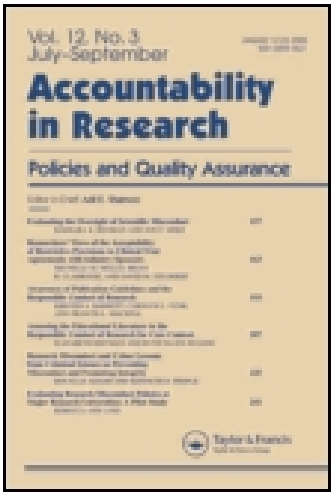

Accountability in Research: Policies and Quality Assurance
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gacr20

Commentary: Perverse Incentives or Rotten Apples?

Lex M. Bouter Ph.D.[ab]
[a] Department of Epidemiology & Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[b] Department of Philosophy, Faculty of Humanities, VU University, Amsterdam, The Netherlands
Published online: 30 Jan 2015.

To cite this article: Lex M. Bouter Ph.D. (2015) Commentary: Perverse Incentives or Rotten Apples?, Accountability in Research: Policies and Quality Assurance, 22:3, 148-161, DOI: 10.1080/08989621.2014.950253

To link to this article: http://dx.doi.org/10.1080/08989621.2014.950253

Accountability in Research, 22:148–161, 2015

ISSN: 0898-9621 print / 1545-5815 online
DOI: 10.1080/08989621.2014.950253

Commentary: Perverse Incentives or Rotten Apples?

Lex M. Bouter, Ph.D.[1,2]

[1]Department of Epidemiology & Biostatistics, VU University Medical Center, Amsterdam, The Netherlands
[2]Department of Philosophy, Faculty of Humanities, VU University, Amsterdam, The Netherlands

Around 2% of the investigators admit to have falsified or fabricated data at least once. Also, 34% report to have been guilty to one or more questionable research practices, such as doing many statistical analyses and to publish only what fits their theoretical framework. Prevention of questionable research practices is very important. Universities should ensure that the training is in order and the research culture is adequate, and they should critically look at perverse incentives, such as a too high publication pressure, but also by ensuring proper guidelines, and by having a fair and transparent procedure for suspected violations of scientific integrity.

Keywords: peer review, questionable research practices, research integrity, research misconduct, responsible conduct of research, selective reporting

INTRODUCTION

Scientists are both ordinary and extraordinary people. Like all of us, they steer their own course through life, keeping in mind what is best for their fellow man and often what is best for themselves. As a rule, they are highly motivated to advance their field and are committed to the cause of good education. But they are also exposed to temptation. After all, they are only human. It would make a wonderful theme for an exciting movie or a compelling book. The novel is

This article is based on my inaugural lecture (in Dutch) formally starting my chair on Methodology and Integrity, which was presented at VU University Amsterdam on May 2, 2014. The Dutch version of the lecture was distributed as a booklet among the audience and is submitted for publication to the Dutch and Flemish journal for management in higher education *Thema*. The original Dutch text and the English translation of the inaugural lecture were made available as PDF in the repository of VU University Amsterdam.

Address correspondence to Lex M. Bouter, Department of Epidemiology & Biostatistics, VU University Medical Center, P. O. Box 7057, 1007 MB, Amsterdam, Netherlands. E-mail: lm.bouter@vumc.nl

perhaps the best form for investigating the essence of what scientists do, and why they do it.

Here I will limit myself to a single example (Goodman, 2006). A postdoctoral researcher has produced spectacular results. They lead to a rapid publication in *Nature* and a substantial new research grant. This generates all kinds of media attention, and the entire laboratory switches its focus to the follow-up research. A colleague is not able to reproduce the results and finds indications that data has been selectively omitted. She gradually finds herself taking on the role of the "whistleblower." The USA Office for Research Integrity (ORI) concludes that there is evidence of scientific misconduct, but on appeal this ruling is annulled on procedural grounds. Meanwhile, a senator has seized on the case as part of his crusade against science. Even on the last page, the reader is still not able to get to the bottom of what really happened.

It is an impressive book. The plotline shows how ambition, external pressure, negligence, and lack of supervision can lead to misconduct among scientists. It also shows that there are many shades of grey along the spectrum that runs from complete integrity to research misconduct. Under a magnifying glass, irregularities in daily practice soon become visible, but interpreting them is tricky. The author's exceptional achievement lies in credibly conveying the motivations and emotions of everyone involved: the postdoc, the whistleblower, and the supervisors. There are only losers, and the damage to an individual's reputation depends only to a limited extent on the facts. It is a book I can wholeheartedly recommend.

SHADES OF GREY

My argument primarily concerns those shades of grey. I will focus on how scientists conduct themselves, and discuss how to promote desirable behavior and combat undesirable behavior. I will call for efforts to prevent violations of academic integrity to be stepped up through education and quality control in the workplace. I will advocate targeted scientific research, and I will present my views on the role of universities and of the agencies funding research. I will conclude that we need to focus our full attention not only on the bad apples, but especially on the perverse incentives that exist in today's academic world.

The preoccupation with publishing vast amounts and achieving frequent citations may well be such a perverse incentive. From 1982, I have seen the gains in quality, relevance, and efficiency that this incentive has brought about. But it is true: you can overdo things. The argument that pressure to publish is now working as a perverse incentive in many disciplines seems to me to be defensible. One significant finding in this regard is that over half of the medical professors in the Netherlands experience the pressure to publish as excessive and a quarter of them meet the clinical definition of burnout (Tijdink et al.,

2013). That does not alter the fact that the vast majority of scientists behave with integrity and not succumb to the pressure to publish.

Diederik Stapel was clearly a bad apple (Stapel Investigation, 2012). Personally, he believes that perverse incentives played a major role in his going off the rails. In his autobiography (Stapel, 2012) he gives an alarming account of how easily he was able to keep on fooling himself and others. It is astounding how primitive his deception was and how long it took for him to be unmasked. Stapel was clearly an extreme case. It goes without saying that such cases must be detected and addressed. But more importantly, it is vital that we work to strengthen the collective resistance to perverse incentives among all researchers. It is all about the everyday dilemmas, about the human tendency to cut corners where possible, and about the shades of grey in actual behavior on a sliding scale.

There is little debate at either end of the spectrum of academic integrity. Research that is carried out in complete accordance with the rules is the norm. Falsifying and fabricating data and committing plagiarism constitute very serious wrongdoings. They are nothing short of research misconduct. But between these two points, there is an extensive area which covers all kinds of questionable research practices. These often involve the violation of basic methodological principles. For instance, carrying out a whole range of statistical analyses and only publishing what suits your needs; or focusing on other research questions than those the study was designed to address, without due disclosure; or summarizing existing knowledge on the basis of preconceptions; or refraining from publication if a research project has failed in the eyes of the researcher or the sponsor.

My point is that things often go wrong in that grey area due to a failure to apply state-of-the-art methodology. This can happen because people do not know how things should be done; or they do know but believe that there is nothing wrong with cutting corners here and there; or they realize that there is serious wrongdoing involved, but they proceed anyway in order to draw the preferred conclusions because they are already heavily invested in a particular theory; or because they believe that this will increase the chance of publication or of obtaining a follow-up grant; and so on and so forth. At the light grey end of the spectrum of questionable research practices, this often involves methodological principles that are still open to discussion and that are sometimes the object of considerable differences of opinion between disciplines. At the dark grey end of the spectrum, there is no room for discussion. There, it is clear that what people are doing is simply wrong, and the individuals involved usually know this all too well!

It is difficult to accurately quantify people's behavior on this spectrum (Martinson et al., 2005). Some years ago, Fanelli (2009) published a meta-analysis of the best available estimates. When asked, around 2% of the researchers admitted to having falsified or fabricated data at least once, and

34% admitted to having been guilty of questionable research practices at least once. These figures rise to 14% and 72%, respectively, when the same questions are asked about the conduct of colleagues. To my mind, these percentages are not only high, but also extremely worrying.

ROOM FOR IMPROVEMENT

My favorite definition of science is the following one: "Science is the struggle for truth against methodological, psychological and sociological obstacles" (Fanelli and Ioannidis, 2013). The business of science is tough enough as it is, even without violations of integrity. This is something not everyone is aware of, so allow me to elaborate for a moment.

John Ioannidis (2005) wrote a controversial article summarizing what many methodologists already knew: most published findings of empirical research are incorrect. The probability that a statistically significant positive finding is consistent with the truth depends on three factors. The first factor is the *power* of the study, that is to say the probability that the study will actually detect an existing positive association. The second factor is the prevalence of true positive associations in the relevant field of research. The third factor is the probability of a positive finding as a result of bias, for example, due to errors in the research design or to selective reporting. Ioannidis (2005) shows that the probability of a statistically significant positive finding being consistent with the truth can vary from 85% in a large well-designed randomized clinical trial to 0.1% in exploratory analyses of large databases. His analysis demonstrates that a positive finding is less likely to be correct as the number of research units and the observed effect grow smaller. This is also true when the number of statistical tests is larger; when there is greater subjectivity in choosing the research design, definitions, outcomes, and analytical methods, and when substantial interests, financial and otherwise, are involved.

Iain Chalmers and Paul Glasziou (2009) have reached similar conclusions on different grounds. They distinguish four problems within the grey area of questionable research practices. First, they show that researchers often choose to investigate questions that are of little relevance. Second, they argue convincingly that research design often leaves a lot to be desired. Over half of the studies carried out are not founded on a systematic review of what is already known about the topic. In addition, the measures taken to prevent avoidable bias are often insufficient. The third problem is that the results published represent less than half of the studies actually carried out. The fourth problem concerns the shortcomings in the quality of the publications that do appear in print. Over one third of the interventions are described in insufficient detail and over half of the outcomes measured are not reported. Chalmers and Glasziou (2009) conclude that all this can generate "avoidable waste" of up to 85%.

Both of these analyses are obviously debatable. My intention here is simply to illustrate the forces at work within that grey area and to substantiate my position that there is plenty of room for improvement, in particular through improved application of key methodological principles. Incidentally, it is worth pointing out that in the area of questionable research practices, it is often not possible to distinguish between research that has been poorly designed and carried out on the one hand, and dubious integrity on the part of the researchers on the other hand. This is a far simpler matter when it comes to research misconduct. In such cases, integrity is definitely found wanting, although without a confession from the suspect, it is often difficult to prove that the fraud was deliberate.

Peer review is the dominant and-–according to many-–the best kind of quality assessment available in science. It takes place in advance, when project proposals are assessed and subject to a medical-ethical review, and it takes place afterwards, when manuscripts, research groups, and researchers are evaluated. However, the objectivity of peer review is not beyond reproach.

It is difficult to give equal weight to all of the relevant aspects, panels are often one-sided in their composition, and the panellists' own views and interests are often too dominant. This means that genuinely innovative and excellent proposals are sometimes given too few opportunities (Nicholson and Ioannidis, 2012). Diversity is an important factor, and not only for peer review. In research teams and nomination committees, a diverse composition is also the best way to avoid tunnel vision and collective blind spots. Outsiders can often shed light on elements that are taken for granted within a discipline. This offers a valuable opportunity to improve on practices that are less than ideal.

SELECTIVE REPORTING

Peer review is not well equipped to detect questionable research practices and scientific misconduct. The findings of John Bohannon (2013) in this regard are downright alarming. He sent a fabricated manuscript containing unacceptable errors to over 300 journals. Over half of the journals accepted the manuscript for publication, in some cases even though the referees had pointed out one or more of the key shortcomings of the manuscript.

There is every reason to take a more critical look at the performance of peer review in the publication culture. What can reasonably be expected from the reviewers of a manuscript? Should they check references and repeat analyses? Should their reports be made public? And who reviews the quality of the reviewers? The roles of editors and publishers also deserve further consideration. Should they actively seek out plagiarism, check whether all relevant conflicts of interest have been reported, and verify whether all authors meet the applicable criteria for authorship?

The authors of systematic reviews are probably the most critical readers of scientific articles. Systematic reviews are a great help in showing what we already know and what we have yet to find out. They are also a good way to identify the methodological shortcomings of existing research. This means that systematic reviews provide a solid foundation for new research, both in terms of the research question and the research design. However, the contribution that systematic reviews make to the detection of possible violations of academic integrity is modest at best. Showing where a specific study deviates in terms of method and results may expose sloppy science or worse.

The Achilles' heel of systematic reviews lies in publication and reporting bias. After all, if not all research results are published, there is a distinct possibility of presenting a distorted picture (Dwan et al., 2008). The only remedy for this is to register all studies and publish all research protocols (Chan et al., 2014). At present, this still happens far too infrequently (Wager and Williams, 2013; Chalmers et al., 2013). Even when we look at registered randomized clinical trials involving over 500 participants, 30% have still not been published five years after the completion of data collection (Jones et al., 2013). There is a world to be won in this respect. The recycling of published research results without acknowledgment also poses a threat to the validity of meta-analyses. It is often far from easy to identify such recycling, which means that the same participants may appear twice or even three times in the meta-analysis.

A disturbing article about reporting bias in the management sciences was recently published (O'Boyle et al., 2014). Out of nearly 2,000 hypotheses researched in 142 dissertations, only one-third were presented in the scientific articles that described the outcomes of these dissertations. Relatively often, there were changes in the statistical significance between dissertation and article: a change from nonsignificant to significant was over four and a half times more common than the other way around. Only 40% of the 1,333 hypotheses which appeared in the dissertations but not in the articles were statistically significant, compared to 70% for the 333 new hypotheses which only appeared in the articles. The authors show that by manipulating hypotheses, variables, and data, nonsignificant findings were transformed into significant findings on a major scale. There is little reason to believe that such practices are restricted to the management sciences.

Making the data files on which a publication is based available to everyone would be a major step forward. It is a requirement that journals are more inclined to make nowadays. However, I am not in favor of simply granting public access to data files across the board. Researchers should first be given the opportunity to publish on their work themselves. Without adequate knowledge of the structure of a data file, the chance of errors is considerable. There is also a risk of tendentious and malicious use by third parties. For, as many researchers know, "If you torture your data enough, nature will always confess" (Coarse, 1982).

PREVENTION

This concludes my review of the problems we face. Now it is time to say something about the possible solutions. Preventing integrity issues through education and training is one important step (Steneck, 2013; Godecharle et al., 2013; Kornfeld, 2013; DuBois and Duecker, 2009). Desirable behavior, questionable research practices, and scientific misconduct should all be explicitly and extensively addressed. They are all part and parcel of the broader context of academic development, by which I mean it exists in close conjunction with the philosophy of science, scientific ethics, and the teaching of methodology. For as I mentioned earlier, methodology is at the heart of many integrity problems. In this regard, knowledge transfer and the teaching of skills are not the most important factor. What is crucial is a focus on the day-to-day dilemmas that surround the practice of research (Grinnell, 2013): recognizing these dilemmas in your own work and that of others, learning to reflect on them, and learning that it is normal to discuss questionable behavior. Education should lead to resilience; it should help us identify perverse incentives and resist them; and it should contribute to a culture of responsible conduct in research (Kalichman, 2014). That is what it is all about.

This cannot be achieved with a few lectures; it calls for blended learning that combines online education, exercises, and workgroup discussions. Peer-to-peer feedback and moral case deliberation (Widdershoven, 2014) can help to make day-to-day dilemmas a topic of discussion, especially among Ph.D. students, postdocs, and their supervisors. There is excellent teaching material available—that is not the problem—but it is only being used to a very limited extent and generally speaking there is no coherent policy. That is also true of my own university. I believe that this has to change, and I am happy to do all I can to help make that change in the coming years.

For permanent academic staff, training in the field of scientific integrity is even more important than for Ph.D. students and postdocs. After all, they are the role models who show how people deal with day-to-day dilemmas in practice. That is something no course can achieve (Anderson et al., 2007). Therefore, I believe that training for supervisors and co-supervisors should no longer be optional and should lead to a license-to-supervise. Regular seminars on current topics, new regulations, and relevant research keep the subject alive and provide the necessary in-service training. For example, the recent advisory letter from the Royal Netherlands Academy of Arts and Sciences (KNAW, 2014) on the correct reuse of previously published material would make a good subject for such a seminar.

But there is an area that may be even more important than education, and that is quality control in the workplace. For it is in the workplace that things stay on the straight and narrow or take a wrong turn. It is essential to create a culture in which dilemmas are discussed and where people help each other

to avoid pitfalls. It is all about combining the intrinsic motivation to do honest research and the extrinsic factors designed to promote such research. The application of clear and explicit guidelines provides the foundation. Not that these should be followed blindly. The underlying principle is always "comply or explain," and guidelines should of course be firmly anchored in the relevant international, national, and institutional codes.

A good example has already been set by EMGO+ (2014). This research institute has over 10 years experience of working with these guidelines and the internal audits based on them. All new employees are given an introduction in how to use the guidelines. The audits focus on specific research projects or themes that run through various projects. For example, how data is stored, and how published analyses can be reproduced. Experience has shown that young researchers, as well as funding organizations and review committees, greatly appreciate this approach.

Of course, the culture in the workplace depends on so much more than the availability of guidelines and sound quality control. A thorough understanding of the views of the researchers is crucial, as is discovering what they perceive to be perverse incentives. It is therefore important to bring these aspects into focus (Martinson et al., 2013; Crain et al., 2013), so that policies and educational content can be modified accordingly. Moreover, discussing the results of such a survey is in itself an important intervention that puts the spotlight on scientific integrity.

RESEARCH

Surprisingly little scientific research has been carried out into violations of academic integrity (Huberts, 2014). We do not have a clear picture of how often the various types of questionable research practices and the various forms of scientific misconduct occur. Nor do we know if there are major differences between disciplines. In addition, we know almost nothing of the main causes of these problems. Should we look for them in perverse incentives and the culture of institutions and research groups, or rather in the character flaws of the individual researchers? It is a mystery.

However, there is no shortage of theories. Some authors believe (Stroebe, 2012) that scientists who overstep the mark make a rational decision, weighing up the slim chance of being caught and the limited penalties on the one hand, and the considerable rewards that their inappropriate conduct can bring in terms of prestige, funding, and career advancement on the other hand (Adams and Pimple, 2005). An interesting alternative approach comes from experimental psychology (Ariely, 2012) and is based around the core idea that everyone is inclined to lie and cheat. We constantly fool others and ourselves. But the irrational thing is that we tend to do so in moderation, even if the risk of being found out is negligible. The behavior of role models and what we see happening

around us are what tips the balance. Collectively stepping across a line soon creates a new standard. Creative and innovative thinkers are thought to have a greater ability to justify their own questionable research practices (Gini and Ariely, 2012). If that is true, then outstanding talents are more at risk!

I believe that there is an urgent need for sound scientific research to better understand how questionable research practices and research misconduct come about, and to substantiate or indeed disprove the usefulness and necessity of certain methodological principles. But above all to identify the most effective educational and organizational interventions for preventing this inappropriate behavior.

ACTION IS NEEDED

That is not a justification for sitting back and doing nothing. Action is needed, particularly on the part of knowledge institutions and the organizations that fund research. But scientific journals, international scientific associations, and accreditation bodies and national academies of arts and sciences also have a role to play. Universities can and should go the extra mile to safeguard scientific integrity. They can start by making it clear that this matter is important to them. This involves working to ensure a healthy balance between broad support and effective decision-making. We would do well to take the approach used at Aarhus University (2013) in Denmark as our example. Broadly speaking, that is the approach I will now go on to describe.

First, it is important to unambiguously endorse the normative framework of the relevant international and national codes of conduct (VSNU, 2012; ESF and ALLEA, 2011; 2nd World Conference on Research Integrity, 2010; 3rd World Conference on Research Integrity, 2013). Leaders at all levels have to be convinced that acting in accordance with this normative framework is of great importance. According to international consensus, those standards ought to be developed in greater detail for the major disciplines within each institution. A good example of this is the recent Research Code published jointly (VUmc and AMC, 2013) by Amsterdam's two main teaching hospitals, the Academic Medical Center (AMC), and VU University Medical Center (VUmc). A similar research code could be drawn up for the natural sciences, the social sciences, and the humanities. In the workplace, these codes are then converted into concrete guidelines, as in the earlier example from EMGO+. These discipline-specific research codes and their translation into practical guidelines should preferably be dynamic in nature, so that progressive insights and new developments can be rapidly incorporated. Moreover, the process of drawing up and amending these codes and guidelines is at least as important as the result.

In addition, universities have an important duty to implement the preventive measures previously mentioned. All I will add on this subject for

now is that this is far from self-evident! The experiences of the best practice institutions I have visited have taught me that the proposed approach sometimes sparks resistance and that progress can only be achieved with adequate resources and administrative tenacity. In this context, it is also good to gain an idea of what people in the workplace are thinking. An anonymous survey among academic staff can provide the necessary insight. The method is available and can be implemented at short notice (NCPRE, 2014).

Of course, universities also have a role to play in the mitigation of potentially perverse incentives. Among other things, this can be done by ensuring sufficient diversity in criteria for promotion, career paths, and the composition of selection committees. This will also help avert the danger of tunnel vision and collective blind spots. Simple rules can help reduce risks to integrity. These might include appointing external members and an independent chairperson to a manuscript committee when doctorates are awarded. Binding rules for archiving data, lab journals, and scripts for data analysis are urgently needed. Obviously, the procedures in place to deal with suspected violations of academic integrity should be fair and clear. This includes proper rules on confidentiality, hearing both sides of the argument and clear criteria governing further investigation and any penalties that may follow.

The organizations that fund research also have an important role to play in combating perverse incentives and promoting scientific integrity. Sufficient diversity within programmes, evaluation criteria, and committees can prevent strategic behavior on the part of researchers that might lead to questionable research practices. A monoculture focused on citation scores, short-term economic gain, and government-defined growth sectors may also lead to an underutilization of research funds. Involving young talented researchers and end users when selecting research proposals also increases the opportunities for relevant, excellent, and innovative projects.

I believe that the organizations which provide research funding should be able to make demands of universities (Titus and Bosch, 2010). For example, with regard to the attention devoted to education about academic integrity and quality control in the workplace; or how alleged violations of scientific integrity are dealt with; and how project proposals are motivated and-–once approved–-how they are carried out and reported on. In a critical reflection on the need for structural reforms (Alberts et al., 2014), the voice of the funding organization should also be heard.

As early as in 1987, it was argued (Shamoo and Annau, 1987) that we need data audits. Codes of conduct, education and training, and whistleblowing have a role to play. But in line with what is considered good practice concerning financial arrangements, accountability in research would greatly benefit if both "for cause" and "random" data audits would be performed (Shamoo, 2013). The purpose of such an audit is to check the trajectory from protocol to publication and specifically the degree of correspondence of the published results with the

original data set. Although many scientists would probably not like the idea for various reasons, data audits would quite likely be helpful and cost efficient in the identification and prevention of research misconduct.

Research in the field of academic integrity deserves generous funding. That will certainly help reduce the risks of questionable research practices and scientific misconduct. This is something that society is now demanding, and rightly so. After all, it is reasonable to be held accountable for what we do with the public funds entrusted to us. We might as well get used to it: the ivory tower has become a glass house! Despite the painful incidents of the past few years, public confidence in science is still substantial (Rathenau Instituut, 2013a,b; KNAW, 2013). But to keep it that way, it seems to me essential to improve the way we operate and to communicate clearly about such matters. Greater transparency will also enable us to debunk a number of misconceptions about science and scientists. We should not make things out to be better than they are.

To summarize the current situation: there are dark clouds overhead, but here and there the sun is peeking through. The scientific method is a powerful tool and a vital source of hope for the future. Let us face up to the limitations and the darker side of our scientific endeavors but without becoming bogged down in gloom and nihilism, and let us focus on increasing the probability of appropriate conduct and reducing the risk that scientists will stray from the straight and narrow. At the same time, we should remember that bad apples do exist and represent a problem that needs to be addressed effectively. Nevertheless, our efforts at prevention should focus on the culture in the workplace and on combating perverse incentives.

REFERENCES

2nd World Conference on Research Integrity. (2010). Singapore statement on research integrity. Available at http://www.singaporestatement.org/downloads/singpore%20statement_A4size.pdf. Last accessed July 13th, 2014.

3rd World Conference on Research Integrity. (2013). Montreal statement on research integrity in cross-boundary research collaborations. Available at http://www.cehd.umn.edu/olpd/MontrealStatement.pdf. Last accessed July 13th, 2014.

Aarhus University. (2013). Discussion paper for the establishment of "joint guidelines for responsible conduct of research at Aarhus University." Available at http://www.au.dk/fileadmin/www.au.dk/forskning/Discussion_paper.pdf. Last accessed July 13th, 2014.

Adams, D. and Pimple, K. D. (2005). Research misconduct and crime: Lessons from criminal science on preventing misconduct and promoting integrity. *Accountability in Research*, 12:225–240.

Alberts, B., Kirschner, M. W., Tilghman, S, and Varmus, H. (2014). Rescuing U.S. biomedical research from its systematic flaws. *PNAS Early Edition*. Available at http://www.pnas.org/cgi/doi/10.1073/pnas.1404402111. Last accessed July 13th, 2014.

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., and Martinson, B. C. (2007). What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a national survey of NIH-funded scientists. *Academic Medicine*, 82:853–860.

Ariely, D. (2012). *The (Honest) Truth about Dishonesty: How We Lie–Especially to Ourselves*. New York: HarperCollins.

Bohannon, J. (2013). Who's afraid of peer review? *Science*, 342:60–65.

Chalmers, I. and Glasziou, P. (2009). Avoiding waste in the production and reporting of research evidence. *Lancet*, 374:86–89.

Chalmers, I., Glasziou, P., and Godlee, F. (2013). All trials must be registered and the results published. *British Medical Journal*, 346:f105.

Chan, A. W, Song, F., Vickers, A,. Jefferson, T., Dickersin, K., Gøtzsche, P. C., Krumholz, H. M., Ghersi, D., and Van der Worp, B. H. (2014). Increasing value and reducing waste: Addressing inaccessible research. *Lancet*, 383:257–266.

Coarse, R. H. (1982). *How should economists choose?* Washington, D.C.: American Enterprise Institute for Public Policy Research.

Crain, A. L., Martinson, B. C., and Thrush, C. R. (2013). Relationships between the Survey of Organizational Research Climate (SORC) and self-reported research practices. *Science and Engineering Ethics Sci*., 19:835–850.

DuBois, J. M., and Duecker, J. M. (2009). Teaching and assessing the responsible conduct of research: A Delphi consensus panel report. *The Journal of Research Administration*, XL:49–71.

Dwan, K., Altman, D., Arnaiz, J. A., Bloom, J., Chan, A. W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Helm, E., Gamble, C., Ghersi, D., Ioannidis, J. P. A., Simes, J., and Williamson, P. R. (2008). Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS ONE*, 3:e3081.

EMGO+. (2014). *Quality Handbook*. Available at http://www.emgo.nl/kc. Last accessed July 13th, 2014.

ESF and ALLEA. (2011). The European code of conduct for research integrity. Available at http://www.nsf.gov/od/iia/ise/Code_Conduct_ResearchIntegrity.pdf. Last accessed July 13th, 2014.

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*, 4:e5738.

Fanelli, D. and Ioannidis, J. P. A. (2013). U.S. studies may overestimate effect sizes in softer research. *Proceedings of the National Academy of Sciences*, 110: 15031–15036.

Gini, F. and Ariely, D. (2012). The dark side of creativity: original thinkers can be more dishonest. *Journal of Personality and Social Psychology*, 102:445–459.

Godecharle, S., Nemery, B., and Dierickx, K. (2013). Integrity training: Conflicting practices. *Science*, 340:1403.

Goodman, A. (2006). *Intuition*. New York: Random House.

Grinnell, F. (2013). Research integrity and everyday practice of science. *Science and Engineering Ethics*, 19:685–701.

Huberts, L. (2014). *The Integrity of Governance: What It Is, What we Know, What Is Done, and Where to Go*. Basingstoke: Palgrave Macmillan.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 8:e124.

Jones, C. W., Handler, L., Crowell, K. E., Keil, L. G., Weaver, M. A., and Platts-Mills, T. F. (2013). Non-publication of large randomized clinical trials: Cross sectional analysis. *British Medical Journal*, 347:f6104.

Kalichman, M. (2014). Rescuing responsible conduct of research (RCR) education. *Accountability in Research*, 21:68–83.

KNAW. (2013). Vertrouwen in wetenschap. Available at https://www.knaw.nl/en/news/publications/vertrouwen-in-wetenschap. Last accessed July 13th, 2014.

KNAW. (2014). Correct citation practice. Available at https://www.knaw.nl/en/news/publications/correct-citation-practice?set_language=en. Last accessed July 13th, 2014.

Kornfeld, D. S. (2013). Integrity training: misconduct's source. *Science*, 340:1403–1404.

Martinson, B. C., Anderson, M. S., and De Vries, R. (2005). Scientists behaving badly. *Nature*, 435:737–738.

Martinson, B. C., Thrush, C. R., and Crain, A. L. (2013). Development and validation of the Survey of Organizational Research Climate (SORC). *Science and Engineering Ethics*, 19:813–834.

NCPRE. (2014). Survey of organizational research climate. Available at https://nationalethicscenter.org/sorc. Last accessed July 13th, 2014.

Nicholson, L. M. and Ioannidis, J. P. A. (2012). Conform and be funded. *Nature*, 492:34–36.

O'Boyle, E. H., Banks, G. C., and Gonzalez-Mulé, E. (2014). The Chrystalis effect: How ugly initial results metamorphosize into beautiful articles. *Journal of Management*. Published online March 19th, 2014. Available at http://jom.sagepub.com/content/early/2014/03/18/0149206314527133. Last accessed July 13th, 2014.

Rathenau Instituut. (2013a). Hoeveel vertrouwen hebben Nederlanders in wetenschap? Available at http://www.rathenau.nl/publicaties/publicatie/hoeveel-vertrouwen-hebben-nederlanders-in-wetenschap.html. Last accessed July 13th, 2014.

Rathenau Instituut. (2013b). Geachte wetenschap. Available at http://www.rathenau.nl/publicaties/publicatie/geachte-wetenschap.html. Last accessed July 13th, 2014.

Shamoo, A. E. (2013). Data audit as a way to prevent/contain misconduct. *Accountability in Research*, 20:369–397.

Shamoo, A. E. and Annau, Z. (1987). Ensuring scientific integrity. *Nature*, 327:550.

Stapel, D. (2012). *Ontsporing*. Amsterdam: Prometheus.

Stapel Investigation. (2012). Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. Available at https://www.commissielevelt.nl. Last accessed July 13th, 2014.

Steneck, N. (2013). Global research integrity training. *Science*, 340:552–553.

Stroebe, W., Postmes, T., and Spears, R. (2012). Scientific misconduct and the myth of self-correction in science. *Perspectives on Psychological Science*, 7:670–688.

Tijdink, J. K., Vergouwen, A. C. M., and Smulders, Y. M. (2013). Publication pressure and burn out among Dutch medical professors: a nationwide survey. *PLoS ONE*, 8:e73381.

Titus, S. and Bosch, X. (2010). Tie funding to research integrity. *Nature*, 466:436–467.

VSNU. (2012). The Netherlands code of conduct for scientific practice. Available at http://www.vsnu.nl/files/documenten/Domeinen/Onderzoek/The_Netherlands_Code_of_Conduct_for_Scientific_Practice_2012.pdf. Last accessed July 13th, 2014.

VUmc and AMC. (2013). Research code. Available at http://www.vumc.com/branch/research-code-VUm-AMC/Contents. Last accessed July 13th, 2014.

Wager, E. and Williams, P. (2013). "Hardly worth the effort"? Medical journals' policies and their editors' publishers' views on trial registration and publication bias: Quantitative and qualitative study. *British Medical Journal*, 247:f5248.

Widdershoven, G. (2014). Moreel beraad als interventie: ethische dilemma's in de hedendaagse geneeskunde. In: Koster, E., ed. *Wat is wetenschap? Een filosofische inleiding voor levenswetenschappers en medici*. Amsterdam: VU University Press, pp. 345–359.