

An Investigation of Retracted Articles in the Biomedical Literature

John M. Budd

University of Missouri
303 Townsend Hall
Columbia, Mo 65211
Phone: 573.882.3258
Fax: 573.884.4944
Email: buddj@missouri.edu

Zach Coble

Digital Scholarship Specialist
New York University Libraries
New York, NY 10012
Email: zach.coble@nyu.edu

Alison Abritis

Retraction Watch
New York, NY
Email: aabritis@gmail.com

from the years 2010 through 2014 (or a total of 1,307 papers), and (b) examination of the citations to articles retracted in the years 2001 through 2005 (a total of 265 retracted papers and approximately 6,400 citations). This project builds upon previous work and includes categorizations both of the retraction statements and of the citations (the previous work is described below). The import of these retractions for scientific communication is examined.

ABSTRACT

A major challenge to formal scientific communication is the retraction of published works. This study includes a detailed analysis of retracted articles in biomedical literature, including categorization of the reasons for retraction. The examination covers the years 2010 through 2014. Analysis also includes citations to articles retracted between 2001 and 2005. The totality of the investigation is couched within the context of communication in the biomedical sciences and, to a lesser extent, of the formulation of theories of citation.

Keywords

Scientific communication; retraction of publication; citation analysis

INTRODUCTION

At times, when something has been determined to be seriously problematic with a published article, that article might be officially retracted. This action means that the article should no longer be considered as valid research. A paper may be retracted by one or more of the authors, by the journal's editor or publisher, or by a third party. When retraction occurs the paper is marked as being retracted in databases such as MEDLINE. Any search—by subject, keyword, title word, or author—word usually yields the paper clearly marked as being retracted, usually along with the statement of retraction. The act of retraction signals to anyone retrieving the document that the research conducted should not be treated as valid.

The present project entails a twofold purpose: (a) examination of articles that have been retracted

BACKGROUND: CITATION THEORY

Some years ago, Harriet Zuckerman (1987) wrote, "By now, it may be redundant to say that a theory of citation is badly needed" (p. 336). The claim should be taken on its face, even though there are efforts at theory formulation that pre-date Zuckerman's statement. For example, Norman Kaplan addressed norms of citation as early as 1965, and Eugene Garfield wrote about reasons for citing in that same year. The reasons for citing were reprinted in a later publication by Garfield (1994):

1. Paying homage to pioneers.
2. Giving credit for related work (homage to peers).
3. Identifying methodology, equipment, etc.
4. Providing background reading.
5. Correcting One's own work.
6. Correcting the work of others.
7. Criticizing previous work.
8. Substantiating claims.
9. Alerting researchers to forthcoming work.
10. Providing leads to poorly disseminated, poorly indexed, or uncited work.
11. Authenticating data and classes of fact—physical constants, etc.
12. Identifying original publications in which an idea or concept was discussed.
13. Identifying the original publication describing an eponymic concept or term as, e.g., Hodgkin's disease, Pareto's Law, Friedel-Crafts Reaction, etc.

14. Disclaiming work or ideas of others (negative claims).
15. Disputing priority claims of others (negative homage) (pp. 451-52).

Others, such as Emanuela Rivera (2013) suggest novel was of looking at the function of citations in scientific literature. She (2013) suggests that, following Luhmann's social systems theories, that citations constitute autopoietic systems, wherein "Citation are assumed to play a fundamental role in the scientific communities viewed as self-organizing and self-referential system" (p. 1443). An essential aspect of autopoietic systems are that they are "organizationally closed but structurally open" (Rivera, 2013, p. 1444). An element of the systemic nature of this relationship is that meaning is generated interactionally, and that meaning entails examining many possibilities among alternative possibilities" (Rivera, 2013, p. 1444). There are indeed possibilities when it comes to citing particular works; Garfield's reasons for citing provide some of the complexity that can affect decision making. Rivera (2013) sums up by saying that, "Conceiving of science as an autopoietic cognitive domain, and communicative events as the elements of this system, allows us to interpret citations as devices through which the structuration process in scientific communities is accomplished (p. 1449).

A couple of other writings on citation theory will be mentioned here. For example, according to Wouters (1999), a citation theory should be able to: "1. Provide a theoretical foundation for citation analysis; justify, in general terms, the use of S&T indicators in science policy; 3. provide a theoretical explanation of the citing behavior of scientists and scholars" (p. 561). Wouters (1999) continues that one form of peer review "is the assessment at regular intervals, of research groups, university institutions and the national contributions to specialties as a whole. These evaluations are again based on expert judgment" (p. 565). That said, bibliometric indicators, according to Wouters (1999), provide another mode of assessment—a tangible one—that is exercised by the community. The opinion of the individual scientist may be diminished in such a collective mode of evaluation.

There is a characteristic of these rather emblematic efforts at deriving theories of citation. They are, to varying degrees, idealistic. Underlying, but usually unstated, assumptions, include a belief that citers act out of very specific motivations that can be reflected in the work of

Garfield and others. There is also an assumption that citers evaluate the merits of works and determine, on the basis of the evaluation, which works to cite. As is mentioned above, the assumptions frequently are not made explicit; the tacit nature of the assumptions introduce shortcomings to many theoretical stances. How can we account for some perversions to the ideal of the social system? That is, can a theory be developed that account for social underhandedness that includes, for examples, the reasons for the retraction of some scientific works (misconduct, in particular) and the continued citation of those problematic works?

RETRACTION AS A PHENOMENON

In recent years a substantial amount of attention and analysis has been devoted to retractions, their categorization, and their subsequent citations. A major concern has been on the ethics of producing works requiring retraction. S. M. Yentis (2010), editor of the journal, *Anaesthesia*, says that this is something related to, but distinct from research ethics; it is a matter of publication ethics. The two forms of ethics are, however, ineluctably connected; in most instance the research lapses are related to the desire, or perceived need, to publish. Therefore it is difficult to make a clear distinction between the two sub-branches of ethics. Steen (2011) concentrates on one specific instance of retraction, that of Wakefield and colleagues, in which a connection between certain vaccines is related to instances of autism. Steen (2011) points out,

In the United Kingdom, there were no mumps cases in 2003 but widespread vaccine rejection led to 63,500 mumps cases in 2005. In the United States, the MMR [mumps-measles-rubella] vaccination rate decreased from 93% before the publicity to 79% in 2003. Subsequently, there was a mumps outbreak in the United States, and the number of cases in 2006 was 21-fold higher than in the prior year (p. 2).

The incident Steen writes of is a stark example of the effects of a breach of ethics and the toll it can take on humans. The "debate" over vaccination continues to be a public health issue. One may infer an action undertaken by many journals regarding policies that both alert authors to retraction and point to what actions will be taken when retraction is deemed necessary. Resnik (2015) and colleagues provide indication that retraction policies have become much more

common: “The most important finding of our study is that 3 times as many journals in our sample (65%) had a retraction policy compared to an earlier study in a similar group of journals (21%)” (p. 138). The policies are apparently necessary, since the numbers of retractions are increased year over year (see Steen and colleagues, 2013).

The phenomenon of retraction can lead to certain frustrations. Van Norden (2011) speaks to the problem: “Other frustrations include opaque retraction notices that don’t explain why a paper has been withdrawn, a tendency for authors to keep citing retracted papers long after they’ve been red-flagged” (p. 26). Van Norden (2011) provides evidence for the growing number of retractions, especially since the year 2001. According to Fanelli (2009), a weighted average of nearly 2% of scientists admitted to falsifying or fabricating data in their publications. The percentage may not seem high, but it is a problematic amount of falsified work. It is also a higher level than is represented by retractions. In a more recent work, Fanelli (2013) writes, “There is no conclusive evidence that the prevalence of misconduct is higher today than in the past. However, the scientific profession is changing in ways that, growing evidence suggests, could increase the expression of unconscious biases, questionable practices, and possible misconduct (p. 5). Whether the incidences of misconduct are increasing, or there is a greater sensitivity to the kinds of misconduct that may exist, the fact remains that the numbers of retractions have been increasing in recent years. That this is cause for concern is reflected in the numbers of citations received by the retracted publications.

Fanelli (2015) and colleagues have also examined reasons for misconduct and factors that lead to the retraction of publication. They conclude that such things as pressure to publish and the gender of authors do not play a significant role in the phenomenon. Rather, the presence and the efficacy of misconduct policies on campuses appear to be the most important factors. They write that “our results suggest that policies to reduce pressures to publish might be, as currently conceived, ineffective, whereas establishing policies and structures to handle allegations of scientific misconduct, promoting transparency and mutual criticism between colleagues, and bolstering training and mentoring of young researchers might best protect the integrity of future science” (p. 14). Their findings notwithstanding, notable figures guilty of

misconduct—such as John Darsee and Dieterick Stapel—cite the pressure to publish as an important motivator to falsify and fabricate data.

SOME PREVIOUS WORK

There have been some rather extensive studies of the phenomenon of retraction and also the numbers and kinds of citations received by those retracted articles. Budd and colleagues (1999) examined retraction revealed by a search of the MEDLINE database (of papers published from 1966 through 1997). The 235 retracted articles were cited more than 2,000 times, and very few acknowledged that papers had been retracted. A substantial number were positive citation that treated the retracted papers as substantively valid. More recent work by Budd and colleagues (2011) extends the study, particularly of retraction statements, to the year 2009. Misconduct and presumed misconduct accounted for 58% of the retractions. Results regarding citations to retracted articles mirrored those of the earlier study; few acknowledged the retraction and many were substantively positive.

Wager and Williams (2011) have also analysed retractions to biomedical papers. Their study Reports an increase in the number of retractions (year-to-year) from 1988 to 2008). As is the case with the present study, the authors use the language of the retraction statements to arrive at the categorization of retractions (see below for a description of the method used in this study).

THE PRESENT STUDY

This study extends the previous studies, noted above. In particular, the present study entails examination of retraction statements from 2010 through 2014, and citations to retracted publications from 2001 through 2005. This structure emulates Budd, Coble, and Anderson (2011), where more recent retracted items were analysed and earlier citation patterns were examined. The investigation of the retraction statements is undertaken in order to discern reasons for retraction (misconduct, error, or other reasons). As is the case with previous studies, the examination of citations begins one year after the appearance of the retraction statement. This allow some time for articles in the publication pipeline to appear (where authors probably could not be aware of the retraction).

The retraction statements are examined in detail in order to determine who issued the statement. The investigation is also intended to categorize the nature of the retraction: misconduct, unavoidable error, tainted biological sample, duplication of publication, or other reasons. The categories of the statements can then be compared with those of other studies to discover whether there are consistencies of reasons. The citations to the retracted items are studied within the context of the citing paper. This means of examination can uncover the usage of the retracted paper: acknowledgement of the retraction, tacit positive citation (mention of the paper as, for example, part of a literature review), or substantive positive citation (mention as part of usage in methodology, results, or other meaningful way). These results can also be compared to those of other studies.

RETRACTIONS

The retracted items from 2010 through 2014 are analysed for the reasons for retractions. The data that are analysed are drawn from a search of the database MEDLINE. The database has a subject heading "retraction of publication;" using this term for searching yields the items that are analysed, both for the recent retractions and for the items whose citations are analysed. Using this tactic would allow anyone who would also analyse retractions in the biomedical literature to replicate the processes employed here. For the study of retracted items, MEDLINE yields the statements of retraction as well, so the reasons for retractions can be analysed as they appear in the retracting journals. Again, follow-up or extension of the study is made possible by the use of this means of data gathering. At some times the retractions are due to unavoidable error, and the authors are doing the ethical thing to retract the paper. Error can occur in a number of ways, including data collection and analysis, contaminated samples being received from third parties, or other unavoidable reasons. These instances have nothing to do with misconduct, and should not be mistaken for misconduct.

On the other hand, there are numerous instances of misconduct. As is apparent from the data below, the reasons for misconduct can be varied. Duplication is a major cause for retraction. So can falsification and fabrication, as well as plagiarism. There are no excuses for such actions and they certainly are not unavoidable. The people retracting such items are usually one or

more of the authors (when misconduct is discovered by the individuals) or by editors of the journals in question. The data reveal that misconduct is a serious problem and one that contaminates the literature.

2010

Not Available: 25
 Error, problems with method, analysis, interpretation: 20
 Error, problems with data: 25
 Error, problems with sample: 5
 Unable to reproduce results: 21
 Scientific misconduct, Plagiarism: 32
 Scientific misconduct, Duplication: 52
 Scientific misconduct, Fabrication/falsification/manipulation: 22
 Scientific misconduct, Submitted without consent of governing organization: 4
 Scientific misconduct, Authorship issues: 8
 Scientific misconduct, Conflict of interest: 0
 Scientific misconduct, Other: 0
 Scientific misconduct, presumed: 34
 Accidental duplicate publication, authors' fault: 3
 Accidental duplicate publication, editors'/publishers' fault: 14
 Unclear: 9
 No reason given: 11
 Other/idiosyncratic: 5
Total: 290

2011

Not Available: 21
 Error, problems with method, analysis, interpretation: 19
 Error, problems with data: 17
 Error, problems with sample: 2
 Unable to reproduce results: 20
 Scientific misconduct, Plagiarism: 54
 Scientific misconduct, Duplication: 79
 Scientific misconduct, Fabrication/falsification/manipulation: 39
 Scientific misconduct, Submitted without consent of governing organization: 95
 Scientific misconduct, Authorship issues: 3
 Scientific misconduct, Conflict of interest: 3
 Scientific misconduct, Other: 2
 Scientific misconduct, presumed: 29
 Accidental duplicate publication, authors' fault: 0
 Accidental duplicate publication, editors'/publishers' fault: 11
 Unclear: 2
 No reason given: 44
 Other/idiosyncratic: 22

Total: 462

2012

Not Available: 9
Error, problems with method, analysis, interpretation: 35
Error, problems with data: 17
Error, problems with sample: 1
Unable to reproduce results: 22
Scientific misconduct, Plagiarism: 53
Scientific misconduct, Duplication: 131
Scientific misconduct, Fabrication/falsification/manipulation: 132
Scientific misconduct, Submitted without consent of governing organization: 16
Scientific misconduct, Authorship issues: 15
Scientific misconduct, Conflict of interest: 3
Scientific misconduct, Other: 26
Scientific misconduct, presumed: 62
Accidental duplicate publication, authors' fault: 0
Accidental duplicate publication, editors'/publishers' fault: 16
Unclear: 1
No reason given: 46
Other/idiosyncratic: 33
Total: 618

2013

Not Available: 21
Error, problems with method, analysis, interpretation: 36
Error, problems with data: 48
Error, problems with sample: 4
Unable to reproduce results: 15
Scientific misconduct, Plagiarism: 53
Scientific misconduct, Duplication: 136
Scientific misconduct, Fabrication/falsification/manipulation: 163
Scientific misconduct, Submitted without consent of governing organization: 16
Scientific misconduct, Authorship issues: 18
Scientific misconduct, Conflict of interest: 0
Scientific misconduct, Other: 5
Scientific misconduct, presumed: 48
Accidental duplicate publication, authors' fault: 0
Accidental duplicate publication, editors'/publishers' fault: 0
Unclear: 0
No reason given: 36
Other/idiosyncratic: 14
Total: 613

2014

Not Available: 9

Error, problems with method, analysis, interpretation: 52
Error, problems with data: 45
Error, problems with sample: 6
Unable to reproduce results: 23
Scientific misconduct, Plagiarism: 65
Scientific misconduct, Duplication: 67
Scientific misconduct, Fabrication/falsification/manipulation: 68
Scientific misconduct, Submitted without consent of governing organization: 15
Scientific misconduct, Authorship issues: 11
Scientific misconduct, Conflict of interest: 2
Scientific misconduct, Other: 10
Scientific misconduct, presumed: 52
Accidental duplicate publication, authors' fault: 0
Accidental duplicate publication, editors'/publishers' fault: 10
Unclear: 0
No reason given: 43
Other/idiosyncratic: 30
Total: 508

We can see from the above data that 2,491 articles were retracted in a relatively short period of time. This figure can be compared with that found in Budd, Coble, and Anderson (2011): 1,164 retractions from 1997 to 2009. The increase in retraction in recent years, as opposed to the immediate previous time period, is striking. Corbyn, (2009) also noted an increase in the occurrence of retractions. What cannot be determined is whether there are greater instances of such things as error and misconduct, or there is a greater sensitivity to those phenomena (and a related increase in reporting). It is a fact that the numbers of retracted papers represent a small portion of all articles published in the biomedical literature, but the number is still troubling. What may be most troubling is that 64.8% of the retractions are due to misconduct or presumed misconduct. What may be more troubling is that the number has been increasing in recent years. Budd, Coble, and Anderson (2011) report that, in their study, only 55% of retractions were due to misconduct. The reasons for the increases can only be a matter for speculation, but some stated rationales include the pressure in the academy for faculty to publish more and more. That, however, is an excuse and not a reason for the action.

The categorization of misconduct is derived directly from the statements that are published in

the retracting journals. An example of a clear statement of retraction is:

The undersigned persons, being the authors of the article herein specified, of their own free will have decided to retract the said article. This article is a limited review of articles previously published in Brazilian scientific journals. It contains no reports on newly obtained scientific results and was written merely for the information of readers of Revista Brasileira de Cirurgia Cardiovascular. The reason for retraction is that an honest error was committed in the method used for collecting articles to be included. The retraction is formally approved by the Editor of Revista Brasileira de Cirurgia Cardiovascular/Brazilian.

At times the “presumed misconduct” category contains statements that indicate actual misconduct:

This article has been retracted at the request of the authors because there is no clarity about the authorship of this paper.

At times there is clear evidence of plagiarism; here is an example:

This article has been retracted at the request of the editor as the authors have plagiarised part of several papers that had already appeared in several journals. One of the conditions of submission of a paper for publication is that authors declare explicitly that their work is original and has not appeared in a publication elsewhere. Re-use of any data should be appropriately cited. As such this article represents a severe abuse of the scientific publishing system. The scientific community takes a very strong view on this matter and we apologise to the readers of the journal that this was not detected during the submission process. From a limited, non-exhaustive check of the text, several elements of the text had been plagiarised from the following list of sources: Dihydroxyacetone — Wikipedia, the free encyclopedia”

Every statement of retraction is carefully examined so that the classification of the retractions can be made with confidence. This is a time-consuming process, but there is no

alternative to the close reading of the statements. In many ways, this mode of classification relies upon content analysis. Klaus Krippendorff (1980) provides one of the most thorough guides to this methodology. The content analysis provides the methodological guide to the creation of the categories that are mentioned above. That is, the retraction statements include the reasons for the retraction in almost all instances. Where there is any ambiguity relating to classification, the retraction statement provides clear clues as to the actual reasons for the retraction. In this way, content analysis is the appropriate method for categorizing the reasons for retraction.

CITATIONS

Examination of the phenomena of retractions is half of the complete analysis of the complexity of scientific communication. The foregoing does display some of the most serious problems that plague the biomedical literature, especially with respect to misconduct. The other half of the analysis has to do with citations received by the retracted items. The question to be asked is: Do authors cite retracted papers as though those retracted items represented valid work? The second part of the present study attempts to answer that question. This portion is a detailed examination of citations to retracted articles to determine whether the practice of citation is, to some extent, oblivious to the phenomenon of retraction. To accomplish this objective, the literature that is analysed is earlier than that studied for the instances of retraction. The reason for doing the analysis in this manner is to allow for a record of citations to be present for analysis. Another reason is that this tactic builds upon previous analyses (see below).

The analysis of citations followed some of the same methodological considerations as those employed in the identification of retraction statements. That is, the MEDLINE database (via OVID) was searched, using the “retraction of publication” subject heading. SCOPUS was also used to retrieve source items. This lends consistency to the two phases of this study. The total number of retracted articles, as is noted below, is 265. In order to examine the citations to the retracted articles, and to attempt to ensure that there could be awareness of the retraction, especially if the citing author searched the MEDLINE database for secondary literature, the analysis begins with citations dating one year *after* the appearance of the retraction statement.

For example, if the statement of retraction appears in 2001, the citations examined date from articles published in the year 2002 or later. If the MEDLINE database were searched by the citing authors, the retraction statement's reference would appear along with the reference to the original retracted article.

Tacit positive citations usually are reflected by mention of the retracted paper in a way that can be presumed to be accepting of the content. At times the retracted article is cited along with other articles. Here is an example:

Not all dogs with brain tumors will develop seizures, hence certain factors must exist that promote epileptogenesis: however the pathogenesis of tumour-related seizures is poorly understood (Beaumont and Whittle 2000, Schaller and Rüegg 2003, van Beemen and others 2007, Shamji and others 2009).

Substantive uses of the content of the retracted paper, for employment as methodology, affirmation of results, or other usage that signals a genuine usage of the retracted paper. One example is:

Recently, the leucocyte-associated immunoglobulin-like receptor 1 (LAIR-1) was identified as a novel extracellular ligand of EpCAM (Meygaard, et al, 2001).

Occasionally, but infrequently, the retraction of an article is noted in a citing article. The notation is usually straightforward, such as:

The initial publication by Meygaard, et al. has been retracted.

The present study, as is mentioned above, includes an examination of citations covering the years 2001-2005 to retracted articles. Every retracted paper was examined using the database Scopus ® and the citations to each retracted paper was analysed. The citations were recorded according to the above categorizations. The following presents the analysis:

Total Articles:	265
Tacit Cites:	3,946 (80.25%)

Substantive Cites:	767 (15.60%)
---------------------------	--------------

Retraction Noted:	204 (4.15%)
--------------------------	-------------

Fifteen of the articles had no citations, but of the 250 that did, the mean citations per retracted article is 19.67. A total of 1,424 citing papers could not be analysed; the principal reason for this is that the citing papers were published in languages other than English and were not included in databases that were available for examination. One other phenomenon should be mentioned; there were 64 instances of self-citation. The vast majority of the self-citations did not make mention of the fact that the cited article had been retracted.

It is our contention that these results uncover a profound problem with scientific communication. Since the retracted items are almost always accompanied by a designation that the paper has been retracted (frequently along with the statement of retraction), anyone searching a database like MEDLINE should become aware of the retraction status of articles in question. This fact notwithstanding, the retracted articles continue to be cited at a considerable rate. More specifically, approximately 6,400 citations are received by the retracted papers. If those citing papers that could not be located are included, the mean number of citations received by each article that had citations is about 25.6. Many are cited more than 100 times each.

The numbers of citations received is cause for concern, but the reality is that almost all of the citations (roughly 96%) are positive in some way. This figure suggests that citing authors are rather uncritical when it comes to incorporating the secondary literature into their own work. More careful searching and verification would clearly indicate that source items had been retracted. With the increasing numbers of retractions, plus the increases in retractions due to misconduct, it behoves authors to take care with their own work. A small percentage acknowledge the fact that the cited paper has been retracted. One could ask why there are so many positive citations to retraction papers, but, at this point, causes are a matter of speculation. To reiterate, the data collected allow one year for retraction statements to have been published and indexed. It may be that citing authors are retrieving papers though some alternative means, such as reviewing the citations in published papers and not consulting databases

prior to including retracted items in the citing works. It is recommended that researchers do consult the databases to be as certain as possible that retracted papers are not cited.

Another matter of speculation is the life that retracted articles have in the overall literature. If the reason for retraction happens to be scientific misconduct, the continued life of the retracted items could have serious ramifications. While scientific research is supposed to be self-correcting, that premise, as we shall see, deserves scrutiny in itself. It may be possible that erroneous work or work resulting from misconduct could work their way into research that has clinical implications. A major reason for the conduct of the present work is to alert researchers and others to the potential implications of continued use of problematic research.

CITATION THEORY REVISITED

Given the results presented above, it would appear that theoretical constructs regarding citations would require some revision. For example, the reasons for citing, noted by Garfield (1965), do not capture all of the characteristics of citing behavior. The assumptions that underlie most theories of citation action appear to have a commonality regarding intentionality. Searle (1983) expounds upon the nature of intentionality and emphasizes that having a belief or desire directed *at* something. He writes, “Intentionality is that property of many mental states and events by which they are directed at or about objects and states of affairs in the world” (p. 1). Along with the directedness is that beliefs embody truth conditionals. The continued citation of retracted articles indicates that the belief structure related to citation betrays the truth condition. What this means is that citation behavior may not be beholden to the tenets of intentionality.

Another challenge to the scientific literature is the seeming reality of irreproducible results. As Gunn says,

From alarming estimates derived from studies by Bayer (F. Prinz et al. *Nature Rev. Drug Discov.* 10, 712; 2011) and Amgen (C. G. Begley and L. M. Ellis *Nature* 483, 531–533; 2012) that some 60–70% of biomedical research papers may contain irreproducible results, it would seem that our time would be better spent investigating experimental

irreproducibility rather than hunting down fraudsters (p. 483).

This problem signals another concern for citation theory; if results in any of the sciences cannot be reproduced then there should be some allowance in theory. Particularly, citation theory should include consideration of the documents that can be cited. In other words, citation theory in the future will have to include more than just the reference to works, but the works themselves. There is a need for a scholarly communication theory, accompanied by a body of empirical research, such as that presented here.

Further, the theories that address information can apply to this matter as well. To use just one example, Budd (2011) incorporates meaning and truth into a statement of theory of information. Again, citation practice is not in keeping with this theoretical expression. A revised theory of citations (which will not be presented here) will have to incorporate the non-intentional and non-truth conditional elements of actual behavior. The results presented here show that these characteristics are not necessarily behavioral aberrations. Whether a future theory concentrates on what citation action *ought* to be, or on what the action *is*, the intentional and informational groundings should be addressed.

SOME FUTURE WORK

The retractions and citation data have considerable import for the development of theories of citation activity and, more broadly, for scholarly communication in general. The data presented here can help inform that theory.

More specifically, future work, particularly on retractions, can add examination of whether the retracted items received funding from any agencies that provide support for research. This analysis of data can be especially informative when it comes to papers where misconduct is present. It may be that a substantive sum of money has gone to researchers who have been guilty of one of the forms of misconduct characterized in this paper. It may be that funding agencies would adjust practices if it were known that some of their funding was expended on work shown to be guilty of misconduct.

John M. Budd is a Professor Emeritus with the University of Missouri. He has worked on retractions to scientific literature for a number years.

Zach Coble is a Digital Information Specialist with New York University. He has collaboration with John Budd on a previous study of retracted items, presented at the 2011 ACRL National Conference.

Alison Abritis is a researcher with Retraction Watch in New York City. Her work involves many projects dealing with retractions in the scientific literatures.

REFERENCES

Budd, J. M. (2011). Meaning, truth, and information: Prolegomena to a theory. *Journal of Documentation*, **67**(1), 56-74.

Budd, J.M., Sievert, M.E., Schultz, T.R. & Scoville, C. (1999). Effects of article retraction on citation and practice in medicine. *Bulletin of the Medical Library Association*, **87**(4), 437-43.

Budd, J.M., Coble, Z.C. & Anderson, K.M. (2011). Retracted publications in biomedicine: Cause for concern. Philadelphia, PA: Association of College & Research Libraries National Conference, http://www.ala.org/acrl/sites/ala.org/acrl/files/content/conferences/confsandpreconfs/national/2011/papers/retracted_publicatio.pdf.

Corbyn, Z. (2009). Retraction up tenfold. *Times Higher Education*, <https://www.timeshighereducation.com/news/retractions-up-tenfold/407838.article>.

Fanelli, D. (1999). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One*, **4**, 1-11.

Fanelli, D., Costas, R. & Larivière, V. (2015). Misconduct policies, academic culture, and career stage, not gender or pressures to publish, affect scientific integrity. *PLoS ONE*, **10**(6), 1-18.

Fanelli, D. (2013). Why growing retractions is (mostly) a good sign. *PLoS ONE*, **10**, 1-6.

Garfield, E. (1965). Can citation analysis be automated? In *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings*, ed. by M. E. Stevens, V. E. Giuliano & L. B. Heilprin, p. 189-92. Washington, DC: National Bureau of Standards.

Garfield, E. (1994). When to cite. *Library Quarterly*, **66**(4), 449-58.

Gunn, W. (2014). Fraud is not the big problem. *Nature* **505**, 483.

Kaplan, N. (1965). The norms of citation behavior: Prolegomena to the f Van Norden, R. (2011). The trouble with retraction. *Nature*, **478**, 26-28.

Krippendorff, K. (1980). *Content analysis*. Thousand Oaks, CA: Sage.

Resnik, D. B. Wagner, E. & Kissing, G. E. (2015). Retraction policies of top scientific journals ranked by impact factor. *Journal of the Medical Library Association*, **103**(3), 136-39.

Rivera, E. (2013). Scientific communities as autopoietic systems: The reproductive function of citations. *Journal of the American Society for Information Science & Technology*, **64**(7), 1442-53.

Searle, J. R. (1983). *Intentionality: An essay on the philosophy of mind*. Cambridge: Cambridge University Press.

Steen, R. G. (2011). Retractions in the medical literature: Who is responsible for scientific integrity? *AMWA Journal*, **26**(1), 2-7.

Steen, R. G., Casadevall, A. & Fang, F. C. (2013). Why has the number of scientific retractions increased? *PLOS One*, **8**(7), 1-9.

Wager, E. & Williams, P. (2011). Why and how do journals retract articles? An analysis of MEDLINE retractions 1988-2008. *Journal of Medical Ethics* **37**(9), 567-70.

Wouters, P. (1999). Beyond the holy grail: From citation theory to indicator theories. *Scientometrics*, **44**(3), 561-80.

Yentis, S. M. (2010). Another kind of ethics: From corrections to retractions. *Anaesthesia*, **65**, 1163-66.

Zuckerman, H. (1987). Citation analysis and the complex problem of intellectual influence. *Scientometrics*, **12**(5), 329-38.