Conceptual Debates and Empirical Evidence About the Peer Review Process for Scholarly Journals,' *Journal of Professional Nursing* 27, 3 (2011): 168–73
140. Herbert Van De Sompel, 'What Functions Do We Take for Granted in Print?' *Nature* (2006), available at http://www.nature.com/nature/peerreview/debate/nature05008.html (accessed 30 September 2013)
141. Anthony van Raan, 'Advanced Bibliometric Methods As Quantitative Core of Peer Review Based Evaluation and Foresight Exercises,' *Scientometrics* 36, 3 (1996): 397–420

142. Susan van Rooyen, Nick Black, and Fiona Godlee, 'Development of the Review Quality Instrument (RQI) for Assessing Peer Reviews of Manuscripts,' *Journal of Clinical Epidemiology* 52, 7 (1999): 625–9
143. Matthew Verleger, Heidi Diefes-Dux, Matthew Ohland, Mary Besterfield-Sacre, and Sean Brophy, 'Challenges to Informed Peer Review Matching Algorithms,' *Journal of Engineering Education* 99, 4 (2010): 397–408
144. Elizabeth Walsh, Maeve Rooney, Louis Appleby, and Greg Wilkinson, 'Open Peer Review: A Randomised Controlled Trial,' *British Journal of Psychiatry* 176 (2000): 47–51
145. Ludo Waltman, Nees van Eck, Thed van Leeuwen, Martin Visser, and Anthony van Raan, 'On the Correlation between Bibliometric Indicators and Peer Review: Reply to Opthof and Leydesdorff' (2011), available at http://arxiv.org/abs/1105.5316 (accessed 3 November 2014)
146. Ellen J. Weber, Patricia Katz, Joseph F. Waeckerle, and Michael L. Callaham, 'Author Perception of Peer Review: Impact of Review Quality and Acceptance on Satisfaction,' *Journal of the American Medical Association* 287, 21 (2002): 2790–3
147. Michael Wood, Martyn Roberts, and Barbara Howell, 'The Reliability of Peer Reviews of Papers on Information Systems,' *Journal of Information Science* 30, 1 (2004): 2–11
148. Paul Wouters, 'Citation Cycles and Peer Review Articles,' *Scientometrics* 38, 1 (1997): 39–55
149. Mahmoud Mohamed Abdelfattah Youssef, 'Peer Review of Manuscripts Submitted to Medical Journals,' *Middle East Fertility Society Journal* 17, 2 (2012): 139–43
150. Harriet Zuckerman and Robert Merton, 'Patterns of Evaluation in Science: Institutionalisation, Structure and Functions of the Referee System,' *Minerva* 9, 1 (1971): 66–100