

Article

Measuring Scientific Misconduct—Lessons from Criminology

Felicitas Hesselmann¹, Verena Wienefoet¹ and Martin Reinhart^{1,2,*}

¹ Institute for Research Information and Quality Assurance, Schützenstraße 6a, 10117 Berlin, Germany; E-Mails: hesselmann@forschungsinfo.de (F.H.); wienefoet@forschungsinfo.de (V.W.)

² Humboldt-Universität zu Berlin, Institut für Sozialwissenschaften, Unter den Linden 6, 10099 Berlin, Germany

* Author to whom correspondence should be addressed; E-Mail: martin.reinhart@hu-berlin.de or reinhart@forschungsinfo.de; Tel.: +49-30-2064-11756; Fax: +49-30-2064-17799.

Received: 28 February 2014; in revised form: 26 June 2014 / Accepted: 27 June 2014 /

Published: 3 July 2014

Abstract: This article draws on research traditions and insights from Criminology to elaborate on the problems associated with current practices of measuring scientific misconduct. Analyses of the number of retracted articles are shown to suffer from the fact that the distinct processes of misconduct, detection, punishment, and publication of a retraction notice, all contribute to the number of retractions and, hence, will result in biased estimates. Self-report measures, as well as analyses of retractions, are additionally affected by the absence of a consistent definition of misconduct. This problem of definition is addressed further as stemming from a lack of generally valid definitions both on the level of measuring misconduct and on the level of scientific practice itself. Because science is an innovative and ever-changing endeavor, the meaning of misbehavior is permanently shifting and frequently readdressed and renegotiated within the scientific community. Quantitative approaches (*i.e.*, statistics) alone, thus, are hardly able to accurately portray this dynamic phenomenon. It is argued that more research on the different processes and definitions associated with misconduct and its detection and sanctions is needed. The existing quantitative approaches need to be supported by qualitative research better suited to address and uncover processes of negotiation and definition.

Keywords: scientific misconduct; scientific retractions; scientific fraud; criminology; labeling theory; methodology

1. Introduction

Studying a problem scientifically involves measuring things quantitatively [1]. Approaching a problem quantitatively promises to establish two foundations from which further research will be possible and justifiable. First a valid and reliable measurement shows the extent of the problem and helps estimating whether more scientific attention is warranted. Second, knowing the extent of a problem from scientific measurement is helpful in convincing other scientists and the public that more research on the problem is needed and should be supported (intellectually and financially). This strategy is as frequently used to introduce research on scientific misconduct as on many other topics. However, our argument will be that especially for research on scientific misconduct this strategy is problematic. It is problematic because statistics on misconduct are prone to biases well-known from Criminology, which find little attention in the literature on scientific misconduct. The measurement of the prevalence of retracted papers will serve as a case in point. The aim of this paper is, thus, to present an argument why measuring scientific misconduct quantitatively should not be first on our research agenda. Decades of research in Criminology and the Sociology of Deviance will serve as a background to support this argument.

2. Measuring Misconduct

Researchers seeking to appraise the prevalence of scientific misconduct try to do so by either looking at known instances of fraud and error, such as retracted articles, or by conducting surveys of scientists asking about their experience with misconduct. These strategies are comparable to the methods and resulting statistics in Criminology that are used to uncover the unknown “dark figure” of crimes in a population. Both the rate of retractions and the number of scientists admitting to or suspecting their colleagues of misconduct are subject to systematic errors. Because Criminology, with its long-standing tradition of investigating these errors [2], has produced notable insights into the problems associated with these methods, the knowledge gained by criminological research will serve as a reference point for our further argument.

Calculating some form of a retraction rate to assess the prevalence of misconduct constitutes a prevailing line of research [3–13]. Many studies focus on data of retracted articles retrieved via databases such as the Web of Science, PubMed, and the like. As with a crime rate, these observed data do not accurately characterize the true rate of scientific misconduct [2,14], but are the joint outcome of a number of distinct data generating processes that each produce a characteristic distortion (see also [15,16]). For the retraction rates, there are at least four different processes that need to be separated:

First, there is the actual occurrence of misconduct, the process many studies seek to explain or measure and that will result in fraudulent publications. Drawing an analogy to Criminology, this process represents the occurrence of criminal acts, which is at the heart of many theories trying to explain why some people resort to crime while others do not [2].

Second, there is the process of detection of misconduct. This process entails arising suspicions, allegations, and subsequent investigations and is comparable to the reporting of crimes to the police and the subsequent recording of these crimes by the police [15]. Thus, in addition to the crimes that

form the official police statistics, there are infractions that are never detected by anyone, crimes detected but not reported, and crimes reported but not recorded by the police [14,15,17]. Hence, crimes officially recorded by the police pose only a fraction of all crimes committed. Moreover, the crimes detected and recorded might misrepresent the structure of incidents, because some crimes or forms of misconduct can be more difficult to detect, minor infringements might simply be shrugged off and show up nowhere in the data [14,15,18], or because some people are more successful at obscuring their deeds or intimidating possible witnesses. In science, there is evidence that many researchers who witness scientific misconduct refrain from reporting the incident [19–21]. To our knowledge, no research has investigated whether reported infractions differ systematically from unreported infractions.

The third process is the retraction of a publication, which can be compared to the sentencing in criminal procedure. There are many factors that determine the outcome of a trial, of which the actual crime is only one. For the case of scientific misconduct, many factors associated with the decision to retract an article (or not) remain unclear: A number of studies show that many journals do not have clear policies for dealing with misconduct [22–25] and handle allegations inconsistently and mostly on a case-by-case basis [26–28]. Additionally, even though resources such as the Committee on Publication Ethics (COPE) guidelines do exist, a notable number of editors are unaware of these guidelines [27]. Even in cases when a finding of misconduct has been established by the Office of Research Integrity (ORI), some affected articles are neither retracted nor corrected [29–32]. Criminological research shows that the aggregate rate of incarceration in a country varies independently from the rate of crimes recorded by the police [33] and exhibits varying and mostly inconsistent patterns of correlation with other measures of crime [34]. It seems optimistic to assume that this would not hold for scientific misconduct and to conclude that the development of retractions was directly relational to the development of misconduct. Many researchers thus acknowledge the fact that there is no way of knowing whether for instance, retractions have risen over the past decades due to a rising prevalence of misconduct or a rising awareness and scrutiny.

The fourth process that ultimately influences the records of retractions and is often overlooked is the way in which journals and databases identify retracted publications once an article has been retracted. To a lesser extent, data entry errors or misidentification also influence official police records [17], but rigid regulations and guidelines for record-keeping and considerable efforts for harmonizing international data (e.g., by EUROSTAT) make this problem less severe for criminological research as compared to research on scientific misconduct. There can be major inconsistencies as to how retracted articles are marked in the databases and whether they are marked at all. It is currently far from clear how many of the publications that were formally retracted will show up as “retracted articles” in one of the databases and how many remain to appear unmarked and “valid”. Moreover, some journals may label a retraction by adding the term “retraction” to the title, while others may use different terms [29]. The number of retracted articles will, hence, vary greatly by the search term researchers use to locate retractions in the database.

We can, thus, learn from Criminology that measuring rates of misconduct should separate at least four different processes related to misconduct. What we can measure depends on the actual occurrence of misconduct (1); the detection of misconduct (2); the sentencing of misconduct (3); and the recording of misconduct (4). The problem posed by the entanglement of these four processes

reaches much further than just an underestimation of the prevalence of misconduct. As every process systematically influences the probability that a given instance of misconduct will be selected into the sample of studied retracted articles, this selection bias will result in biased estimates in a regression analysis [35]. Because these selection rules are currently unknown, both the size, as well as the direction, of this bias remain unknown, rendering the analyses mostly meaningless.

The concurrence of these processes alone makes inferences about scientific misconduct extremely difficult, however, the current measurements also suffer from the fact that misconduct seems to be a moving target. In the absence of a generally valid definition of misconduct, both official institutions [36] and researchers apply a variety of definitions with differing scopes. Hence studies of scientific misconduct often do not measure the same thing. Looking at the analysis of retractions, researchers differ, e.g., as to whether they include plagiarism as a form of misconduct [7,11,37] or not, rather defining it as a form of error [12,38]. This inevitably results in differing numbers of retractions due to misconduct (see also [39]). Accordingly, numbers of retractions are not comparable across studies. Differing definitions of prohibited acts also trouble criminological research; hence many researchers consider it futile to compare raw crime rates across countries or even within a country over long periods of time. Growing (or declining) rates of retractions may not only be caused by greater (lesser) scrutiny or propensity to act on the part of reviewers and editors [6], but also by a changing range of behaviors that are considered to be misconduct. Because there is so little information on what behaviors exactly prompt journals to retract articles [5,28], we do not know if and how this definition has changed over time.

In the current situation, given only the retractions as they appear in the databases it is impossible to untangle these influences to make any statement about the prevalence or structure of either misconduct, the risk of its detection, or the probability that detected misconduct is sanctioned. It is crucial to gain a better insight into these distinct mechanisms, both in the form of theoretical consideration and empirical evidence, before analyses based on the rate of retractions can be meaningful.

As one reaction to the outlined shortcomings of official measures of crimes, criminological research has turned to self-report measures to gauge the true extent of criminal behavior [40,41]. This strategy is also applied when examining scientific misconduct [42,43]. The problems associated with self-report measures of delinquency continue to be of great interest in Criminology, as they may result both in under-reporting [17,40,41] and over-reporting of misbehavior [17,40]. Respondents might not want to tell the truth about their actual involvement in criminal behavior [40], or they might simply have forgotten about some acts [17]. Accurate recall of own behaviors poses a very serious problem in the research of misconduct using self-reports: Questions pertaining to only vaguely defined periods (*i.e.*, “ever” [44]) or to very large time-spans (*i.e.*, “during the last 10 years” [45]; “since entering medical school” [46]) will generally not produce very reliable or, for that matter, comparable answers ([41]; see [47] for a short overview). The difference between self-reports and official records furthermore varies, among others, by number of arrests [40,41], gender, and type of offense [41], hence producing systematically biased estimates. Additionally, for the case of self-reports, the problem of definition becomes crucial. The phenomenon of over-reporting arrests is commonly seen as resulting from respondents misleadingly counting all contacts with the police as arrests, even if they were not officially arrested [17,40]. Particularly, with regard to scientific misconduct, the absence of a stable definition might result in different people reporting the same act under different names. What

might seem as falsification of data in one discipline, might be considered only sloppiness or even accepted practice in another discipline. This problem becomes obvious in the study by Titus, Wells, and Rhoades [20], where 24% of reported incidents did not actually meet the federal criteria of misconduct.

3. Defining Misconduct

The problems associated with defining scientific misconduct thus need further elaboration. Following Becker that “deviant behavior is behavior that people so label” ([48]: p. 9) and that the simplest conception of deviant behavior is a statistical one which defines everything as different, which lies too far from the average [48], presents us with problems regarding definition on two different levels. One level is represented by observers of scientific misconduct (e.g., Science Studies or administrators) and the other by the daily practices of scientists. On both levels, labeling occurs with specific restrictions but not independent of one another.

For the level of observation of misconduct the most severe restriction seems to lie in the fact that the labels for misconduct must come from scientific practice itself. Politicians, administrators or Science Studies scholars lack authority to label a specific scientific practice as deviant since only scientists themselves possess the necessary qualifications to do so. We are not talking about forms of misconduct that can be defined rather easily like outright plagiarism or fabrication here but about the many routine scientific practices that operate in a “grey zone”. For such practices to become problematic and labeled as deviant the scientific community on the level of scientific practice has to act first and only then are we able to use the same label on the level of observation. To give an example, before certain forms of image manipulation could be labeled as misconduct, e.g., by journals, funding organizations or the ORI the scientists themselves had, first, to engage in these practices and then, second, to negotiate what the community rejects as too much “beautification” [49]. On the level of observation, definitions of deviance are thus always dependent on scientific practices and, in some sense, too late [50]. This aspect raises problems especially for research on misconduct because it cannot rely on an external source for distinguishing misconduct from regular behavior. Furthermore, labels for misconduct vary from one discipline to the next, presenting Science Studies with a difficult task in establishing a precise and generalizable terminology [49,51]. This may, in part, explain the variance in terminology in research on scientific misconduct presented above.

When switching to the level of scientific practice and the ways scientists negotiate the permissible and non-permissible forms of behavior on a day-to-day basis, the picture becomes more complicated. Here, these problems of definition can pose restrictions for scientific practice in that the capacity to innovate may be affected. In Criminology a classical idea is that innovation often comes along with the rejection of institutional norms and the establishment of new forms of behavior which, over time and through negotiation, become permissible forms of behavior ([52]: p. 231). This is not to say that for innovation to occur norms must always be broken. However, there is a tradeoff between rigid control of behavior according to institutional norms and the possibilities for innovation. Especially for the case of science, where we expect scientists to deliver new and innovative knowledge, that we sometimes call “revolutionary”, a context that does not allow for non-conformist approaches may stifle innovation. Only the resulting processes of negotiation separating the permissible from

the non-permissible will establish what can count as new and regular facts, theories, and methods [53]. Again, this is not to say that plagiarism, fabrication, and falsification should be tolerated but rather that efforts to prevent or sanction misconduct may have the non-intended consequence of rendering certain forms of scientific innovation more unlikely.

One can, thus, see that the problems associated with defining scientific misconduct lead to questions of labeling and ultimately to the processes of negotiation in different scientific fields about the boundaries of permissible practice. The dynamics of these processes and their specificity to certain research fields should deter from prematurely committing to overly precise or general definitions of scientific misconduct. Otherwise, research on scientific misconduct may suggest a more objective picture of scientific practice than it actually can deliver. One conclusion that can be drawn from this and from the experiences from Criminology is that more qualitative approaches are needed to analyze the processes of occurrence, detection, punishment, and recording of scientific misconduct in the places where these processes happen.

The line drawn between scientific misconduct and accepted research practice can be seen as a divide not between two types of research but essentially between science and non-science. “Boundary work” [54] can be a useful concept to study the ways scientists separate the permissible from the non-permissible. Cases of (alleged) misconduct, like the one by Cyril Burt, a famous psychologist who was posthumously accused of having falsified and fabricated data, give a rich picture of the complex social processes that lead to questionable practices and the ways these are detected and labeled as misconduct. In the case of Burt, his studies on the cognitive capabilities of twins, arguing that intelligence is mainly determined by genes, seem to have highly unlikely samples and results. However, even within Psychology, consensus has not been forthcoming on whether this qualifies as misconduct or not. It seems that cases of possible misconduct can be highly contested because they stand for more than just the one case. They are always also cases where the limits of the cognitive authority of a discipline or science in general are questioned [55]. We would thus suggest to continue existing research from Science and Technology Studies dealing with boundary-work [54], which not only broaches the issue of scientific misconduct but also the demarcation between science and non-science ([56]: p. 16). This line of research, as well as the literature on retractions and scientific misconduct, could benefit from including criminological approaches.

4. Conclusions

Making use of an analogy to criminological research and criminal statistics, this article addressed the problem of confounding different processes that influence the final number of retractions that appear in scientific databases. Moreover it discussed the problem of changing and fuzzy definitions of misconduct that can hardly be represented by quantitative data.

Because currently very little is known both about the outlined processes of occurrence, detection, punishment, and publication of misconduct and about the negotiation of definitions within the scientific fields, raw numbers of retractions or self-reported misbehaviors are not very telling. Every process systematically influences the probability that a given instance of misconduct will be selected into the sample of studied retracted articles and causes amongst others biased estimates. They will most likely

paint an inaccurate picture both of the amount and the characteristics of misbehaviors and the scientists committing them.

More research is needed to illuminate these intermediary processes. Because of the explorative nature of such research and because the objectives of research are dynamic and deeply embedded within structures of everyday action, reasoning, tacit knowledge and sense-making, the methods of choice at this point would have to be qualitative (see, e.g., [57–59]). The definitions of misconduct, for instance, cannot be addressed by imposing pre-established definitions on the phenomenon in question, as is necessary in statistical analyses, but call for an open-ended approach that aims at discovering rather than confirming the various meanings of the label “misconduct”. Qualitative research can deliver an empirically grounded overview that should serve as a starting point for further quantitative analyses.

Moreover we argue to altogether shift the research focus to address the processes of negotiation at the various levels—in the scientific community (micro), in institutions like universities and journals (meso), and in more general public discourses (macro)—more prominently. Instead of being solely conceptualized as error or noise that will taint quantitative analyses and that more research simply seeks to partial out, it could be noted that, just like the processes influencing criminal statistics, they are “an aspect of social organization and cannot, sociologically, be wrong.” ([14]: p. 734). Social processes that shape definitions of scientific misconduct are not just sources of error or noise for researchers who are trying to measure misconduct but should be seen as an important resource and topic for further research. In fact, it is these very processes of negotiating definitions and boundaries of scientifically correct and scientifically fraudulent (or erroneous) behavior that lie at the heart of the scientific enterprise and therefore should be a major focus of the study of scientific misconduct [54,60]. An analysis of these processes should not simply aim at producing more reliable statistics in the end, but seek to shed light on the structures and mechanisms of scientific knowledge production.

Acknowledgments

We would like to thank Natascha Trutzenberg for her help with retrieving and managing the literature for this paper and Marion Schmidt for bibliometric advice. This paper originates from project No. 01PY13009 funded by the German Ministry of Education and Research. The content of this paper is solely the responsibility of the authors.

Conflict of Interest

The authors declare no conflict of interest.

References

1. Porter, T.M. *Trust in Numbers*; Princeton University Press: Princeton, NJ, USA, 1996.
2. Biderman, A.D.; Reiss, A.J. On Exploring the “Dark Figure” of Crime. *Ann. Am. Acad. Pol. Soc. Sci.* **1967**, *648*, 1–15.

3. Budd, J.M.; Sievert, M.; Schultz, T.R. Phenomena of retraction. Reasons for retraction and citations to the publications. *JAMA* **1998**, *280*, 296.
4. Davis, P.M. The persistence of error: A study of retracted articles on the Internet and in personal libraries. *J. Med. Libr. Assoc. (JMLA)* **2012**, *100*, 184–189.
5. Decullier, E.; Huot, L.; Samson, G.; Maisonneuve, H. Visibility of retractions: A cross-sectional one-year study. *BMC Res. Notes* **2013**, *6*, doi:10.1186/1756-0500-6-238.
6. Fanelli, D. Why growing retractions are (mostly) a good sign. *PLoS Med.* **2013**, *10*, e1001563.
7. Fang, F.C.; Steen, R.G.; Casadevall, A. Misconduct accounts for the majority of retracted scientific publications. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 17028–17033.
8. Foo, J.Y. A Retrospective Analysis of the trend of retracted publications in the field of biomedical and life sciences. *Sci. Eng. Ethics* **2011**, *17*, 459–468.
9. Furman, J.L.; Jensen, K.; Murray, F. Governing knowledge in the scientific community: Exploring the role of retractions in biomedicine. *Res. Policy* **2012**, *41*, 276–290.
10. Grieneisen, M.L.; Zhang, M.; von Elm, E. A Comprehensive Survey of Retracted Articles from the Scholarly Literature. *PLoS One* **2012**, *7*, e44118.
11. Nath, S.B.; Marcus, S.C.; Druss, B.G. Retractions in the research literature: Misconduct or mistakes? *Med. J. Aust.* **2006**, *185*, 152–154.
12. Redman, B.K.; Yarandi, H.N.; Merz, J.F. Empirical developments in retraction. *J. Med. Ethics.* **2008**, *34*, 807–809.
13. Steen, R.G. Retractions in the scientific literature: Is the incidence of research fraud increasing? *J. Med. Ethics* **2011**, *37*, 249–253.
14. Black, D.J. Production of Crime Rates. *Am. Sociol. Rev.* **1970**, *35*, 733–748.
15. Boivin, R.; Cordeau, G. Measuring the impact of police discretion on official crime statistics: A research note. *Police Q.* **2011**, *14*, 186–203.
16. O'Brien, R.M. Police productivity and crime rates: 1973–1992. *Criminology* **1996**, *34*, 183–207.
17. Kirk, D.S. Examining the divergence across self-report and official data sources on inferences about the adolescent life-course of crime. *J. Quant. Criminol.* **2006**, *22*, 107–129.
18. Agnew, R.; Brezina, T. *Juvenile Delinquency: Causes and Control*, 4th ed.; Oxford University Press: New York, NY, USA, 2012.
19. Fanelli, D.; Tregenza, T. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS One* **2009**, *4*, e5738.
20. Titus, S.L.; Wells, J.A.; Rhoades, L.J. Repairing Research Integrity. *Nature* **2008**, *453*, 980–982.
21. Pryor, E.R.; Habermann, B.; Broome, M.E. Scientific misconduct from the perspective of research coordinators: A national survey. *J. Med. Ethics* **2007**, *33*, 365–369.
22. Atlas, M.C. Retraction policies of high-impact biomedical journals. *J. Med. Libr. Assoc. (JMLA)* **2004**, *92*, 242–250.
23. Bosch, X.; Hernández, C.; Pericas, J.M.; Doti, P.; Marušić, A.; Manzoli, L. Misconduct policies in high-impact biomedical journals. *PLoS One* **2012**, *7*, e51928.
24. Resnik, D.; Peddada, S.; Brunson, W. Research misconduct policies of scientific journals. *Account. Res.* **2009**, *16*, 254–267.
25. Resnik, D.; Patrone, D.; Peddada, S. Research Misconduct Policies of Social Science Journals and Impact Factor. *Account. Res.* **2010**, *17*, 79–84.

26. Enders, W.; Hoover, G.A. Whose line is it? Plagiarism in economics. *J. Econ. Lit.* **2004**, *42*, 487–493.
27. Wager, E.; Fiack, S.; Graf, C.; Robinson, A.; Rowlands, I. Science journal editors' views on publication ethics: Results of an international survey. *J. Med. Ethics* **2009**, *35*, 348–353.
28. Williams, P.; Wager, E. Exploring Why and How Journal Editors Retract Articles: Findings from a Qualitative Study. *Sci. Eng. Ethics* **2013**, *19*, 1–11.
29. Elia, N.; Wager, E.; Tramèr, M.R.; Wray, K.B. Fate of Articles That Warranted Retraction Due to Ethical Concerns: A Descriptive Cross-Sectional Study. *PLoS One* **2014**, *9*, e85846.
30. Neale, A.V.; Dailey, R.K.; Abrams, J. Analysis of Citations to Biomedical Articles Affected by Scientific Misconduct. *Sci. Eng. Ethics* **2010**, *16*, 251–261.
31. Neale, A.V.; Northrup, J.; Dailey, R.; Marks, E.; Abrams, J. Correction and Use of Biomedical Literature Affected by Scientific Misconduct. *Sci. Eng. Ethics* **2007**, *13*, 5–24.
32. Resnik, D.B.; Dinse, G.E. Scientific retractions and corrections related to misconduct findings. *J. Med. Ethics* **2012**, *39*, 46–50.
33. Lappi-Seppälä, T. Penal policy in Scandinavia. In *Crime, Punishment, and Politics in Comparative Perspective*; Tonry, M., Ed.; The University of Chicago Press: Chicago, IL, USA, 2007; Volume 36, pp. 217–295.
34. Lappi-Seppälä, T. Explaining imprisonment in Europe. *Eur. J. Criminol.* **2011**, *8*, 303–328.
35. Heckman, J.J. Sample Selection Bias as a Specification Error. *Econometrica* **1979**, *47*, 153–161.
36. Macilwain, C. Scientific Misconduct: More Cops, More Robbers? *Cell* **2012**, *149*, 1417–1419.
37. Mongeon, P.; Larivière, V. *The Collective Consequences of Scientific Fraud: An Analysis of Biomedical Research*; Gorraiz, J., Ed.; In Proceedings of the International Conference on Scientometrics and Informetrics; Austrian Institut of Technology: Vienna, Austria, 2013; pp. 1897–1899.
38. Steen, R.G. Retractions in the scientific literature: Do authors deliberately commit research fraud? *J. Med. Ethics* **2011**, *37*, 113–117.
39. Bilbrey, E.; O'Dell, N.; Creamer, J. A novel rubric for rating the quality of retraction notices. *Publications* **2014**, *2*, 14–26.
40. Krohn, M.D.; Lizotte, A.J.; Phillips, M.D.; Thornberry, T.P.; Bell, K.A. Explaining Systematic Bias in Self-Reported Measures: Factors that Affect the Under- and Over-Reporting of Self-Reported Arrests. *Justice Q.* **2013**, *30*, 501–528.
41. Maxfield, M.G.; Luntz Weiler, B.; Spatz Widom, C. Comparing Self-Reports and Official Records of Arrest. *J. Quant. Criminol.* **2000**, *16*, 87–110.
42. Martinson, B.C.; Anderson, M.S.; de Vries, R. Scientists behaving badly. *Nature* **2005**, *435*, 737–738.
43. Böhmer, S.; Neufeld, J.; Hinze, S.; Klode, C.; Hornbostel, S. *Wissenschaftlerbefragung 2010. Forschungsbedingungen von Professorinnen und Professoren in Deutschland*, iFQ-Working Paper No.8; Institut für Forschungsinformation und Qualitätssicherung: Bonn, Germany, 2011.
44. List, J.; Bailey, C.; Euzent, P.; Martin, T. Academic economists behaving badly? A survey on three areas of unethical behavior. *Econ. Inq.* **2001**, *39*, 162–170.
45. Gardner, W.; Lidz, C.W.; Hartwig, K.C. Authors' reports about research integrity problems in clinical trials. *Contemp. Clin. Trials* **2005**, *26*, 244–251.

46. Geggie, D. A survey of newly appointed consultants' attitudes towards research fraud. *J. Med. Ethics* **2001**, *27*, 344–346.
47. Schwarz, N. Cognitive aspects of survey methodology. *Appl. Cogn. Psychol.* **2007**, *21*, 277–287.
48. Becker, H.S. *Outsiders: Studies in the Sociology of Deviance*; Free Press: New York, NY, USA, 1973.
49. Frow, E.K. Drawing a line: Setting guidelines for digital image processing in scientific journal articles. *Soc. Stud. Sci.* **2012**, *42*, 369–392.
50. Vaughan, D. The Dark Side of Organizations: Mistake, Misconduct, and Disaster. *Annu. Revie. Sociol.* **1999**, *25*, 271–305.
51. Knorr-Cetina, K. *Epistemic Cultures: How the Sciences Make Knowledge*; Harvard University Press: Cambridge, MA, USA, 1999.
52. Merton, R.K. *Social Theory and Social Structure*; Free Press: New York, NY, USA, 1968.
53. Collins, H.M. *Changing Order: Replication and Induction in Scientific Practice*; University of Chicago Press: Chicago, IL, USA, 1992.
54. Gieryn, T.F. Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists. *Am. Sociol. Rev.* **1983**, *48*, 781–795.
55. Gieryn, T.F.; Figert, A.E. Scientists protect their cognitive authority: The status degradation ceremony of Sir Cyril Burt. In *The Knowledge Society*; Böhme, G., Stehr, N., Eds.; Sociology of the Sciences; Springer: Dordrecht, The Netherlands, 1986; pp. 67–86.
56. Gieryn, T.F. *Cultural Boundaries of Science: Credibility on the Line*; University of Chicago Press: Chicago, IL, USA, 1999.
57. Flick, U. *Qualitative Sozialforschung: Eine Einführung*; Rororo Rowohlt's Enzyklopädie; Orig.-Ausg., vollst. überarb. und erw. Neuausg; Rowohlt-Taschenbuch-Verl.: Reinbek bei Hamburg, Germany, 2007; Volume 55694.
58. Denzin, N.K.; Lincoln, Y.S. *The SAGE Handbook of Qualitative Research*, 3rd ed.; Sage Publications: Thousand Oaks, CA, USA, 2005.
59. King, G.; Keohane, R.O.; Verba, S. *Designing Social Inquiry: Scientific Inference in Qualitative Research*; Princeton Paperbacks; Princeton University Press: Princeton, NJ, USA, 1994.
60. Pinch, T.J. Normal Explanations of The Paranormal: The Demarcation Problem and Fraud in Parapsychology. *Soc. Stud. Sci.* **1979**, *9*, 329–348.