

Financial Conflicts of Interest and Criteria for Research Credibility

Kevin C. Elliott

Received: 1 October 2013 / Accepted: 1 October 2013 / Published online: 17 October 2013
© Springer Science+Business Media Dordrecht 2013

Abstract The potential for financial conflicts of interest (COIs) to damage the credibility of scientific research has become a significant social concern, especially in the wake of high-profile incidents involving the pharmaceutical, tobacco, fossil-fuel, and chemical industries. Scientists and policy makers have debated whether the presence of financial COIs should count as a reason for treating research with suspicion or whether research should instead be evaluated solely based on its scientific quality. This paper examines a recent proposal to develop criteria for evaluating the credibility of research without considering its source of funding. It concludes that proposals of this sort are likely to be either ineffective or impractical in many cases. Nevertheless, this does not imply that all research funded by those with an interest in the outcome must be placed under a cloud of suspicion; there are conditions under which research is at much more serious risk of being corrupted than in other cases, and attention to these conditions can guide productive responses to financial COIs.

1 Introduction

A wide range of scholars and concerned citizens have recently raised worries about the influences of powerful interest groups on scientific research. Books like *Merchants of Doubt* (Oreskes and Conway 2010), *Doubt is Their Product* (Michaels 2008), *Bending Science* (McGarity and Wagner 2008), *The Truth about the Drug*

K. C. Elliott (✉)
Department of Philosophy, University of South Carolina, Columbia, SC 29208, USA
e-mail: ke@sc.edu

K. C. Elliott
Lyman Briggs College and Department of Fisheries and Wildlife, Michigan State University,
East Lansing, MI 48825, USA
e-mail: kce@msu.edu

Companies (Angell 2005), *Deceit and Denial* (Markowitz and Rosner 2002), *Is a Little Pollution Good for You?* (Elliott 2011), *The Price of Truth* (Resnik 2006b), and *Global Spin* (Beder 2000) have drawn attention to a persistent pattern of industry efforts to produce misleading research and to confuse the public about hazards associated with smoking, acid rain, the ozone hole, climate change, second-hand smoke, industrial chemicals, and pharmaceuticals. Moreover, these strategies are not limited to industry; when other entities, such as government agencies, have had access to adequate resources and incentives, they have employed similar strategies (Michaels 2008; Resnik 2006b). Insofar as these strategies tend to paralyze and misdirect public policy making in response to major issues, they represent a significant social problem. Because of their expertise in evaluating scientific methodologies and reflecting on the relationships between science and society, philosophers of science are well-placed to help develop effective policy responses to this problem.

Unfortunately, while a number of potential policy responses to financial conflicts of interest in research have been proposed, most face significant problems. Universities, under pressure from government funding agencies, typically have conflict-of-interest policies in place (Resnik 2006b). Nevertheless, the effectiveness of these policies is questionable (Curzer and Santillanes 2012; Elliott 2008), and they do not address the large amounts of research that take place outside the university setting. Many experts have responded to problems with financial conflicts of interest by calling for governments to provide more funding for policy-relevant research, with the hope that this research would be more reliable (Angell 2005; APHA 2003; Elliott 2011; Shrader-Frechette 2007). For the time being, however, industry still funds the majority of scientific research and development in countries like the United States (Elliott 2011), and it funds the vast majority of safety studies for new pharmaceuticals and industrial chemicals (Conrad and Becker 2011). Others have suggested strategies that individual researchers can take to help them maintain their integrity in the face of conflicts of interest (Curzer and Santillanes 2012). But while these strategies might be helpful in some cases, they are of limited value when scientists are not fully aware of their biases or when they are not motivated to correct them.

In the face of these difficulties, various commentators have come to widely differing conclusions about the extent to which research should be treated with suspicion when financial conflicts of interest are present (see Resnik and Elliott 2013). A recent interchange in the journal *Environmental Health Perspectives* highlights these debates. On one hand, Conrad and Becker (2011) argued that research should be evaluated independently of its funding source. On the other hand, Patrice Sutton et al. (2011) and Tweedale (2011) responded that there are good empirical reasons to consider funding sources as a potential source of bias.

The present paper addresses these issues by evaluating the relative merits of employing criteria that ignore funding sources versus employing criteria that include it as part of an analysis of study credibility. Section 2 provides an overview of current trends regarding financial COIs, including empirical evidence concerning their effects on research. The next section examines Conrad and Becker's (2011) attempt to formulate criteria without considering funding sources and concludes that this approach is likely to be either ineffective or impractical in many cases.

Section 4 then considers the possibility of including sources of funding as a criterion for assessing research credibility. The paper concludes that one can refine this criterion by identifying some funding conditions that are more worrisome than others, and this refinement suggests fruitful strategies for eliminating the conditions that are most likely to result in questionable research.

2 Background on Financial Conflicts of Interest

A common definition of a conflict of interest is “a set of conditions in which professional judgment concerning a primary interest (such as patients’ welfare or the validity of research) tends to be unduly influenced by a secondary interest (such as financial gain)” (Thompson 1993; see also Davis 1982; Resnik 2006b). While financial conflicts of interest arise in a variety of settings, they have recently received increasing attention in the context of scientific research. This is partly because of high-profile scandals in which powerful industries were found to have suppressed scientific information, produced questionable research, and used PR organizations and think tanks to disseminate misleading information about their products (McGarity and Wagner 2008). The tobacco industry has received a great deal of attention as a pioneer that developed many of these strategies in their present form during the latter half of the twentieth century (Proctor 2012), but similar approaches were taken at the same time by other industries, including those using lead, asbestos, and vinyl chloride (Markowitz and Rosner 2002; McGarity and Wagner 2008). More recently, the fossil fuel industry has come under scrutiny for its efforts to disseminate inaccurate claims about climate change (Oreskes and Conway 2010), and the pharmaceutical industry has been assailed from many quarters because of evidence that it has suppressed unwelcome findings and inappropriately manipulated the interpretation of studies (Angell 2005; Schafer 2004).

In general, the worry about financial conflicts of interest in science is that the secondary interest of financial gain can cause those involved in research to sacrifice the primary interest of upholding important scientific norms (e.g., reporting data accurately and completely, interpreting data according to community conventions, and acknowledging major value judgments or interpretive assumptions) (McKaughan and Elliott 2013; Wilholt 2009). In their book *Bending Science* (2008), Tom McGarity and Wendy Wagner summarize many of the different ways that science can suffer as a result. They conceptualize scientific practice as a pipeline moving through several steps: conducting research, interpreting data, engaging in peer review at scientific journals, scrutinizing research after publication, and evaluating it in expert panels. McGarity and Wagner emphasize that interest groups can inappropriately influence science at each stage.¹ They can start by influencing the framing of questions, the way studies are designed, and the manner in which data

¹ As the later sections of this paper emphasize, industry is not the only interest group that inappropriately influences science. While industry has received a great deal of attention (in part because of the prodigious financial resources at its disposal), government agencies, labor organizations, and citizens’ groups sometimes employ very similar strategies.

are interpreted. If unwanted findings still emerge, they can conceal the findings through mechanisms such as gag contracts and trade secrets. If an unwelcome finding is published, they can produce new research or reanalyze the data in an effort to discredit it. Finally, special panels and conferences can be created to emphasize preferred research, and PR firms can be used to disseminate this information to the public at large.

Instances of these strategies are now commonplace. For example, dioxins are particularly potent carcinogens, and yet Monsanto and BASF managed to confuse regulators for years based on major studies that they published in journals like *Scientific American*, *Science*, and the *Journal of the American Medical Association*. These studies purported to show few ill effects from the compounds, and yet they were later found to contain crucial data falsifications (Beder 2000, 141–143). In the pharmaceutical industry, ghostwriting—in which prominent academics are paid to put their names on papers that are largely designed, performed, and written by industry-affiliated groups—appears to have become widespread. Although it is difficult to collect evidence about such activities, one study found that more than 50 % of the articles published on the antidepressant Zoloft between 1998 and 2000 were ghostwritten. Unsurprisingly, the ghostwritten articles contained more favorable evaluations of Zoloft than the others, but they were also cited more often and published in more prestigious journals (Healy and Catell 2003). One might wonder how these studies could be published in such high-quality venues, but Richard Smith, the former editor and chief executive of the *British Medical Journal* group, argues that this should not be surprising. Pharmaceutical companies have the money and expertise to perform studies that look excellent but that generate desired results by using data selectively or strategically interpreting the findings (Smith 2005).

In addition to all these specific examples of interest-group strategies and activities, there is also a growing body of empirical evidence that highlights a systematic correlation between the outcomes of research and the interests of those funding it. So far, these empirical studies have been most common in the biomedical domain, where researchers have compared a number of studies in which the funder had a financial interest in the study outcome (e.g., it manufactured one of the drugs under investigation) to similar studies in which the funder did not have such an interest. A meta-analysis of eleven previous studies that performed comparisons of this sort found that research studies funded by those with an interest in the outcome were almost four times as likely to yield results that favored the funders than comparable studies that were not funded by an interested party (Bekelman et al. 2003; see also Sismondo 2008). Other research outside the biomedical domain has yielded similar results. A comparison of 206 previous nutritional studies on the health effects of soft drinks, juice, and milk found that research funded by an entity with a financial stake in the outcome was four to eight times as likely to produce results favorable to the funder as compared to similar research funded by a disinterested party (Lesser et al. 2007). Similarly, an analysis that compared previous studies of four worrisome industrial chemicals found that 6 out of 43 studies funded by industry organizations yielded evidence that the chemicals were

problematic, whereas 71 out of 118 studies funded by organizations without a major stake in the outcome provided evidence of harm (Fagin et al. 1999, 51).²

Admittedly, a number of different causal factors can lie behind these sorts of correlations (Schafer 2004; Sismondo 2008). Some of these factors do not challenge the credibility of individual research studies. For example, academic scientists might be more likely to publish their findings when they obtain evidence of harm than when they do not, whereas privately funded scientists might be inclined to do the opposite. Similarly, government-funded scientists might typically have less money to spend on drug comparisons than privately funded scientists; therefore, the privately funded scientists might tend to perform larger studies that are more likely to generate positive results (i.e., results that show a difference in the effectiveness of two treatments) (Hochster 2008). Nevertheless, given the additional evidence that there has been ghost writing, selective reporting of data, and strategic design and interpretation of studies by industry groups, these correlations between funding sources and study results do raise concerns about the impact of financial COIs on the credibility of scientific research (Sismondo 2008).

As noted in the introduction, a number of commentators have responded to these concerns by suggesting (either explicitly or implicitly) that research funded by those with an interest in the outcome should be treated with skepticism. For example, numerous authors have argued that governments should provide more funding for research through agencies like the U.S. National Institutes of Health (NIH) or National Science Foundation (NSF) so that decision makers do not have to depend to such a great extent on research funded by an interested party (see e.g., APHA 2003; Elliott 2011; Shrader-Frechette 2007). Prominent journals, such as the *Journal of the American Medical Association* (JAMA), have also enacted policies requiring that publications include at least one author separate from industry who is primarily responsible for data collection, management, and analysis (DeAngelis and Fontanarosa 2008). In response to this skepticism, others have insisted that this focus on the source of funding is illegitimate and that research should be evaluated solely or primarily based on its scientific quality (Borgert 2007; McCarty et al. 2012). The Society for Toxicology has explicitly stated: “Research should be judged on the basis of scientific merit, without regard for the funding source or where the studies are conducted (e.g., academia, government, or industry)” (Society of Toxicology 2008).

With the goal of evaluating research without considering its source of funding, figures like Conrad and Becker (2011) have proposed other criteria that can be used for assessing research credibility. This poses an ideal opportunity for philosophers of science to perform socially engaged work. Assessing the quality of scientific methodology and the level to which scientific conclusions are confirmed has been a traditional focus of the philosophy of science. With this in mind, the next two sections of the paper examine the criteria for credibility proposed by Conrad and

² It should be noted that two of the 118 studies listed as being funded by organizations “without a major stake in the outcome” were actually supported by labor organizations, which could potentially be characterized as having an interest in finding that particular chemicals are harmful.

Becker as well as the feasibility of considering funding sources as an additional criterion for credibility.

3 Evaluating Criteria for Research Credibility

As the introduction to this paper emphasized, there are no ideal solutions to the problems associated with financial COIs in scientific research. However, as we have already seen, Conrad and Becker (2011) suggest that the most promising approach is to develop criteria that can be used to evaluate the credibility of scientific studies, independently of their source of funding. Their article provides an excellent opportunity to evaluate this strategy for assessing credibility, because they synthesize five previously proposed sets of criteria. Thus, while Conrad and Becker's piece provides only one example of an attempt to formulate criteria for research credibility, it draws on a good deal of the cutting-edge work on this topic. Based on their analysis, Conrad and Becker argue that there is a growing consensus on ten criteria. Unfortunately, we will ultimately see that they are less effective than they initially appear.

Conrad and Becker claim that research is credible if it "faithfully reflects what was observed" and "is not the result of unconscious bias or intentional manipulation, and is thus believable" (Conrad and Becker 2011, 758). In order to develop their criteria for credibility, they draw from three papers that were presented in a session at the annual Meeting of the Society for Risk Analysis in December 2009. These papers were based on work by the Bipartisan Policy Center (BPC), by the International Life Sciences Institute (ILSI), and by Henry and Conrad (2008). In addition, Conrad and Becker pull from criteria developed by the International Agency for Research on Cancer (IARC) and the Federation of American Societies for Experimental Biology (FASEB). They identify ten criteria, each of which appears in at least one of the five sources, and most of the criteria appear in multiple sources. The resulting list (in the order provided by Becker and Conrad, which does not appear to be in order of importance) looks like this:

1. Disclosure of funding sources and other competing interests
2. PI is legally guaranteed: (a) freedom to publish; (b) authority to analyze and interpret results; (c) control of study design
3. Public release of data and methods
4. Factual and transparent research objective and appropriate research design
5. Peer review
6. Prior listing in a public registry (where exists)
7. No linkage of remuneration with outcome of experiment
8. Disclosure of paid "name lending"
9. Maintain clarity between CRO and academic auspices
10. External review of research program

Conrad and Becker emphasize that their list of criteria does not include the source of funding for a study; they argue that this would be "antithetical to science," insofar as it would put too much emphasis on the circumstances of the scientist rather than

letting the scientific facts “speak for themselves” (Conrad and Becker 2011, 758; Borgert 2007).

The remainder of this section evaluates Conrad and Becker’s list of ten criteria to see if they are adequate to replace the consideration of funding sources. Let us start by trying to get clearer about how exactly each of the criteria provide evidence for a study’s credibility. Conrad and Becker are arguably mixing together three different sorts of criteria, which assess three distinct phenomena (see Table 1). First, some of the criteria are what I will call “confirmatory,” insofar as they assess whether a study’s conclusions are well supported by the available evidence. A second set of criteria are “preventive,” insofar as they assess whether steps have been taken to prevent or discourage problematic biases. Finally, criteria in the third set are “facilitative,” insofar as they assess whether the conditions are optimal for the confirmatory criteria in the first category to be applied.

Consider first the confirmatory criteria. Of the ten proposed by Conrad and Becker, the criterion that most obviously assesses whether a study’s conclusions are well confirmed is whether it has a factual and transparent research objective and appropriate research design for achieving that objective. Two other criteria could also be classified in the confirmatory category, but in a more indirect manner. The criteria of peer review and external review of the research program do not directly indicate that the conclusions of a study are well confirmed, but the idea behind peer review and external review is that studies will not get past these hurdles if they do not have appropriate research designs for achieving their objectives. Surprisingly, though, the other seven criteria are not easy to place in this category.

Most of them fall in the preventive category, insofar as they assess the extent to which steps have been taken to prevent or discourage problematic biases. For example, the criterion that the PI be legally guaranteed control over publication, interpretation of results, and study designs seems to be designed to prevent funders from having the power to influence studies in inappropriate ways. The criterion of whether a study is listed in a public registry is also designed to prevent a particular sort of problem, namely, publication bias. By listing a study in advance, those funding the study cannot as easily “bury” it, even if it yields results that the funder does not like. The criterion that forbids linking remuneration to study outcomes is also designed to lessen the motivation for those performing studies to obtain specific

Table 1 Categorization of Conrad and Becker (2011) criteria for research credibility

Confirmatory	Preventive	Facilitative
Factual and transparent objective and appropriate research design	PI legally guaranteed control	Disclosure of funding sources
	Prior listing in public registry	Disclosure of name lending
Peer review	No linking of remuneration with study outcome	Maintain clarity between CRO and academic auspices
External review of research program	Disclosure of funding sources	Public release of data
	Disclosure of name lending	
	Maintain clarity between CRO and academic auspices	

results. Finally, there are three criteria that focus on the disclosure of information (i.e., about funding sources, name lending, and investigators' relationships with CROs and academia). At least part of the motivation for requiring the disclosure of this information is so that researchers who have these COIs are kept on their toes and avoid being flagrantly biased (Elliott 2008).

In addition to the goal of preventing serious biases, the criteria that focus on disclosing information arguably have another purpose. They also serve as "facilitative" criteria, insofar as they help to assess whether conditions are optimal for confirmatory criteria to be applied. One of the goals of requiring that information about funding sources and name lending be disclosed is that those who receive this information can then scrutinize the resulting studies with extra care. Thus, these criteria help to promote an environment in which the scientific quality of research will be evaluated carefully. The criterion that fits even more perfectly in the facilitative category, however, is the public release of data and methods. The release of this information does not guarantee that a study is of high quality, but it does make it much more feasible for others to evaluate the quality of the study. And, in fact, a number of commentators have recently argued that this is a crucial strategy for responding to financial COIs (McGarity and Wagner 2008; Michaels 2008).

With this understanding of how the various criteria work, we are better equipped to evaluate their effectiveness and to determine whether they can be used as an adequate alternative to considering funding sources. Two lessons emerge from this analysis. First, most of the criteria are actually fairly similar to the criterion of funding sources, insofar as they measure phenomena that are merely correlated with credibility rather than measuring credibility itself. Second, there is very little evidence that the criteria are effective at identifying research that is not credible. With these lessons in mind, this section concludes that it is advisable for those seeking to assess the credibility of research to supplement Conrad and Becker's current list of criteria with the additional criterion of funding sources in order to improve their assessments.

Let us first consider the nature of these criteria. We have seen that they vary in terms of what they assess. Ideally, criteria for assessing credibility would measure a characteristic that is actually constitutive of credibility. For example, the credibility of a study consists in the trustworthiness of the evidence reported by the study and the extent to which the evidence actually supports the conclusions that are drawn from it. Therefore, the most ideal criteria for assessing a study's credibility would actually measure the quality of the evidence that it reports and the level of support that the evidence provides for the study's conclusions. In the case of Conrad and Becker's ten criteria, however, most of them are fairly indirect, in the sense that they measure phenomena that are merely correlated with credibility to some extent. For example, the facilitative criteria assess the extent to which other experts are given the information that they need in order to determine whether studies are credible. The availability of this sort of information does not, by itself, indicate anything about a study's credibility. Researchers could perform a very poor study and still provide the information needed for others to assess its quality. It seems plausible that scientists are more likely to perform high-quality studies if others can easily

assess the quality of their work, but this merely establishes a *correlation* (not a constitutive relationship) between the facilitative criteria and study credibility.

The preventive criteria are also fairly indirect, in the sense that they merely assess whether steps were taken to prevent biases; they do not assess whether particular biases were actually eliminated from the studies under consideration. For example, it does seem plausible that studies are less likely to be inappropriately manipulated if the PIs are legally guaranteed control over the analysis and interpretation of the results. But there is no inherent connection between the PIs' control over the studies and their credibility; PIs could still either intentionally or unintentionally engage in poor study interpretations even when they have complete control. Therefore, as in the case of the facilitative criteria, the preventive criteria are not constitutively related to study credibility but are merely correlated with it.

At first glance, the category of confirmatory criteria appears to be far preferable to the other two categories, insofar as the criteria in this category are designed to assess whether a study's conclusions are well supported by the evidence. Therefore, one might think that they would have a constitutive relationship with credibility. Unfortunately, the situation is not that simple. Some of the criteria that are labeled as "confirmatory" for the purposes of this paper actually focus on proxies for evidential support rather than evidential support itself. The criteria of peer review and external review arguably fit this description; the extent to which a study has undergone peer review or external review serves as a proxy for the evidential quality of a study, insofar as reviewers will typically not approve a study if it is poorly designed or interpreted. But because these criteria focus on proxies, the relationship between the criteria and study credibility is again a matter of correlation rather than a constitutive relationship.³

Even if the other confirmatory criteria focus on proxies rather than evidential support itself, the final one (namely, having a factual and transparent research objective and appropriate research design) does appear to have a genuine constitutive relationship with study credibility. In large part, what it means for a study to be credible is that it is designed in a manner appropriate to fulfill its objective. However, while the criterion itself appears to have a constitutive relationship with the concept of credibility, it is still very difficult to measure the extent to which a study actually does have an appropriate research design. Conrad and Becker offer four suggestions, each of which turns out to involve a somewhat unreliable proxy measure.

Their first suggestion is to examine whether a study conformed to good laboratory practice (GLP) guidelines. However, GLP guidelines were developed primarily as a way to maintain the integrity of scientific data in response to scandals in which industry-funded studies included fabricated or falsified information (Michaels 2008; Myers et al. 2009). Therefore, the GLP guidelines do not ensure that studies are interpreted correctly or that they are designed appropriately to support the conclusions drawn from them; they ensure only that the reported study

³ Some commentators might argue that if criteria focus on proxies for evidential quality rather than evidential quality itself, they should not even be placed in the category of confirmatory criteria. But this is just a semantic question of how one chooses to define the categories; the crucial issue is to clarify how each of Conrad and Becker's ten criteria actually does or does not support study credibility.

design was actually followed and that the study data were reported accurately (Henry and Conrad 2008).⁴ Conrad and Becker might respond that most GLP studies are conducted according to agency-approved, validated study protocols, which helps to ensure appropriate study designs. But these protocols still provide enough flexibility so that one can influence study design and interpretation in a manner that favors one's preferred outcomes. For example, when performing chemical safety studies, researchers typically have leeway to choose from among several strains or species of animals, and in some cases they also have significant latitude when selecting doses and statistical analyses (Elliott and Volz 2012).

Conrad and Becker's second suggestion for identifying studies with appropriate research designs is to look for studies that employ good epidemiological practices (GEPs). Unfortunately, these guidelines fall prey to the same problem as GLPs; namely, they are geared primarily toward ensuring that reported study designs are actually followed, not to ensure that those study designs adequately support the conclusions that are drawn from them. Moreover, in his discussion of financial conflicts of interest in safety studies for industrial chemicals, David Michaels (2008) argues that GEPs were developed at least partially because of instigation from industry groups. According to Michaels, these groups found that they could use GEPs as a ploy to require high standards of evidence for acknowledging causal relationships between hazardous substances and health effects, even if lower standards of evidence would be appropriate in many social contexts.

Conrad and Becker's third suggestion for identifying studies with appropriate research designs is to consider whether outside experts have been involved in choosing them. But there are two problems with this suggestion. First, we have already seen that those with financial COIs can strategically marshal groups of experts who are sympathetic to their point of view and who will advocate for study designs that tend to support preferred conclusions (Michaels 2008). The second problem is that even when study protocols are specified by government regulations and are thus less likely to reflect the preferences of a narrow range of interest groups, there is still generally some design flexibility that can be exploited by those with financial COIs (Elliott and Volz 2012). Unfortunately, Conrad and Becker's fourth suggestion is inadequate as well. They suggest that adherence to the federal Common Rule, including use of an Institutional Review Board to evaluate studies, can be an additional guide to identifying appropriate study designs. However, while IRB scrutiny can be helpful to some extent, it (as well as the Common Rule as a whole) is focused primarily on protecting the rights and safety of human research participants and only secondarily on ensuring optimal study designs.

In sum, the first lesson to draw from this scrutiny of Conrad and Becker's criteria is that all three categories (facilitative, preventive, and confirmatory) rely primarily on correlations between what the criteria measure and what it actually means for a study to be credible. But there are good reasons why Conrad and Becker (2011)

⁴ One might argue that if following GLP guidelines ensures that the evidence reported in a study is trustworthy, then it does in fact provide important information related to a study's credibility. This may indeed be the case, but it still provides only part of the evidence needed to evaluate a study's credibility. A study with very trustworthy data but an inappropriate study design could not be classified as providing credible conclusions.

adopted the sorts of criteria that they did. While the ideal approach for determining whether a study is credible would be to focus directly on the quality of the evidence reported and the extent to which it supports the study's conclusions, this approach is impractical when evaluating large numbers of studies. Regulatory agencies are overburdened as it is, so they often have to rely on the sorts of criteria proposed by Conrad and Becker (e.g., requiring that submitted studies conform to standardized protocols). And even if the experts in regulatory agencies had more time to scrutinize studies, they might still miss design flaws that serve the goals of powerful interest groups. For example, Richard Smith, the former medical journal editor who was mentioned earlier in this paper, admitted that it was almost impossible for the expert peer reviewers associated with his journal to keep up with the clever strategies employed by pharmaceutical companies when testing their products. He noted, "there are many ways to hugely increase the chance of producing favorable results, and there are many hired guns who will think up new ways and stay one jump ahead of peer reviewers" (2005). Considering that expert scientists face these sorts of difficulties when attempting to evaluate the credibility of studies directly, it is no wonder that regulatory agencies, citizens, policy makers, and even other scientists depend on the sorts of criteria proposed by Conrad and Becker.

Based on this analysis, one might conclude that even though Conrad and Becker's ten criteria are based solely on correlations with credibility, we should remain satisfied with them. But the second lesson to draw from this analysis of the criteria is that there is very little evidence that they are effective in distinguishing studies that are credible from those that are not. In principle, the fact that the criteria depend on correlations between credibility and other phenomena need not imply that the criteria are ineffective. One could potentially develop criteria that rely on tight correlations between two phenomena. But it is not clear that the criteria identified by Conrad and Becker identify sufficiently tight correlations to provide reliable evidence concerning a study's credibility.

Consider first the facilitative and preventive criteria. Disclosure of various pieces of information (e.g., funding sources, name lending, the PI's status with CROs or academia) is central to many of these criteria. But psychological research suggests that when people disclose conflicts of interest, they may actually feel more comfortable about being biased than when they do not disclose them (Cain et al. 2005; Elliott 2008). This research also indicates that those who find out about these COIs have a great deal of difficulty determining how much to discount the information that they receive from those with the COIs (Cain et al. 2005; Loewenstein et al. 2012). Therefore, both the preventive and facilitative effectiveness of these disclosure-based criteria is in doubt. As for the criterion of disclosing the raw data from a study, it is a helpful facilitative criterion (i.e., it indicates that the conditions are fairly optimal for assessing a study's credibility), but there is little evidence to show that it is correlated with credibility itself. One might suggest that PIs are likely to be intimidated by the prospect of disclosing their raw data and that they will engage in unbiased interpretations of their data as a result. But at present this is merely speculation and is not established by empirical evidence; we have already seen that the disclosure of information may not have the effects that we expect (Cain et al. 2005).

As for the criteria of guaranteeing the PI's control over studies and eliminating links between funding and study outcomes, their effectiveness rests on the assumption that PIs are able and motivated to avoid problematic biases. But this assumption is dubious. For one thing, when industry hires organizations like product defense companies to perform research, it is understood that the companies' investigators will be motivated to produce results that are favorable to industry, even if they are not contractually obligated to do so (Michaels 2008). And even if the investigators attempt to be unbiased, psychological research indicates that people significantly overestimate the extent to which they can overcome the subtle, unconscious effects of COIs (Loewenstein et al. 2012; Moore et al. 2005). Finally, the criterion of listing studies in a registry is, at best, effective only in eliminating one particular source of bias (namely, publication bias). Thus, it is an unreliable indicator of study credibility as a whole.

Conrad and Becker's confirmatory criteria are little better. With respect to peer review, there is a growing body of literature suggesting that it has not been studied adequately and that it is not very reliable or effective in weeding out poor-quality studies (Elliott 2011, 104; McCarty et al. 2012). Another problem with both peer review and external review is that industry groups have allegedly created their own journals and conferences and blue-ribbon panels in order to "rubber stamp" findings that they find appealing (Michaels 2008). Finally, we have seen that the proxies that Conrad and Becker employ to assess whether a study has an appropriate research design for achieving its objective are likely to be unreliable. Whether a study abides by GLP or GEP standards or has undergone IRB review, this has little to say about whether the study is designed adequately. When a study has received external review, this may be a more reliable indicator that the study design is appropriate, but we have seen that this depends on the independence and thoroughness of the external review.

So, to sum up this evaluation of Conrad and Becker (2011) article, we have found that it is more difficult than it initially appears to develop an effective set of criteria for evaluating the credibility of research studies without looking at funding sources. All of the criteria that Conrad and Becker identified from previous literature depend on correlations between credibility and other phenomena. This is understandable, because directly assessing the credibility of studies on a large scale would be impractical. Nevertheless, these criteria are not just based on correlations; they are based on presumed correlations that are not well supported by available evidence. Therefore, it seems unwise to dismiss an additional criterion that appears to have at least some empirical support.

But Conrad and Becker might respond that using funding sources as a criterion of credibility would be antithetical to science. In their article, they appeal to Christopher Borgert's claim that scientific facts should speak for themselves and that the scientist should be regarded as a mere accessory in the collection of this evidence (Borgert 2007; Conrad and Becker 2011, 758). This is a very dubious claim. First, scientific facts do not speak for themselves; researchers have to actively collect, characterize, analyze, and interpret their data (Longino 2002). Second, there are few activities more central to science than identifying and correcting for sources of error (Mayo 1996). Given that social scientists have found that scientists suffer

from a wide variety of biases and cognitive shortcomings, including overconfidence biases and influences from conflicts of interest (Cain et al. 2005; Elliott 2009; Loewenstein et al. 2012; Moore et al. 2005; McGarity and Wagner 2008), it would actually be antithetical to science to ignore these biases and fail to respond to them. Just as it is legitimate (and, indeed, advisable) for scientists to identify and compensate for biases or systematic errors that show up in their measuring instruments, it is legitimate to take steps to identify and compensate for biases in scientists themselves.⁵

A final objection to including the source of funding as an additional criterion when considering study credibility is that we do not have precise data about how well it is correlated with credibility. After all, there are presumably many cases in which funders have an interest in the outcome of research and in which the research that they fund is highly credible (Resnik and Elliott 2013). Nevertheless, there is pervasive evidence that funders with a stake in the outcome of research have often influenced the research in ways that diminish its credibility (see Sect. 2, as well as Angell 2005; Markowitz and Rosner 2002; McGarity and Wagner 2008; Michaels 2008). Our knowledge of the frequency of these influences in various contexts is, admittedly, seriously limited. But this objection (i.e., that we do not have precise information about the correlation between funding sources and study credibility) would be compelling only if we could make use of other criteria that would serve us better. It turns out that almost all the other criteria proposed by Conrad and Becker suffer from the same problem (and perhaps to an even greater degree)—we do not have good empirical evidence about the correlations between study credibility and legally guaranteeing PIs control over their studies or performing peer review or disclosing funding sources or maintaining clarity between CRO and academic auspices. In this situation, all we can do is to employ as many potentially helpful criteria as possible and to refine them over time.

4 Refining the “Funding” Criterion

The previous section argued that the attempt to ignore research funding and to scrutinize the credibility of studies more directly is not promising. The criteria proposed by Conrad and Becker depend on correlations between credibility and other phenomena; they do not measure study credibility directly. Moreover, there is little empirical evidence to support the strength of these correlations. Therefore, given the evidence that funding sources can have a significant influence on research results, it seems wise to include funding source as an additional criterion for assessing credibility. But we also saw at the end of the previous section that this criterion, like the others proposed by Conrad and Becker, could benefit from a great deal of refinement and elaboration. If one were to treat all research funded by an interested party with suspicion, one would be forced to be suspicious of a very large quantity of research, even though much of it is actually credible (Resnik and Elliott 2013). And even if one did treat it with suspicion, it is not clear what one would do.

⁵ I thank David Resnik for helpful insights related to this paragraph.

We have already seen that government agencies do not have resources to scrutinize it all in detail. And it would be totally impractical to reject all this research entirely, because industry pays for well over half of all R&D expenditures (Elliott 2011).

A promising approach for addressing this difficulty is to develop a more refined criterion that focuses on more than just the source of funding for a study; we should identify more detailed conditions under which financial COIs are most likely to corrupt the quality of scientific work (Resnik and Elliott 2013). As a starting point for further work on this issue, consider three conditions that are all present in previous cases where financial COIs have been particularly problematic:

1. Scientific findings are ambiguous or require a good deal of interpretation or are difficult to establish in a straightforward manner.
2. Individuals or institutions have strong incentives to influence those scientific findings in ways that damage the credibility of the research.
3. Individuals or institutions that have incentives to influence those scientific findings also have adequate opportunities to influence them.

There may be other conditions that are also sufficient to damage the credibility of research.⁶ However, all three conditions are met in virtually every high-profile case in which financial COIs have caused significant problems (e.g., climate change, the ozone hole, acid rain, smoking, second-hand smoke, industrial chemicals, and pharmaceuticals). It is crucial to remember, however, that research performed under these three conditions can still be exemplary. The presence of these three conditions does not by any means guarantee the corruption of research. But if one is trying to identify the circumstances under which corruption is most likely to occur, these appear to be important conditions to look for.

One major benefit of clarifying these conditions is that it enables those who are responding to financial COIs to be somewhat more precise in identifying areas of research where they might suspect a lack of credibility. By paying attention to these conditions one can avoid unfairly targeting all industry-funded work as unreliable. For example, a great deal of industry research does not meet the second condition (i.e., that there be incentives to damage the credibility of research). As long as scientific mistakes harm research funders in the same ways that they affect the general public, then there is little incentive to damage the credibility of research. For example, if scientific mistakes will result in a product that obviously malfunctions, then the industry that funds the scientific work will have an incentive to get the science right, or nobody will buy the product. Thus, when companies are doing research to improve the performance of new transistors or fuel cells or smart phones or batteries or stain-resistant pants or comfortable shoes, there is often little reason

⁶ While this paper proposes three conditions that help to determine whether financial COIs challenge study credibility, Resnik and Elliott (2013) propose five conditions. Nevertheless, the conditions in both papers are closely related. The second condition in this paper (that individuals have strong incentives to influence scientific findings) is elaborated in terms of three criteria in Resnik and Elliott (2013): whether the funders have a significant financial stake in study outcomes, whether their financial interests coincide with the goal of producing credible research, and whether the funders have a history of influencing research. The third condition in this paper (that individuals have opportunities to influence scientific findings) is elaborated in terms of two conditions in Resnik and Elliott (2013): whether it is easy to manipulate research and whether there are oversight mechanisms to prevent bias.

to be suspicious of the quality of this work. But when companies turn to studying subtle environmental or human health risks associated with these products, then there are clear incentives to corrupt the research, because many harmful effects of the products will not be easy for consumers to recognize.⁷

By being more precise in identifying worrisome areas of research, those who respond to financial COIs can also avoid focusing solely on industry-funded work and failing to identify other suspicious scientific work. For example, some areas of government-funded science are likely to meet the three conditions that tend to threaten the credibility of research. Consider cases in which the Department of Energy funds research to determine whether its facilities are disposing of nuclear waste safely, or in which the Department of Defense studies whether its installations pose hazards to those living in the vicinity, or in which the Department of Agriculture examines the health effects of meat and dairy products. In these cases, government agencies experience worrisome incentives that are similar to those experienced by industry (Michaels 2008), although political pressures may be even more significant than financial ones for these agencies. Non-profit citizens' groups and labor organizations also sometimes face similar incentives, such as when environmental organizations are motivated by ideological considerations to exaggerate the hazards of chemicals or activities. However, attention to the three conditions also suggests that at least some of these groups may not pose as widespread or as intense a threat to research credibility as industry or government. This is partly because they often have fewer financial resources (and thus less opportunity to influence research) and partly because they often have less money at stake in the outcome (and thus less incentive to influence research).⁸ Nevertheless, it would clearly be valuable to engage in further social scientific research to determine how likely it is for research to be compromised when various sorts of incentives are present, such as the ideological incentives experienced by those affiliated with environmental organizations, or the political pressures faced by government agencies, or the financial incentives experienced by those who work in industry.

A second benefit of clarifying the three conditions under which financial COIs are particularly worrisome is that policy makers and administrators can take active steps to prevent one or more of the conditions from being met, thereby lessening the amount of research that suffers from a lack of credibility. To see how one could identify and alter one or more of the three conditions, let us consider the process by which industrial chemicals and pharmaceuticals are regulated. Currently, there are two general approaches to regulating chemicals in the United States (Cranor 2011). For pesticides and pharmaceuticals, which are regulated under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) and the Federal Food, Drug, and Cosmetic Act (FD&C), manufacturers must submit a package of safety studies so that the chemicals can be approved before being marketed. For most other chemicals, which are regulated under the Toxic Substances Control Act (TSCA), no

⁷ I thank Martin Carrier and Torsten Wilholt for very helpful insights on this issue.

⁸ While these citizens' groups do not have as many financial resources as industry, they can partially compensate for this limitation with powerful media and communications strategies. And because they are more likely to be trusted by the public than industry, these groups can have a significant impact on public risk perceptions.

pre-approval is required before marketing them, and the EPA has to find evidence that they are harmful in order to remove them from the market (Cranor 2011). While some incentives are different in the two cases, they are similar in that the manufacturers fund the vast majority of safety tests. In the past, manufacturers would have performed the studies “in-house,” whereas in recent years they have moved toward outsourcing these studies to contract research organizations (CROs) (Lenzer 2008; Mirowski and Van Horn 2005; Volz and Elliott 2012).

This regulatory situation meets all three conditions that can challenge the credibility of research. First, providing definitive scientific evidence that particular chemicals cause harm is exceedingly difficult. In general, this first condition (i.e., that scientific evidence is ambiguous) may be the most difficult of the three discussed in this paper to avoid; after all, science-studies scholars have emphasized that if the stakes are high enough, almost any research can be called into question and challenged (Sarewitz 2004). But in the case of chemical safety testing, a variety of factors make the science particularly difficult. For one thing, it is generally unethical to deliberately expose humans to significant quantities of potentially toxic chemicals under controlled conditions. While it might be acceptable to employ very low doses that are not expected to cause significant harm (Resnik 2006a), such studies are likely to underestimate risks, especially those associated with higher levels of exposure. Therefore, scientists are forced to employ sub-optimal techniques for identifying potential hazards. These techniques can include a wide range of *in vitro* assays, animal experiments, and epidemiological studies. Whatever the technique, there is often room for differences in interpretation, for reanalyzing data, and for altering studies in ways that promote a preferred outcome (Michaels 2008). Moreover, once chemicals are approved, any harmful effects from them are often exceedingly difficult to trace back to their sources, so it could be years (if ever) before various effects are convincingly traced back to them (Cranor 2011).

The second condition (i.e., that individuals or institutions have incentives to influence the research outcomes) is also easily satisfied in these cases. Pesticide manufacturers have often invested tens of millions of dollars in R&D costs by the time they submit their products for regulatory approval, and for pharmaceutical manufacturers the dollar figures are often in the hundreds of millions. Given that it is so difficult to identify harmful effects once products are on the market, there are huge financial incentives for companies to influence safety studies so that they can get their products approved. One might think that the threat of tort lawsuits would provide incentives against introducing harmful products to the market, but scholars have argued that the current incentive structure of the tort system tends to be inadequate for this purpose (Cranor 2008; McGarity and Wagner 2008).

The third condition (namely, that those with incentives to damage the credibility of research have opportunities to do so) is also satisfied in the case of chemical safety testing. When companies perform safety studies in-house, the same researchers who have the power to make crucial interpretive and study-design choices also have incentives to satisfy their superiors. One might think that these problems would be reduced by outsourcing safety studies to CROs. However, while this creates an institutional separation between those with the most incentive to approve products (i.e., the management of chemical or pharmaceutical companies)

and those who make crucial study-design choices (i.e., employees of CROs), it does not eliminate the problems. The remaining difficulty is that the business model for CROs revolves around developing long-term relationships with chemical and pharmaceutical companies so that they can continue to obtain their business. Thus, the management and employees of CROs still end up with both the opportunity and the incentive to make subtle methodological choices in ways that benefit the chemical manufacturers (Lenzer 2008; Volz and Elliott 2012). It is crucial to note in this context that the employees of chemical companies and CROs can be influenced by COIs and financial incentives even if they have the very best of intentions to resist those influences. Psychologists tell us that we are all much worse at resisting the influences of COIs than we typically think (Cain et al. 2005; Loewenstein et al. 2012; Moore et al. 2005).

But chemical safety testing also illustrates how an awareness of the three conditions under which financial COIs are most worrisome can assist in altering incentive structures and institutions so that one or more of the conditions can be eliminated. One of the most obvious ways to solve the problem is to eliminate the first condition; this requires changing regulatory policies so that they either depend less on scientific results or so that they depend on much more straightforward findings. An excellent example of this strategy comes from the Toxic Use Reduction Act that the state of Massachusetts passed in 1989. This law created a list of suspicious chemicals, and companies that use large amounts of those chemicals have to prepare a plan that documents why they need to use them and whether they have other options. They also have to make public the quantity of those chemicals being used. The beauty of this approach is that it shifts the regulatory focus away from debates about the details of scientific studies on these suspicious chemicals. Instead of incentivizing industry to “manufacture uncertainty” and to grind regulatory decision making to a halt by exploiting ambiguous science about those substances (Michaels 2008), it encourages industry to identify new alternatives.

Unfortunately, one might still worry that in the process of searching for new alternatives, the chemical industry would once again encounter ambiguous evidence about the safety of new products, and they might again work to manipulate this science. This problem illustrates that in today’s knowledge-intensive economy, it is difficult to avoid depending on scientific information when making decisions, and so it is also important to examine ways to avoid the second and third conditions that can contribute to compromised science. In the case of chemical safety testing, the second condition (i.e., the presence of incentives to corrupt science) is somewhat difficult to change, insofar as there are major incentives to influence studies so that they downplay subtle risks. One could alleviate some of these incentives by shifting from a post-market regulatory scheme to a pre-market scheme (see Cranor 2011). By doing so, companies would have less incentive to “manufacture uncertainty” to keep worrisome chemicals on the market. Another strategy would be to tweak legal policies so that those who manipulate science are more likely to lose tort suits, to pay stiff penalties, to face public embarrassment, and even to suffer criminal penalties (Cranor 2011; McGarity and Wagner 2008). However, neither of these strategies is likely to entirely eliminate incentives for corrupting science.

The third condition (i.e., that those with incentives also have the opportunity to influence the science) may be somewhat easier to change, at least in the case of chemical safety studies. Numerous authors have argued that safety testing for pharmaceuticals and industrial chemicals should be set up in such a way that those with the greatest interest in the outcome have less control over the studies (see e.g., Cranor 2011; Krimsky 2003; McGarity and Wagner 2008; Michaels 2008; Volz and Elliott 2012). A promising avenue for accomplishing this would be to require that any safety studies performed for regulatory purposes be controlled by a government agency [e.g., the U.S. Environmental Protection Agency (EPA)] or an international organization [e.g., the Organization for Economic Cooperation and Development (OECD)]. Industry would provide the government agency or international organization with funds to perform the study, but the agency would choose the study design and contract with academic labs or CROs to perform them. On this scheme, industry would still have significant incentives to influence the studies, but they would not have control over them. And those with the control (the academic labs or CROs) would have fewer incentives to guide the studies toward predetermined outcomes, because they would continue to receive contracts from the EPA or the OECD no matter what sorts of results they obtained. Of course, industry could still attempt to exert control over the government agency that chooses study designs, but hopefully steps could be taken to insulate those involved in these decisions from such influences.

This analysis of strategies for altering the incentive structure surrounding chemical safety studies provides a model for analyzing other areas of science to eliminate conditions that create the most worrisome COIs. Nevertheless, one might raise a final objection to employing the refined criterion of credibility developed in this section. Namely, even though it exonerates a great deal of industry-funded science (e.g., research to develop products whose malfunctions will be readily apparent), it still leaves almost all industry-funded human-health and environmental-safety studies under a cloud of suspicion. One might think that it is unacceptable to be suspicious of such a broad swath of science, especially because the solutions offered for alleviating this suspicion (e.g., creating a buffer between funders and those performing the studies) are likely to be politically difficult to implement. But this objection is not very compelling. If there are good reasons to be suspicious of the science that industry produces for regulatory purposes, then we should not dismiss this conclusion just because it is unappealing. If we had further criteria for distinguishing the industry-funded studies that are credible from those that are not, then it would certainly make sense to focus only on the studies that are worrisome. But so far it does not appear that we have better criteria. Thus, rather than sticking our heads in the sand and ignoring this problem, the realization that we have reasons to be suspicious of so much scientific research should motivate us to take action. Even if we are not able to enact all the policies proposed in this paper, we can at least explore other creative responses. For example, in some cases industry groups have collaborated with other stakeholders to jointly fund and design scientific studies that are widely trusted (see e.g., Busenberg 1999; Douglas 2005). These sorts of win-win solutions deserve much more attention in the future.

5 Conclusion

The impact of financial COIs on scientific research is a significant social challenge, and philosophers of science are well-equipped to help address a number of the issues raised by this challenge. This paper has tackled the question of how to develop criteria for evaluating the credibility of research. Based on an analysis of Conrad and Becker's (2011) synthesis of five previous sources, it argued that current efforts to evaluate the credibility of research without considering the source of funding are likely to be unreliable or impractical. Of the ten criteria that Conrad and Becker pulled from previous sources, most are "preventive" or "facilitative" and do not measure phenomena that are constitutive of credibility. Even the "confirmatory" criteria that seem to be focused directly on assessing credibility end up depending on proxy measures. But this is not entirely Conrad and Becker's fault; the problem is that regulatory agencies do not have enough resources to engage in the painstaking work of directly assessing the quality of research.

Given that efforts to assess research credibility without considering funding sources are likely to be impractical or unreliable, this paper argued that citizens and policy makers should take funding sources into account as an additional criterion. Moreover, one can refine this criterion by considering the extent to which three conditions are met:

1. Scientific findings are ambiguous or require a good deal of interpretation or are difficult to establish in an obvious and straightforward manner.
2. Individuals or institutions have strong incentives to influence research findings in ways that damage the credibility of research.
3. Individuals or institutions that have incentives to influence research findings also have adequate opportunities to influence them.

When all three conditions are met, it is reasonable to be suspicious of the credibility of the resulting work. Based on this criterion, most industry-funded human-health and environmental-safety studies performed for regulatory purposes fall under a shadow of suspicion. But this should motivate industry, policy makers, and other stakeholders to develop creative solutions that eliminate one or more of the conditions that generate this concern.

Acknowledgments I would like to thank the participants and attendees at the University of Cincinnati colloquium for Socially Engaged Philosophy of Science, the SPSP pre-conference workshop on Science, Policy, and Values at the University of Toronto, and the Science Studies Program at the University of California San Diego for helpful comments on earlier versions of this paper. I received particularly helpful feedback from Craig Callender, Jane Maienschein, David Resnik, David Volz, and two anonymous referees at various points in the development of the paper.

References

- Angell, M. (2005). *The truth about the drug companies: How they deceive us and what to do about it*. New York: Random House.
- APHA (American Public Health Association). (2003). *Supporting legislation for independent post-marketing phase IV Comparative evaluation of pharmaceuticals*. Washington, DC: APHA.

- Beder, S. (2000). *Global spin* (rev ed.). White River Junction, VT: Chelsea Green.
- Bekelman, J., Lee, Y., & Gross, C. (2003). Scope and impact of financial conflicts of interest in biomedical research. *Journal of the American Medical Association*, 289, 454–465.
- Borgert, C. (2007). Conflict of interest or contravention of science. *Regulatory Toxicology and Pharmacology*, 48, 4–5.
- Busenberg, G. (1999). Collaborative and adversarial analysis in environmental policy. *Policy Studies*, 32, 1–11.
- Cain, D., Loewenstein, G., & Moore, D. (2005). The shortcomings of disclosure as a solution to conflicts of interest. In D. Moore, D. Cain, G. Loewenstein, & M. Bazerman (Eds.), *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy* (pp. 104–125). New York: Cambridge University Press.
- Conrad, J., & Becker, R. (2011). Enhancing credibility of chemical safety studies: Emerging consensus on key assessment criteria. *Environmental Health Perspectives*, 119, 757–764.
- Cranor, C. (2008). *Toxic torts: Science, law, and the possibility of justice*. New York: Cambridge University Press.
- Cranor, C. (2011). *Legally poisoned: How the law puts us at risk from toxicants*. Cambridge, MA: Harvard University Press.
- Curzer, H., & Santillanes, G. (2012). Managing conflict of interest in research: Some suggestions for investigators. *Accountability in Research*, 19, 143–155.
- Davis, M. (1982). Conflict of interest. *Business & Professional Ethics Journal*, 1, 17–27.
- DeAngelis, C., & Fontanarosa, P. (2008). Impugning the integrity of medical science: The adverse effects of industry influence. *Journal of the American Medical Association*, 299, 1833–1835.
- Douglas, H. (2005). Inserting the public into science. In S. Maasen & P. Weingart (Eds.), *Democratization of expertise? Exploring novel forms of scientific advice in political decision making* (pp. 153–169). Berlin: Springer.
- Elliott, K. (2008). Scientific judgment and the limits of conflict-of-interest policies. *Accountability in Research*, 15, 1–29.
- Elliott, K. (2009). Respect for lay perceptions of risk in the hormesis case. *Human and Experimental Toxicology*, 28, 21–26.
- Elliott, K. (2011). *Is a little pollution good for you? Incorporating societal values in environmental research*. New York: Oxford University Press.
- Elliott, K., & Volz, D. (2012). Addressing conflicts of interest in nanotechnology oversight: Lessons learned from drug and pesticide safety testing. *Journal of Nanoparticle Research*, 14, 664–668.
- Fagin, D., Lavelle, M., & the Center for Public Integrity. (1999). *Toxic deception* (2nd ed.). Monroe, Maine: Common Courage.
- Healy, D., & Catell, D. (2003). Interface between authorship, industry, and science in the domain of therapeutics. *British Journal of Psychiatry*, 183, 22–27.
- Henry, C., & Conrad, J. (2008). Scientific and legal perspectives on science generated for regulatory activities. *Environmental Health Perspectives*, 116, 136–141.
- Hochster, H. (2008). The power of “P”: On overpowered clinical trials and “positive” results. *Gastrointestinal Cancer Research*, 2, 108–109.
- Krimsky, S. (2003). *Science in the Private Interest*. Lanham, MD: Rowman and Littlefield.
- Lenzer, J. (2008). Truly independent research? *British Medical Journal*, 337, 7670.
- Lesser, L., Ebbeling, C., Goozner, M., Wypij, D., & Ludwig, D. (2007). Relationship between funding source and conclusion among nutrition-related scientific articles. *PLoS Medicine*, 4, e5.
- Loewenstein, G., Sah, S., & Cain, D. (2012). The unintended consequences of conflict of interest disclosure. *Journal of the American Medical Association*, 307, 669–670.
- Longino, H. (2002). *The fate of knowledge*. Princeton: Princeton University Press.
- Markowitz, G., & Rosner, D. (2002). *Deceit and denial: The deadly politics of environmental pollution*. Berkeley: University of California Press.
- Mayo, D. (1996). *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.
- McCarty, L., Borgert, C., & Mihaich, E. (2012). Information quality in regulatory decision making: Peer review versus good laboratory practice. *Environmental Health Perspectives*, 120, 927–934.
- McGarity, T., & Wagner, W. (2008). *Bending science: How special interests corrupt public health research*. Cambridge, MA: Harvard University Press.
- McKaughan, D., & Elliott, K. (2013). Backtracking and the ethics of framing: Lessons from voles and vasopressin. *Accountability in Research*, 20, 206–226.

- Michaels, D. (2008). *Doubt is their product: How industry's assault on science threatens your health*. New York: Oxford University Press.
- Mirowski, P., & Van Horn, R. (2005). The contract research organization and the commercialization of scientific research. *Social Studies of Science*, 35, 503–548.
- Moore, D., Cain, D., Loewenstein, G., & Bazerman, M. (Eds.). (2005). *Conflicts of interest: Challenges and solutions in business, law, medicine, and public policy*. New York: Cambridge University Press.
- Myers, J., vom Saal, F., Akingbemi, B., Arizono, K., Belcher, S., Colborn, T., et al. (2009). Why public health agencies cannot depend on good laboratory practices as a criterion for selecting data: The case of bisphenol A. *Environmental Health Perspectives*, 117, 309–315.
- Oreskes, N., & Conway, E. (2010). *Merchants of doubt*. New York: Bloomsbury Press.
- Proctor, R. (2012). *Golden holocaust: Origins of the cigarette catastrophe and the case for abolition*. Berkeley: University of California Press.
- Resnik, D. (2006a). Intentional exposure studies of environmental agents on human agents: Assessing benefits and risks. *Accountability in Research*, 14, 35–55.
- Resnik, D. (2006b). *The price of truth: How money affects the norms of science*. New York: Oxford University Press.
- Resnik, D., & Elliott, K. (2013). Taking financial relationships into account when assessing research. *Accountability in Research*, 20, 184–205.
- Sarewitz, D. (2004). How science makes environmental controversies worse. *Environmental Science & Policy*, 7, 385–403.
- Schafer, A. (2004). Biomedical conflicts of interest: A defense of the sequestration thesis—Learning from the cases of Nancy Olivieri and David Healy. *Journal of Medical Ethics*, 30, 8–24.
- Shrader-Frechette, K. (2007). *Taking action, saving lives: Our duties to protect environmental and public health*. New York: Oxford University Press.
- Sismondo, S. (2008). Pharmaceutical company funding and its consequences: A qualitative systematic review. *Contemporary Clinical Trials*, 29, 109–113.
- Smith, R. (2005). Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Medicine*, 2(5), e138. doi:[10.1371/journal.pmed.0020138](https://doi.org/10.1371/journal.pmed.0020138).
- Society of Toxicology. (2008). *Principles for research priorities in toxicology*. Available at: <http://www.toxicology.org/pr/PrinResearch.asp>. Last accessed on December 11, 2012.
- Sutton, P., Woodruff, T., Vogel, S., & Bero, L. (2011). Conrad and Becker's 'Ten Criteria' fall short of addressing conflicts of interest in chemical safety testing. *Environmental Health Perspectives*, 119, a506–a507.
- Thompson, D. (1993). Understanding financial conflicts of interest. *New England Journal of Medicine*, 329, 573–576.
- Tweedale, T. (2011). Enhancing credibility of chemical safety studies: No consensus. *Environmental Health Perspectives*, 119, a507–a508.
- Volz, D., & Elliott, K. (2012). Mitigating conflicts of interest in chemical safety testing. *Environmental Science and Technology*, 46, 7937–7938.
- Wilholt, T. (2009). Bias and values in scientific research. *Studies in History and Philosophy of Science*, 40, 92–101.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.