Authors' Note: In an effort to follow our own advice calling for greater transparency in science and publishing, this is an example of the "transparent peer review" process that is being used in some journals such as Nature Communications and the BioMed Central and PLoS families of journals.

The present article went through 2 rounds of review by 4 anonymous reviewers. The full sets of reviews, responses, original, and revised submissions follow in reverse chronological order.

---

7 December 2018

Responses to reviews of Revision 1:
"Scientific Integrity Issues in Environmental Toxicology: improving research reproducibility, credibility, and transparency"

We greatly appreciate the continuing interest and constructive advice from all four reviewers. Our point-by-point responses and explanations of revisions follows. Original comments are in black text and our responses are indented and in blue colored text.

Associate Editor Comments

Comments to the Author:
Thank you again for submitting your paper to IEAM. As the assigned editor, I apologize for the 12-day delay in completion of the second review. That was my fault.

The paper still needs minor revisions, but it does not have to go back to the peer reviewers again. Please address the reviewers' comments on the second draft.

Two reviewers had comments about the tone of the manuscript. Here are a couple examples:

"...the opening sentence about elections is weak... it sets an unfortunate political tone that implies bias."

"Line 760-765: Unnecessary ridicule of researchers that lacks a reference. Remove or rephrase."

These are valid criticisms that should be addressed to improve the impact of the paper.

> The relevant sections of those sentences have been removed to appease the reviewers' criticisms.

****

Reviewer(s)' Comments to Author:

Reviewer: 1

Comments to the Author
I noted that the authors added "chemistry" to the title but forgot to also add that to the rest of the manuscript. Many of the recommendations and examples only address ecotoxicologists, e.g. the section "Education as the way forward".

We also added a specific section on chemistry issues. Changing most usages to “ecotoxicology and chemistry” would be ponderous, and some aspects are specific to ecotoxicology.

Line 66. Remove "ample" for a more balanced view.

We acknowledge that Reviewer 1 may not accept our examples and arguments that “*there is ample room for improvement within our discipline.*” We think the word “ample” is appropriate and was kept

Line 381. Add missing full stop after "breastfeeding". Good catch

Figures: Consider removing them completely. I understand your wish to make the text funny but I do not think that type of figures belongs in scientific papers.

The goal with this paper boils down to a hope that many will critically consider some of these points and spur further discussions and refinement of potential improvements. The approach of using cartoons to illustrate serious points and writing in a less formal way than is customary in science policy literature is aimed at hopefully hooking interest. We realize that by attempting to introduce elements of whimsy or humor into the arguments, some will be put off, but we hope more might continue reading and appreciate it. We think that a long, sincere essay of this sort could be dreadfully sincere, preachy and set aside by most. For example, journals such as Nature commission cartoons to inject humor into serious topics, such as this, this, or that or this other one. Reviewer #1 and others might think such dressings distract and detract from the serious business of science, but others argue for science writing to be less turgid and a drudge for readers. This one aims for the latter.

Line 488. Is the quote meant to be funny? It does not follow your own argumentation at all. Remove the quote.

Amended the lead in to spell it out – the point is that universities are dependent on external industry research funding, and the need for sufficient funding can trump concerns by administrators over the color of money. And some (obviously not all) readers might appreciate a memorable turn of phrase.

Line 634: To increase the relevance further, scientists should also read regulatory assessments (or is that included in "available literature"?). Suggested reference: https://www.ncbi.nlm.nih.gov/pubmed/28452384

We had overlooked the Ågerstrand paper. It is a good reference and we added it to the “Relevance” subheading.

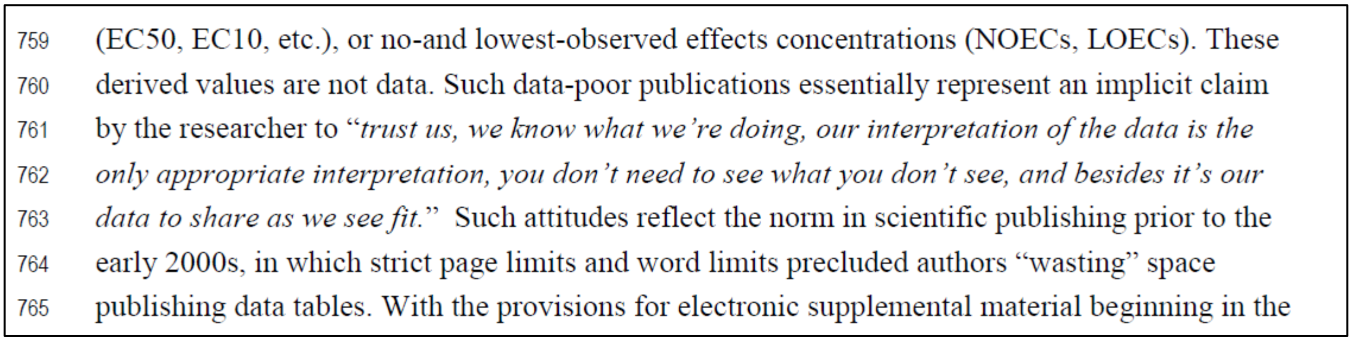
Line 668: Suggested references for this section: https://www.ncbi.nlm.nih.gov/pubmed/29960649 https://www.ncbi.nlm.nih.gov/pubmed/30075188

OK

Line 760-765: Unnecessary ridicule of researchers that lacks a reference. Remove or rephrase.

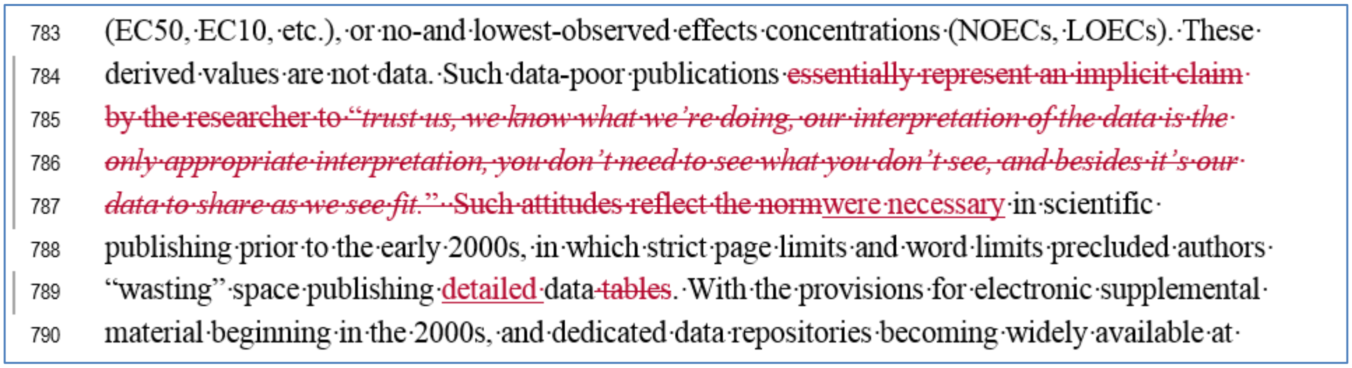
While we are unapologetic about trying to goad authors who can’t be bothered to show their supporting data or work to change their ways, we agree the use of scare quotes confuse our strong criticisms as being a quote from others. Thus, we edited it out

The snip of the original text is pasted below:



(EC50, EC10, etc.), or no-and lowest-observed effects concentrations (NOECs, LOECs). These derived values are not data. Such data-poor publications essentially represent an implicit claim by the researcher to “*trust us, we know what we’re doing, our interpretation of the data is the only appropriate interpretation, you don’t need to see what you don’t see, and besides it’s our data to share as we see fit.*” Such attitudes reflect the norm in scientific publishing prior to the early 2000s, in which strict page limits and word limits precluded authors “wasting” space publishing data tables. With the provisions for electronic supplemental material beginning in the

And the revised text:



(EC50, EC10, etc.), or no-and lowest-observed effects concentrations (NOECs, LOECs). These derived values are not data. Such data-poor publications ~~essentially represent an implicit claim by the researcher to “*trust us, we know what we’re doing, our interpretation of the data is the only appropriate interpretation, you don’t need to see what you don’t see, and besides it’s our data to share as we see fit.*” Such attitudes reflect the norm~~were necessary in scientific publishing prior to the early 2000s, in which strict page limits and word limits precluded authors “wasting” space publishing detailed data ~~tables~~. With the provisions for electronic supplemental material beginning in the 2000s, and dedicated data repositories becoming widely available at

Reviewer: 2

Comments to the Author
The paper is significantly improved and is close to final. In particular, it is better organized and more consistent. However, I have a few specific suggestions.

I still think the opening sentence about elections is weak. I doubt that public opinions on science or expertise decided any recent election. In the case of BREXIT, apparently the major issue was immigration followed by nationalism. Discounting of expertise was a way of diminishing concerns about the economic consequences. Also, as an opening statement, it sets an unfortunate political tone that implies bias.

Removed the election reference

p. 4. I do not agree with the following response to my comment:
123-127 I do not consider this to be a deception. Journals do not want an account of a research program; most readers do not want it either and papers do not claim to present it. The reported methods are the methods that generated the reported results. Presenting all of the mistakes and failed methods would just be confusing. This is from a quote and presumably would be read as that one person’s viewpoint. It is not just “one person’s

viewpoint." You adopted it as your viewpoint when you wrote that Goldstein "put it well." I do not believe that it is desirable to describe an appropriate and generally accepted practice as deception. Nobody is deceived.

> Changed "*Goldstein (1995) put it well*" to "*as Goldstein (1995) put it*" plus a remark that if concise writing neglects to mention statistical fishing or excludes data that don't fit, that could bias the literature.

p. 18, items 1, 2. Field based environmental effects studies are seldom experiments. Most are observational. I suggest studies (as in the previous draft) or investigations. Good point – made the change

p. 18, items 5. You should not hope for a positive relationship, particularly in this paper! You should hope to accurately estimate the actual relationship. (I missed this one last time.)

> Fair point – changed to detecting real relationships if present

Reviewer: 3

Comments to the Author
I think the paper will be publishable, but there are still some suggestions I'd strongly recommend before publication:
(1) Most importantly, there's still ambiguity about what exactly the paper's take-home lesson is regarding how to promote scientific integrity. The last sentence of the abstract made it sound like their focus is on promoting rigor, relevant reproducible research, transparency, and education. But it wasn't clear throughout the rest of the paper if this precise list was indeed the take-home lesson. On p. 17, they listed objectivity as part of their list with rigor, relevance, reproducibility, and transparency, and they didn't mention education (although education was discussed later in the paper). At the end of the paper, on p. 32, they mention objectivity again in their list and don't mention transparency. And then at the very end of the paper there's a list of concrete suggestions. I liked the suggestions, but it wasn't totally clear how these related to all the stuff about rigor, relevance, reproducibility, and transparency. It sounded like maybe the suggestions were focused specifically on promoting transparency, but I think the suggestions could promote all the items on the list. I would frame the list as a set of concrete strategies that different individuals and institutions could take to promote the rigor, relevance, reproducibility, and transparency that the authors are calling for.

> The reviewer is correct that we certainly don't have a plan to solve every behaviour. We have a few concrete suggestions but much of "scientific integrity" are things to consider or watch out for but don't necessarily have prescriptive fixes. While we thought that objectivity ran throughout our other points on minimizing bias, we appreciate the criticism here that for parallel structure it could be good to have something specific to objectivity. So, we added a short section that mostly relies on links to other literature for a more detailed treatment

(2) Given how comprehensive the paper is, I think it's somewhat surprising that it doesn't mention the limitations of disclosure as a response to conflicts of interest. There's important empirical literature suggesting that people who disclose conflicts of interest may feel more comfortable being biased as a result, and those who receive the information about conflicts of interest often don't really know what to do with that information. I still think it makes sense to disclose conflicts of interest, but these cautionary points ought to be noted. (See Cain DM, Loewenstein G, Moore DA. The dirt on coming clean: perverse effects of disclosing conflicts of interest. J Legal Stud 2005;34:1–25.)

> That's a good point. Transparency is well and good but avoiding conflicts of interest is better.

(3) When the authors discuss cases of failing to disclose conflicts of interest on the top of p. 9, I would suggest mentioning the recent case where the journal Critical Reviews in Toxicology issued a correction because authors of articles in a supplement about glyphosate failed to acknowledge assistance provided by Monsanto. That whole case is reminiscent of the "publication planning" by industry that has been widely criticized in biomedical research. That sort of publication planning appears to be happening in toxicology research as well. (See https://www.bloomberg.com/news/articles/2018-09-27/monsanto-s-role-in-roundup-safety-study-is-corrected-by-journal)

> Agreed, that is a relevant recent development in the field. Added citation to, but no discussion of this incident

(4) As I noted earlier, the section beginning on p. 17 about promoting integrity in ecotoxicology is confusing because it gives a list of key concepts and then proceeds to talk about all of them except objectivity. I mentioned this in my previous review. They need to drop the mention of objectivity there or include a section about it like they do with the other concepts. Another organizational weakness of that section is that they have sub-sections at the end on environmental chemistry and critical reviews. I'd be inclined to put those in a separate section because they don't fit with the series of key concepts (rigor, reproducibility, relevance, transparency) that otherwise make up that section.

> As noted above, we added a short section specific to objectivity. Objectivity should run throughout, and the discussion of confirmation bias is effectively on objectivity.

(5) I think it's a mistake to say that education is the only way forward (see p. 31). It makes it sound like the best way to promote integrity is to tell people how to do the right thing, but a lot of psychological research suggests that's not effective. You have to change people's incentives, not just tell them what to do. This is why the NSF changed their science ethics program from a focus on education to a focus on "Creating Cultures for Ethical STEM." The authors can keep that section on education as one piece of the solution, but they shouldn't make it sound like the main solution.

> We agree we shouldn't imply education is the only way forward, but it is a way to foster a culture of science integrity. Revisions include making the title less specific and adding mentions of the importance of culture and enforcement.

(6) Instead of making education sound like the main solution, I think the authors should focus more on their concluding list of strategies. I would frame them as a list of concrete ways to promote the rigor, relevance, reproducibility, and transparency that they called for earlier.

> Same as above

(7) When they talk about reproducibility, it seems like a shame not to mention the paper by Munafo et al. (2017) that provides concrete recommendations for promoting reproducibility: Munafò, M.R., Nosek, B.A., Bishop, D.V., Button, K.S., Chambers, C.D., du Sert, N.P., Simonsohn, U., Wagenmakers, E.J., Ware, J.J. and Ioannidis, J.P., 2017. A manifesto for reproducible science. Nature Human Behaviour, 1(1), p.0021.

> We agree and appreciate having it pointed out to us

Reviewer: 4

[Reviewer 4 carried forward some previous author responses from the first round of reviews. These are indicated with red text.]

Review of IEAM-2018-029-CR-R1

General: Having read the author responses to reviewer comments for all four reviewers I think they have done a fairly reasonable balanced job overall. The manuscript is now much more structured and fluid. Moreover, I appreciate acknowledgment by the authors regarding the need to address educational consideration and focus, with particular mention in the abstract. Consequently I believe the manuscript is acceptable and I have only provided minor comments below, but these are not considered to be mandatory for publication, rather for consideration.

Specific (note these page/line number pertain to the previous version and comments):

Page 14, Page 301-304: Do the author's feel that such disclosures should be universal regardless of funding source and affiliation? It has been my experience that for industry funded projects or collaborations the requisite degree of disclosure (conflicts of interest etc.) are quite onerous relative to other sectors. Would it be reasonable to suggest a more generally prescriptive approach that is consistent across sectors and reflective of equality?

Perhaps we are naïve, but our suggestion is that for most people and situations regardless of where the funding came from, a more prescriptive approach is doubtfully helpful. Instead we argue "simple, unambiguous statements of the funding sources that directly or indirectly allowed the work to be completed should generally be sufficient."

As a personal anecdote, the expectation for industry sponsored research is to not only indicate as much (which makes perfect sense for any funding source) but also to indicate if an author is employed by, or otherwise affiliated with, the sponsor. Additionally, I have recently been requested to detail roles and responsibilities of individual authors in the acknowledgments. I do not necessarily take exception to these requirements in principle, but rather the lack of tripartite equality per se. If the goal is transparency, to be an egalitarian society, then consistency is paramount.

> Hopefully our viewpoint comes through clearly that such disclosures are appropriate for all, although in practice some will be tedious and redundant. No change made.

Page 18, Lines 383-384: Why "particularly when funded by sponsors with financial interests in the findings"? In what situation does a 'funder' not have interest in the findings? This seems to be venturing into the realm of the subjective here

In my (cm) experience with government-funded research (other than NRDA), the funding entity does not stand to directly gain or lose financially from the study outcomes. They have just wanted a reliable answer to their questions, not a particular answer. No changes made.

Honestly that is a bit of a softball response, and I still take exception to the word "particularly", although this may be an artifact of the cited source. Also, emphasizing 'financial incentive' as an agent of bias vulnerability suggests disproportionate influence relative to reputational bias, personal bias, political bias etc. Moreover, although in cases where a funding entity may be considered bias neutral, that neutrality does not necessarily transfer to the recipient of the funding. Anyway, just seems a little loaded to me…

> (In my copy, this phrase was at line 411). That's a fair criticism, especially since the phrase was specific to risk assessment, not studies in general. Removed the sentence fragment about financial bias, since as noted, alternatively agency risk assessors can "be on a mission" or subject to other kinds of bias. It reads better to delete rather than balance with other types of potential bias.

Page 18, Lines 384-389: These are artifacts of scale, resolution, technical precision as well as matters of policy. Risk assessments are often screening level by design in order to minimize type-II errors, but seldom move to higher tiers of refinement because there is no established mechanism for interpretation and incorporation.

That's a perceptive point, but to actually work it into the text would require expansion and additional literature searching and citing. The manuscript had gained weight as it is through responses to review suggestions, and in consideration, we think it best to leave this paragraph as is.

Fair enough, though it is often poorly understood just how conservative screening-level risk assessments are (and they are with good reason), and that in fact the EPA subscribes to a tiered risk assessment framework to potentially address the type I/type II error rate artifact (hypothetically).

> We agree this is an important point, but after some attempts to work it into the paragraph with enough words to make sense, but not too many words, it seemed like it was getting off point and we didn't incorporate those points. Also, we were unable to quickly find a good reference on the point that risk assessments "*seldom move to higher tiers of refinement because there is no established mechanism for interpretation and incorporation*." The risk assessment terminology hadn't been used earlier in the manuscript and would need explanations such as screening levels, and higher tiers. The terminology and definitions vary somewhat between European and American regulatory approaches. We have so much already, getting more into the particulars of risk assessment practices seemed a little too much.
>
> By the way, the line numbers between Reviewer 4 and our proof go out of sync here, for in the author proof, the text mentioning risk assessment uncertainties that corresponds with the comment was at lines 405 to 418 on p. 12 of our version (that is, the printed number 12 on the bottom of the page, not the pdf generated numbers).

Page 31, Lines 650-667: This may be somewhat true for field experiments but less-so for lab-based experiments, especially those conducted under GLP guidelines with extensive validation.

This is encouraging to hear, but we're not sure whether information supporting this observation have made it into the open literature. The Owen (2010) rainbow trout study referenced was conducted under GLP, so we think these points are at least sometimes relevant.

Perhaps a good point of reference would be the requirement within a number of FIFRA ecotox protocols to have a positive control test conducted on a defined schedule to ensure laboratory performance. In such studies the performance of the positive control needs to be consistent with historical data and within the bounds of performance acceptability criteria for the lab to be considered competent to perform the studies. For example, in OPPTS 850.3020 (Honey Bee Acute Contact Toxicity): "A concurrent positive control with a substance of known toxicity is not required. However, a quarterly or semiannual test with a laboratory standard (reference toxicant) is recommended as a means of detecting possible interlaboratory or temporal variation. A laboratory standard is also recommended when there is any significant change in source of bees"

> In the "Reproducibility" section, we added a sentence on the benefit of positive controls.

Page 36, Lines 746-756: Are the authors referring to the actual "raw" data here or summarized data? Best to specify.

Added several sentences explaining that most often, when researchers ask for "raw data" they probably really want a detailed data summary.

Not sure that is entirely accurate, my experience has always been to request, or have requested, the entire raw data compilation.

> In context, we are probably thinking of the same way as Reviewer 4, but the sentence wasn't clearly written. Replaced it with an example of what constitutes "raw" data in a typical toxicity test. Otherwise, we think the wording in the paragraph is OK. "Raw data" may vary depending on the question and context. For example, in a toxicity test, counts or measurements from each replicate and all discrete, measured chemical values would likely be considered "raw data." We (mostly) take the concentrations that the chemists tell us, and don't really want the raw

instrument records of GC-MS time elution graphs or the raw spectral plots from an ICP-MS. If the study were into comparative performance of different chemical methods, "raw data" would mean something else.

Page 37, Line 780: "manipulation" sounds a little shady, do you mean evaluation?
I was thinking more of transformations and standardizations that need to be done before statistical analyses, how to handle missing data, non-detects. No change was made.
Still sounds shady (and not like the Real Slim Shady), perhaps data transformation or handling or something…

Replaced data manipulation with "data reduction and data standardization." We do cite a wide variety of sources, but not The Real Slim Shady.

Page 37, Lines 780-787: This is where establishing criteria for evaluation *a priori* is critical (see Van Der Kraak et al. 2014 for example (Crit Rev Toxicol. 2014 Dec;44 Suppl 5:1-66. doi: 10.3109/10408444.2014.967836)
Already added reference to Moermond, which seems to cover similar ground.
Similar, but not the same in my opinion. I could mistakenly sense some reluctance to cite Van Der Kraak et al., 2014, however, regardless of the subject matter content under evaluation, the principle of the review (i.e. quantitative weight of evidence with a priori established criteria and transparency of documentation) is quite important.

We added the Van Der Kraak citation as the reviewer has persistently suggested. We agree that Van der Kraak did a nice job in their Table 1 of laying out tangible criteria for critically rating literature. Reviewer 4 is perceptive – we were reluctant to cite another salvo between the atrazine combatants, as citation is a form of signaling (as Reviewer 2 noted, taking exception to another one of our citations). Also, we wondered just how "a priori" Van der Kraak's decision criteria truly were since most of the ground evaluated had been well trod by some of those authors before that paper. Nevertheless, we agree that Van der Kraak's Table 1 gives a nice example of a rating scheme that can be applied objectively and transparently in critical reviews.

Editorial Office

Comments to the Author
- Footnotes are not allowed in the text. Please incorporate the footnotes into the main body. OK

- The citation on page 2 is for Huff and Geist (1954) but the reference has the names Huff and Geis. (no "t"). Please correct. Done

- Figures 1 & 2 are low resolution, 72 dots per inch (dpi). Figures should have a minimum resolution of 300 dpi to ensure high quality reproduction in the typeset article.

I reloaded the files with the highest resolution that was provided to me.

****

16 September 2018

Responses to reviews of the original submission of "*Scientific Integrity Issues in Environmental Toxicology: improving research reproducibility, credibility, and transparency*"

Reviewer 1

Comments to the Author
This is a well-written manuscript that deals with an important issue. I have a few comments:
1. Both "scientist" and "researcher" is used in the manuscript. Are these used as synonyms or are you referring to different groups? Sometimes I got the impression that you were talking about academic researchers and sometimes a broader group, i.e. those publishing in peer-reviewed journals. Please clarify.

> We think the context is reasonably clear that we mean "scientists," which is broader than academic researchers, generally referring to anyone who is engaged in scientific pursuits and publishes in peer-reviewed journals.

2. Provide references to the biomedical research mentioned in row 75.

> Most of the 32 cited examples from lines 50-65 relate to biomedical research, as do the 2 citations following in the same paragraph at lines 81-84. Thus, the point seems sufficiently referenced without further cluttering the prose.

3. The whole section on row 392-404 lacks references. This is especially problematic since it includes harsh allegations. Provide references or rephrase.

> Added citations supporting statement that site remediation can be in the hundreds of millions or more (Gustavson 2007, ES&T "Superfund and mining megasites"; NYT – GE spent $1.6B USD dredging PCBs in the Hudson River). The suggestions that such tremendous financial consequences might actually have some effect on how scientific data are scrutinized are hardly "harsh allegations."

4. The section on row 429-445 lacks a critical discussion of the implications (for credibility and scientific integrity) of having industry-funded research (industry control 40% of the research in the US) and education.

> We think we do touch on the dual potential benefits and pitfalls of industry funded research, and have added another sentence on the need and benefit of full transparency

5. The section on row 455-481 should also include a couple of sentences on the numerous cases of misconduct from the chemical industry, to clarify why it could be relevant to consider the funder of the study. Examples of misconduct include perfluorinated chemicals, trichloroethylene, formaldehyde, styrene, dibromochloropropane, 1,3-butadiene, chromium (VI), benzene, vinyl clorid, lead, pharmaceuticals, asbestos, beryllium, and tobacco.

> We've worked in more examples, limited to a few in the open literature which do not appear to be actively disputed. Some of the examples suggested by the reviewer are contested and this review/commentary is not the place for original research of science misconduct allegations.

6. Row 520: Add that the cost of attending meetings and conferences is a major concern for NGOs.

Good point; we added a statement to the text.

7. The section on reproducibility on row 608. Add the reference by Moermond et al. (CRED method) since it includes a list of reporting recommendations to authors of ecotoxicity studies that has the possibility to enhance reproducibility of studies.

Good suggestion, added

8. The "implicit claim by researchers" mentioned on row 710 is an ill-disposed interpretation that should be left out of this manuscript.

It is intended to be provocative to readers that presenting a study without supporting data is effectively a "trust me" claim. Doubtfully all will agree

9. Provide a reference on the attitude change mentioned on row 713.

Changed to more accurately state the historic limited details in scientific papers was due publication constraints rather than changes in attitudes

10. Row 823: Systematic review methodology is now being used also for chemical assessments. Environmental International has a new editor, Paul Whaley, that has addressed this in several publications that should be cited, e.g.

A primer on systematic reviews in toxicology Hoffmann, S., de Vries, R.B.M., Stephens, M.L., Beck, N.B., Dirven, H.A., Fowle, J.R., Goodman, J.E., Hartung, T., Kimber, I., Lalu, M.M., Thayer, K.A., Whaley, P., Wikoff, D., Tsaioun, K. 07/2017 In: Archives of Toxicology. 91, 7, p. 2551-2575. 25 p.
Raising the standard of systematic reviews published in Environment International Whaley, P., Letcher, R.J., Covaci, A., Alcock, R. 12/2016 In: Environment International. 97, 3 p.
Assuring high-quality evidence reviews for chemical risk assessment: five lessons from guest editing the first environmental health journal special issue dedicated to systematic review Whaley, P., Halsall, C.J. 07/2016 In: Environment International. 92-93, 3 p.
Implementing systematic review techniques in chemical risk assessment: challenges, opportunities and recommendations. Whaley et al.

We appreciate having this work pointed out. Added citation to the "Implementing" paper.

****

Reviewer: 2

General Comments:
The manuscript is an excellent review of the issue of scientific integrity. However, it is rather long and loosely organized. What is the purpose other than literature review? The abstract promises a framework, but the word does not even appear in the rest of the paper. The last sentence of the introduction seems to set a goal of relating remedies for nuanced issues in scientific integrity from other sciences to SETAC's science. That is done but not clearly or consistently. (The sentence itself is unclear.) Eleven points related to transparency are presented in the last section. It is not clear whether these are the promised remedies from other sciences or if the remedies are limited to transparency. If the conclusions all concern transparency, why discuss bias, rigor, relevance, etc.? Is transparency the solution to all scientific integrity issues? In sum, the paper reads like a committee effort: lots of good

information that feels somewhat thrown together. One of the authors should do a major rewrite to tighten up the organizational logic.
There are suggestions and recommendations scatter through the text. They do not appear consistently at the end of sections and good ideas (e.g., videos of the methods) are not distinguished from recommended or required practices.

> We've tried to clean up these points to keep the logic logical and the flow better. Specifics follow.

Specific Comments:
Title. Why just toxicology and not Chemistry? The text frequently refers to SETAC. Is integrity not an issue in chemistry or less of an issue?

> That's a fair criticism. We agree that chemistry has its own issues, and added a section on readily available pitfalls in environmental chemistry. We also added "chemistry" to the title of the article.

25 Where is the framework? This is the only place that the term is used.

> Framework was probably not the best term. Removed.

29 How have recent elections shown that large segments of society distrust science? In the U.S., Trump said that global warming is a hoax but I have not seen any evidence that his election hinged on that position. I have not heard of any election in which scientific integrity was a major issue.

> The sentence doesn't imply anything about specific issues being in play in the US elections, but that large fractions of electorates do not trust in "experts." Added citation to Nichols 2017, (Death of Expertise) who specifically discusses the discounting of "experts" in UK's Brexit vote

41-43 This may be intended as a joke. It misrepresents Socrates. His dialogues are exercises in essentialist argument and do not rely on evidence. Perhaps you are thinking of Aristotle who was an empiricist and proto-scientist.

> Fine to remove, as I am not a classical scholar. The point was that he was put to death for his views/teachings which contradicted mainstream thought

45 Archaeoraptor was never a scientific discovery. It was synthesized by a farmer/collector in China and bought by a dinosaur enthusiast in the U.S. Its description was rejected by Nature and Science. It was published by National Geographic without checking the science. Creationists described it as a failure of science, but no scientist who looked at it or at images of it believed it was legitimate. The failure was that nobody warned National Geographic until they were in press and they did not check. Hence, it was never accepted by the scientific community nor by a peer-reviewed scientific journal. This is a picky comment but I am sensitive to this case because it is used regularly by creationists.

> Yes, but the point is none of those examples were valid scientific discoveries, but at some points were presented as such. Amended to "purported discoveries."

50-64 This is a list of bad practices (called concerns) in science. The comment about beer at the end does not belong in that series.

> Fine, deleted. We thought it ironic that studies have been published attributing beer as both causing and preventing cancer questioning the credibility of both. It was an attempt to slip in a bit of humor.

73 maintain should be maintaining
Correction made.

106 You do not need to cite Kolok for "the dose makes the poison." It is a common paraphrase of Paracelsus.

Concur; citation deleted.

115 Lackey's contention that scientists should avoid normative concepts is naive. Environmental laws and regulations contain normative requirements. Scientists are required to operationalize concepts in environmental laws such as biotic integrity, impairment, and toxic effects. Business managers or government attorneys are not going to do that themselves.

Revised to remove the citation here to Lakey's advocacy of nonadvocacy. This is explored more in the section on advocacy.

123-127 I do not consider this to be a deception. Journals do not want an account of a research program, most readers do not want it either and papers do not claim to present it. The methods are the methods that generated the reported results. Presenting all of the mistakes and failed methods would just be confusing.

This is from a quote and presumably would be read as that one person's viewpoint

173 Misplaced period.
Corrected.

244 The primary school norms are cute, but "practice makes perfect" is not reproducibility. I am not sure what is. Perhaps this is an exception to the hypothesis that we learned the needed norms as children. It is listed as a "profession-specific provision" in the first sentence. Perhaps those provisions should not be among the primary school norms.

We agree that alluding to primary school norms and folk idioms only goes so far, but the purpose is to ground our discussion of scientific integrity is universal behaviors, rather than an emphasis on procedural. This and other attempts to bring humor and whimsy into the discussion are done deliberately with the hope that it will increase the likelihood of actually being read and thought about. Some of these attempts will undoubtedly bomb with some readers. Reworded, but largely retained.

328-330 Awkward sentence.
Agree. Rewritten.

349, 352 The "doubtfully" construction is awkward. I had to read the sentence in 349 twice.

Agree. Reworded

409-412 Edit this very long and not entirely grammatical sentence.
Agree. Shortened.

407-421 This seems to imply that scientists employed or contracted by natural resource trustee agencies stand to financially benefit from NRDA. Who has a financial incentive in the U.S.? Not the scientists or lawyers (who are just public service grunts) or regulators (regulators are not trustees so they have no standing). The situation is quite different from a product liability tort where the lawyers get a fraction of the take and expert witnesses are paid to present testimony to support one side. The leadership of the trustee agencies have an incentive to obtain funding to remediate the damages, but it does not line their pockets. NRDA is a scientific integrity issue to the extent that the scientists at trustee agencies are likely to be personally invested in the protection and restoration of natural resources.

However, that is a potential source of bias whether or not financial damages are involved. The situation may be like product liability torts in other countries, but I do not know.

This is a fair point and we revised it to split out NRDA from toxic torts.

482-484 I agree that scientists should not be dismissed based on their employment or funding. However, it is appropriate to look at the body of evidence in a case to determine whether there is a bias associated with the source of funding. If industry funded studies and foundation funded studies consistently come down on one side or the other (e.g., Atrazine), that should be noted and if possible it should be accounted for.

We think we've captured this point in the present text, following this sentence

522-523 Also, environmental advocacy groups conduct or fund little science. They mainly review other people's science, so they do not have much to present at a SETAC conference. They might also feel uncomfortable with the large industry presence. It would be interesting to know whether they go the less industry-dominated societies like the Ecological Society of America or Society for Freshwater Science.

That could be. I (cm) participated in SFS for several years and interacted with a few NGO representatives, who were advocates for insect or species conservation. Cost could be a major factor (noted by Reviewer 1). We haven't researched the point and don't want to go too far into this guessing.

531-532 Fix grammar

Fixed

581 Melvin et al (2009) is not a field study.

Good catch. Reworded. It with was with effluents and intended to inform environmental effects monitoring.

579-607 Parallel structure would make the 8 principles easier to read and understand. For example, only a few are complete sentences and two (4 and 7) are imperative case, but not others.

Rephrased as factors leading to rigor, and reworded to avoid the mix of cases

655-656 I do not have a lot of sympathy for this excuse. If it is impossible to produce the same effect twice in laboratory studies, can it really be called science? But you address that in the next paragraph.

We kept the wording as is. In toxicity testing, even standard test organisms can show quite different responses to very similar test organisms. For instance, USEPA's water quality criteria derivation guidance consider that if normalized toxicity test results of the same species differ by more than a factor of 10, it may be appropriate to reject some values. A factor of 10 is a large difference.

742-743 What if the author dies or is incapacitated? All data should be posted, in my opinion. See 760-762.

We tend to share that view, but are trying to build the case. Added a sentence pointing out that eventually all scientists die and that important data sets need not die with them.

868-883 The laws use normative terms. If an environmental toxicologist says only that something is changed without stating whether the change is adverse, he is not doing his job (note, he and his are indefinite masculine pronouns, not signs of sexism).

We think we've captured this point, and but added the point that laws and international agreements are inherently normative

935 practice (singular)
Change made.
943 Which is it: avoid or not tolerate? They are very different recommendations.

True. We mean not tolerate

944-945 This is unclear.

Deleted this sentence which was the simplest resolution

960 Change to: Workshops and resulting publications.
Done
****

Reviewer: 3

I think this paper provides a nice overview of a lot of material related to scientific integrity in toxicology.
However, I have some concerns and suggestions.
One fundamental concern is that it's not entirely clear to me how all the different parts of the paper fit together. Parts of the paper seem like a review of the literature on scientific integrity related to environmental toxicology, but other parts of the paper seem like more of a commentary on how to promote integrity. Maybe it's OK to have multiple things going on, but I think the authors could do more to clarify how it all fits together and what the new points are that they are trying to defend.
Another concern is that the parts of the paper where they seem to be defending new points could be strengthened. In the abstract and introduction, the authors make it seem like their biggest new contribution is to conceptualize scientific integrity as an extension of personal integrity. Two thoughts:
(1) I think this is a questionable move, because there's so much recognition nowadays that integrity is not just a personal, individual issue but also an institutional, community-level issue. (After all, the authors themselves note all the valuable things that SETAC can do to help promote integrity.) Thus, I think the authors should, at minimum, talk about integrity as a joint individual- and community-level enterprise.

True. Deleted the personal integrity wording from the abstract.

(2) I thought the authors' discussion of the concepts of relevance, rigor, reproducibility, objectivity, and transparency was one of the most interesting parts of the paper, so I think they should consider playing up those concepts in the abstract and introduction as a more central element of their positive account of scientific integrity in toxicology.

Concur; abstract has been significantly reworded.

Speaking of relevance, rigor, reproducibility, objectivity, and transparency, somehow I missed the discussion of objectivity. Did they skip over it accidentally and go directly from reproducibility to transparency? I'd also encourage the authors to consider their list of recommendations at the very end of the paper. It seemed like kind of a mish-mash of different ideas from throughout the paper. Is there any way to systematize or motivate that particular list of ideas a little better?

We think that objectivity is a thread throughout our arguments. Listing specific recommendations is a good idea.

Some smaller points, largely involving literature that might be good to cite or mention:
(1) Sheldon Krimsky's book Science in the Private Interest would be a good piece to cite when talking about concerns that industry-academic partnerships can be problematic (2) Resnik and Elliott have a piece on "Taking Financial Relationships into Account When Assessing Research" that discusses how research shouldn't automatically be dismissed because of the funding source

Concur Krimsky is relevant here. Resnick and Elliott's arguments were already made.

(3) Linda Birnbaum and colleagues have an important paper discussing an important issue that wasn't discussed much in the "relevance" section of the article. Specifically, they point out that a lot of the standardized toxicology studies done for regulatory purposes may not actually provide very good indications of risk in humans with diverse genetics being exposed to chemical mixtures. See Birnbaum et al, "Informing 21st Century Risk Assessments with 21st Century Science"

It is an important paper, but they're taking on huge and somewhat different questions – better rigor in animal models for informing human health risks, better use of epidemiology, better use of molecular tools to predict clinical-level effects. In our context, which emphasizes ecotoxicology, it might confuse readers. We didn't add it.

(4) An interesting opinion piece promoting industry collaborations with academics as beneficial is: Edwards, "Reproducibility: Team up with Industry"

Yes, we appreciate you pointing this article out. While not all aspects of this big-money human health genomics work would apply to the typically much smaller ecotox experimental world, some are probably universal such as mandating data sharing and defining quality criteria. We added mention of some of the relevant features.

(5) In the authors' discussion of transparency, I think they need to provide more emphasis on the problematic fact that most toxicology studies done for regulatory purposes are not published or made public (although there's an important initiative by Bayer to make more of their data available; see https://cropscience-transparency.bayer.com/)

Both of these were mentioned. The Bayer transparency note was promoted from a footnote to the main text.

(6) For the discussion on advocacy, it would be wise for the authors to cite Roger Pielke's book The Honest Broker, which discusses strategies for scientists to avoid falling into inappropriate advocacy

Reasonable point. It's a high-profile book and some have reviewed it favorably. We have added a reference to this text in the discussion about "stealth advocacy"

****

Reviewer: 4

General: I enjoyed reading this article, I found the content and expressed perspectives to be balanced, thoughtful and insightful. Moreover, the multi-partite authorship composition, reflecting the many sectors that constitute the society, conveyed a sense of consensus view. I have provided a number of specific suggestions/comments/questions below, however, the most significant detraction from the manuscript is the lack of consideration/focus on education. SETAC (in my opinion) as an entity has evolved into a more student oriented venue, relatively speaking, and given the greater focus on students (participation and consideration) I was a little surprised at the lack of specific focus on this demographic. I fully support the content and perspectives conveyed by the authors, but I don't think we can achieve broader success as a Society unless there is a concerted effort to instill this ethos in students, which requires a strategy for education. The tenants of science and scientific method need consistent reinforcement beyond just the 'SETAC crowd' and this is best achieved in the classroom in addition to conferences and workshops. I would encourage the authors to consider this point and reflect it as appropriate in the article. Other than that my comments are pretty minor and feel the article is acceptable with revision.

> Upon reflection, we concur we gave the education strategy short shrift. We added a brief section about strategy for education of early-career scientist and reinforcement for latter career scientists.

Specific (note author page and line numbers used, not the automated ones):
Page 1, Line 21: …such as poor reliability/reproducibility and bias.

> OK, changed

Page 2, Line 29: Whatever do you mean?

> Added another citation on that point (book, "Death of Expertise").

Page 2, 3: Although it is eluded to it would be of value to illustrate how sensationalism has been incentivized. Personally I think there is immense pressure put on tenure track professors and tenure package reviews are often gauged by H-index, impact factor etc., and unfortunately this (in my opinion) has led to an incentivizing of sensational results. It is a peculiarity that many sensational (and I don't presume to suggest sensational necessarily equals questionable) research findings are pre-empted by news briefs and press releases before the article is even published (i.e. publicly evaluated). This is not to suggest undue institutional pressure, but given the metrics for evaluation, the more sensational the findings, the greater the media coverage, the higher the proportion of citations, H-index and so on and so forth. Personally I think this is a dubious practice, but it seems to be becoming more common. There was a recent article in the journal Ecotoxicology by Hanson et al entitled "Evidence of citation bias in the pesticide ecotoxicology literature" that was quite interesting and may be of value for the author's to consider as a perspective.

> In early drafts, we went into some detail on publication pressures and practices, but the manuscript was becoming unwieldy with length, so we pulled out most of the publication practice material. We intend to address publication matters more in a separate manuscript. We appreciate pointing out the 2018 article which we hadn't yet seen, and we agree it likely is highly relevant.

Page 9, Line 190: What are the peer-review procedures in place to identify and address FFP? Does this vary by journal on a case by case basis? I can only feature the Andrew Wakefield saga where a full retraction by The Lancet was issued for his falsified 'seminal' research linking vaccines and autism…a perception that unfortunately persists to this day at the expense of public health and safety.

> Unfortunately, peer-review will seldom catch outright fabrication or fraud. This will probably only be found out through post-publication scrutiny.

Page 10, Lines 216-219: What about the Freedom of Information Act (FOIA)? This is an oft-used mechanism to obtain transparency within the Federal Government.

> True, when used to pry raw data or code that would help with analyses and reproducibility and when aimed at scientists. It also is a venue for harassment through vexatious demands for notes, emails, etc. We mention the double-edged sword of transparency in the "Weaponizing scientific integrity section."

Page 11, 233-234: cooking data…I think the degree or imposition of this issue is dependent on whether or not the criteria, assumptions, and uncertainties were disclosed and if this was done so *a priori*

> At those lines, the cleaning or cooking of data is posed as rhetorical question which we come back to in the bias section.

Page 14, Page 301-304: Do the author's feel that such disclosures should be universal regardless of funding source and affiliation? It has been my experience that for industry funded projects or collaborations the requisite degree of disclosure (conflicts of interest etc.) are quite onerous relative to other sectors. Would it be reasonable to suggest a more generally prescriptive approach that is consistent across sectors and reflective of equality?

> Perhaps we are naïve, but our suggestion is that for most people and situations regardless of where the funding came from, a more prescriptive approach is doubtfully helpful. Instead we argue "simple, unambiguous statements of the funding sources that directly or indirectly allowed the work to be completed should generally be sufficient."

Page 16, Line 337: …can influence scientists to modify their perception and thinking

> Concur; change made.

Page 16, Lines 345-346: Again, as described earlier for other sectors, there tends to be a bias towards studies demonstrating significant effects vs those that don't. Obviously if you are a drug maker you're much less likely to publish a non-significant therapeutic effect, which unfortunately is lost information that could have informed other approaches. As with the sensational, perhaps the scientific community should aspire to the more fundamentally neutral, valuing the non-significant equally to the significant.

> We expanded on this a little, pointing out that study results that favor the interests of the funder don't necessarily imply experimental bias, they might just have greater expertise with the chemical.

Page 17, Lines 360-316: This sentence reads oddly/awkwardly

> Checked and edited.

Page 17, Lines 364-365: I can appreciate the obviousness of this example, but the Aviv article was anything but balanced, and conveying citations as if weighted equally in terms of content, rigor, reproducibility etc. is a little disingenuous.

> The goal is to lead readers to sources on a conflict in the science, not to attempt to referee which is more persuasive. However, we concur that the Aviv article could be replaced from more scholarly literature. Removed it, and added peer reviewed sources from both camps: Rohr & McCoy (2010), Bero et al 2016, Hanson et al (2018)

Page 18, Lines 383-384: Why "particularly when funded by sponsors with financial interests in the findings"? In what situation does a 'funder' not have interest in the findings? This seems to be venturing into the realm of the subjective here

> In my (cm) experience with government-funded research (other than NRDA), the funding entity does not stand to directly gain or lose financially from the study outcomes. They just want a reliable answer to their questions, not a particular answer. No changes made.

Page 18, Lines 384-389: These are artifacts of scale, resolution, technical precision as well as matters of policy. Risk assessments are often screening level by design in order to minimize type-II errors, but seldom move to higher tiers of refinement because there is no established mechanism for interpretation and incorporation.

> That's a perceptive point, but to actually work it into the text would require expansion and additional literature searching and citing. The manuscript had gained weight as it is through responses to review suggestions, and in consideration, we think it best to leave this paragraph as is.

Page 19, Lines 405-421: I feel like the concept of litigation has not been adequately addressed thus far (it is only briefly touched on here). Action by litigation is often touted as an effective tool to bring about scientific, regulatory, or policy action, but it is categorically abused, primarily in the U.S. There are cases where litigation is absolutely necessary, particularly in ensuring human and environmental health, but when taken advantage of by thinly veiled interest groups to support litigious enterprises that's problematic and detracts from scientific objectivity and integrity.

> That's a reasonable point. Added a mention of the hired guns and biased research associated with toxic torts in the subsection "*Some particularly challenging situations in ecotoxicology";* also touched on in the "Weaponizing transparency" section

Page 20, Lines 423-428: As a point of reference it may be worth mentioning the Gold King mine spill here as well.

> I had looked into that, but it didn't really fit the point. The references linked to on Mount Polley, in contrast, have quite a "bad science" drama to them. Litigation from Imperial Metals asserted faulty design, bad science and engineering by their consultant Knight Piesold, who designed the tailings facility. Knight Piesold in turn argued that in the years since their design and initial construction, Imperial Metals raised the height of the dam without their consultation, which cut into the safety margins of the dam construction review.

Page 22, Lines 446-466: I agree, but…what about other collaborating entities such as NGOs, do the same issues/considerations apply?

> In concept, they do although in practices NGOs usually don't have funding to support research. The Van Kirk episode (lines 470-480 in the original submission) was from an academic-NGO collaboration. It didn't go well.

Page 23, Line 490-493: This seems like an ideal place to discuss 'research contracts', which could loosely be analogous to a marriage certificate of sorts. Personally, I have been involved in establishing numerous research contracts with a variety of academic institutions and all have advocated for academic freedom. I think there is a perception that such contracts are highly restrictive, but again that has not been my experience.

> We haven't made any efforts to research this point and only have our own experiences which are mostly in line with the comment. Added a line to note that such contracts often establish expectations of academic freedom.

Page 25, Lines 527-529: I'm not sure I agree with tis statement, I don't think it is purely subjective. For example, if you consider the Sir Austin Bradford Hill criteria for causality, these are not purely subjective considerations to gauge scientific merit.

> The "subjective judgment" isn't essential to the sentence's meaning and was removed, to wit: *"While "scientific integrity" ~~is ultimately a subjective judgment that~~ cannot easily be reduced to review checklists, …*

Page 26, Line 552: I think 'sensational' is more accurate and appropriate here than surprising, or at least surprising and/or sensational.

> Agree. Added "sensational" to read "sensational or at least surprising."

Page 26, Lines 549-561: Again suggest to consider Hanson et al. 2018 here (https://doi.org/10.1007/s10646-018-1918-4)

> We aren't really taking on the publication bias issue here. We're a bit cautious invoking that study since the publication bias explored by Hansen et al. focused on papers by some of the same authors or collaborators, which muddies the publication bias angle. We plan to come back to publication bias and more questionable publication practices in a follow-up article.

Page 28, Line 579: First I think it is more fundamental to have a basic understanding of the need for experimentation…why are we asking the question(s) – should precede considerations for design Page 28, Lines 579-592: What about thoroughly vetting the available literature etc. to inform questions and study design? Understanding the issues to as informed questions is the first step in my opinion

> Good point. We added these thoughts as new #1 in the list.

Page 30: Again, designing a robust, reliable, reproducible experiment is dependent on thoroughly understanding the question(s) being posed. Inability to understand the question leads to vague and often misguided experimentation that undermines the scientific process. Asking a novel question does not preclude the application of scientific rigor; we need greater educational training of scientific method.

> Worked these thoughts into new #2.

Page 31, Lines 650-667: This may be somewhat true for field experiments but less-so for lab-based experiments, especially those conducted under GLP guidelines with extensive validation.

> This is encouraging to hear, but we're not sure whether information supporting this observation have made it into the open literature. The Owen (2010) rainbow trout study referenced was conducted under GLP, so we think these points are at least sometimes relevant.

Page 32, 672-682: Suggest to also include an example where a claim of anomalous reporting was in-fact falsified.

> We think that might be kicking a hornet's nest that's best left at a little distance. It's hard to say when incongruent results falsify one another, such as the atrazine and frogs controversy from line 365 (in the original version). Even in some situations which as a near-outsider seem resolved such as the bacteria-arsenic-phosphorus controversy, or inflated pharmaceutical concentrations (see the new material on Environmental Chemistry) despite the failure of other researchers to repeat anything close to the original findings, the original authors seem to still think they were right and the others have technique or other limitations.

Page 36, Lines 746-756: Are the authors referring to the actual “raw” data here or summarized data? Best to specify.

> Added several sentences explaining that most often, when researchers ask for “raw data” they probably really want a detailed data summary.

Page 36, Lines 758-759: Sentence reads oddly/awkwardly

> Reworded. Hopefully it’s clearer

Page 37, Line 780: “manipulation” sounds a little shady, do you mean evaluation?

> I was thinking more of transformations and standardizations that need to be done before statistical analyses, how to handle missing data, non-detects. No change was made.

Page 37, Lines 780-787: This is where establishing criteria for evaluation *a priori* is critical (see Van Der Kraak et al. 2014 for example (Crit Rev Toxicol. 2014 Dec;44 Suppl 5:1-66. doi: 10.3109/10408444.2014.967836)

> Already added reference to Moermond, which seems to cover similar ground.

Page 38, Lines 792-799: So use Google scholar (I agree)? Or is the point to use multiple search resources?

> Yes, one should use Google scholar to discover literature, but we can’t directly say so. US government authors such as the first author are specifically prohibited from endorsing specific products, firms, or services. Regardless of this specific prohibition, search engines and their performance may rapidly change, and more specific recommendations could become quickly dated.

Page 39, Line 817: The EPA Mid-Continent Ecotox Database is quite useful, but there are no strict review criteria for inclusion of data and users ought to be aware of this…just because you can find it in the EPA database doesn’t necessarily mean is credible.

> We wholeheartedly agree, but also think the existing text captures this point adequately that uncritical reliance on this or any secondary sources can introduce or perpetuate errors.

Page 39, Line 828: “science can never answer “should” questions, but can inform issue.

> Good point. Amended to include.

Page 42, Line 896: “Covert advocacy”…what about explicit advocacy? I agree that advocacy should be minimized in principal, but if it is obvious and explicit at least one acknowledges a position.

> We’re making the argument that explicit advocacy is a personal and situational decision.

Page 43, Lines 918-922: Lawsuits by who? I think the authors need to expand a little on the litigation issues as a whole.

> We added one more citation to a harassing lawsuit of a private researcher (Robbins) and earlier had added a short-section on toxic torts, cautioning that litigation-support science is questionable at best. But generally, it’s become a fairly lengthy article as is, and we only scratch the surface of important topics such as this. We think that readers will find the several references are useful leads to more reading.

Page 44: It is a little surprising to see that a focus on education a crucial and fundamental mechanism to ensure scientific method and enhance scientific integrity has been omitted here. I think this is a major oversight and the authors really need to focus on education as a concept and strategy to reinforce not only the tenants of science but the perspectives they have comprehensively conveyed in this manuscript. I think this manuscript is valuable and I largely agree with the authors, but it is not going to useful in practice unless there is a greater focus on students with consistent and compelling messaging.

> We think this is a good point and we were remiss in not being more direct. We added a section on education and enforcement.

Page 44, Point #2: Nice alliteration! ☺

> Rrrr! Thanks.

Page 44, Point #5: …that could contribute to, or be perceived as, bias. Also, as a general principal conflict should be avoided but that is not always possible and requires further consideration of tact, morals and ethics.

> Good point. Reworded as suggested to make it more explicit to avoid conflicts when possible.

Page 45: Another seeming omission is the lack or perspective on enforcement? Self-enforcement, institutional-enforcement, peer-review-enforcement…all of the above?

> Good point. All of the above, plus professional society to some extent. Worked it into the Education section.

Integrated Environmental
Assessment and Management

Scientific Integrity Issues in Environmental Toxicology and Chemistry: improving research reproducibility, credibility, and transparency

Journal:	*Integrated Environmental Assessment and Management*
Manuscript ID	IEAM-2018-029-CR.R1
Wiley - Manuscript type:	Critical Review
Date Submitted by the Author:	16-Sep-2018
Complete List of Authors:	Mebane, Christopher; National Oceanic and Atmospheric Administration, Sumpter , John; Brunel University Fairbrother, Anne; Exponent, EcoSciences Augspurger, Tom; USFWS Canfield, Timothy; US Environmental Protection Agency, National Risk Management Research Laboratory, Ground Water and Ecosystems Restoration Division Goodfellow, William; Exponent Inc Guiney, Pat; Guiney LeHuray, Anne; Chemical Management Associates LLC Maltby, Lorraine; University of Sheffield, Department of Animal and Plant Sciences Mayfield, David; Gradient, Toxicology Ortego, Lisa; Bayer CropScience LP McLaughlin, Mike; CSIRO/University of Adelaide, Schlekat, Tamar; Society of Environmental Toxicology and Chemistry Scroggins, Rick; Environment Canada, Verslycke, Tim; Gradient,
Keywords:	reproducibility, bias, transparency, research integrity, scientific integrity
Abstract:	High profile reports of detrimental scientific practices leading to retractions in the scientific literature contribute to lack of trust in scientific experts. While the bulk of these have been in the literature of other disciplines, environmental toxicology and chemistry are not free from problems. While we believe that egregious misconduct such as fraud, fabrication of data, or plagiarism is rare, scientific integrity is much broader than the absence of misconduct. We are more concerned with more commonly encountered and nuanced issues such as poor reliability and bias. We review a range of topics including conflicts of interests, competing interests, some particularly challenging situations, reproducibility, bias, and other attributes of ecotoxicological studies that enhance or detract from scientific credibility. Our vision of scientific integrity encourages a self-correcting culture promoting scientific rigor, relevant reproducible research, transparency in competing interests, methods and results, and education.

Scientific Integrity Issues in Environmental Toxicology and Chemistry: improving research reproducibility, credibility, and transparency

Christopher A. Mebane[1], John P. Sumpter[2], Anne Fairbrother[3], Thomas P. Augspurger[4], Timothy J. Canfield[5], William L. Goodfellow[6], Patrick D. Guiney[7], Anne LeHuray[8], Lorraine Maltby[9], David B. Mayfield[10], Michael J. McLaughlin[11], Lisa Ortego[12], Tamar Schlekat[13], Richard P. Scroggins[14], Tim A. Verslycke[15]

[1]U.S. Geological Survey, Boise, ID USA; [2]Brunel University London, UK; [3]Exponent Inc., Bellevue, WA, USA; [4]U.S. Fish and Wildlife Service, Raleigh, NC USA; [5]U.S. Environmental Protection Agency, Ada, OK USA; [6]Exponent Inc., Alexandria, VA USA; [7]University of Wisconsin, Madison, WI USA; [8]Chemical Management Associates LLC, Alexandria VA, USA; [9]University of Sheffield, Sheffield, UK; [10]Gradient, Seattle, WA, USA; [11]University of Adelaide, Adelaide, South Australia, Australia; [12]Bayer, Crop Science Division, Research Triangle Park, NC, USA; [13]SETAC, Pensacola, FL USA; [14] Environment and Climate Change Canada, Ottawa, ON, Canada; [15]Gradient, Cambridge, MA, USA

Abstract

High profile reports of detrimental scientific practices leading to retractions in the scientific literature contribute to lack of trust in scientific experts. While the bulk of these have been in the literature of other disciplines, environmental toxicology and chemistry are not free from problems. While we believe that egregious misconduct such as fraud, fabrication of data, or plagiarism is rare, scientific integrity is much broader than the absence of misconduct. We are more concerned with more commonly encountered and nuanced issues such as poor reliability and bias. We review a range of topics including conflicts of interests, competing interests, some particularly challenging situations, reproducibility, bias, and other attributes of ecotoxicological studies that enhance or detract from scientific credibility. Our vision of scientific integrity encourages a self-correcting culture promoting scientific rigor, relevant reproducible research, transparency in competing interests, methods and results, and education.

Introduction

Highly polarized recent elections in Europe and North America have shown that large segments of society are distrustful of scientific and other experts. Some have suggested that we are in a culture in which reality is defined by the observer and objective facts do not change peoples' minds, and those that conflict with one's beliefs are justifiably questionable (Campbell and Friesen 2015; Nichols 2017). Science and scientists have been central to these debates, and the boundaries of science, policy and politics may be indistinct. In a social climate skeptical of science, the easy availability of numerous reports of dubious scientific practices gives fodder to skeptics. Because environmental regulations on use of chemicals and waste management rely heavily on the disciplines of ecotoxicology and chemistry, the integrity of the science is of utmost importance. Here we discuss scientific integrity in the applied environmental sciences, with a focus on ecotoxicology and how the role and culture of the Society of Environmental Toxicology and Chemistry (SETAC) may influence such issues.

Science has long endured questionable science practices and a skeptical public. Galileo's criticisms of prevailing beliefs resulted in his issuing a public retraction of his seminal work. In contrast, purported science "discoveries" such as Piltdown Man, canals on Mars, cold fusion, Archaeoraptor, homeopathic water with memory, arsenic-based life, and many others have not stood the test of time (Gardner 1989; Schiermeier 2012). By 1954, Huff and Geist (1954) illustrated how the presentation of scientific data could be manipulated to become completely misleading yet accurate. Are things worse now? Recent articles in both the scientific literature and popular print and broadcast venues paint a bleak picture of the status of science. One does not have to search hard to find plenty of published concerns about the credibility of science. These include overstated and unreliable results (Harris and Sumpter 2015; Henderson and Thomson 2017; Ioannidis 2005), conflicts of interest (Boone et al. 2014; McGarity and Wagner 2008; Oreskes et al. 2015; Stokstad 2012; Tollefson 2015), profound bias (Atkinson and Macdonald 2010; Bes-Rastrollo et al. 2014; Suter and Cormier 2015a, b), suppression of results to protect financial interests (Wadman 1997; Wise 1997), deliberate misinformation campaigns as a public relations strategy for financial or ideological aims (Baba et al. 2005; Gleick and 252 co-authors 2010; McGarity and Wagner 2008; Oreskes and Conway 2011), political interference with or suppression of results from government scientists (Hutchings 1997; Ogden 2016; Stedeford 2007), self-promotion and sabotage of rivals in hypercompetitive settings (Edwards and Roy 2016; Martinson et al. 2005; Ross 2017), publication bias, peer review and authorship games (Callaway 2015; Fanelli 2012; Young et al. 2008), selective reporting of data or adjusting the questions to fit the data (Fraser et al. 2018), overhyped institutional press releases that are incommensurate with the actual science behind them (Cope and Allison 2009; Sumner et al. 2014), dodgy journals (Bohannon 2013), and dodgy conferences (Van Noorden 2014). [1].

Such published concerns reasonably raise doubts about science and scientists and could even lead some to conclude that the contemporary system of science is broken. In writing this commentary, we attempt to address some prominent science integrity concerns in the context of environmental toxicology and chemistry. In our view, there is ample room for improvement within our discipline, but the science is not broken, and some criticisms are overstated. In writing this commentary, we do not pretend to have solutions that will overturn insidious pressures on scientists and funders for impressive results, or hold some moral high ground making us immune from such pressures ourselves, or that all of our own works are above reproach. Our recommendations are pragmatic, not dogmatic. Our goal is to nudge practices and pressures on

---

[1] Throughout this commentary, citations are intended to be representative, without the "e.g." qualifier, which would otherwise be needed in nearly every instance.

scientists to advance the science, while maintaining and improving credibility through transparency, ongoing review, and self-correction.

Many of the prominent science integrity controversies have been in the high stakes biomedical discipline, and in response that discipline probably has done more self-evaluation and taken more steps toward best practices than most other disciplines. Results of self-reported, anonymous surveys of scientists, mostly in the biomedical fields, have not been reassuring. In a 2002 survey of early and mid-career scientists, 0.3% admitted to falsification of data, 6% to a failure to present conflicting evidence, and 16% to changing of study design, methodology or results in response to funder pressure (Martinson et al. 2005). A subsequent meta-analysis of surveys suggested problems were more common, with close to 2% of scientists admitting to having been involved in serious misconduct, and over 70% reported they personally knew of colleagues committing less severe detrimental research practices (Fanelli 2009). Serious misconduct such as fraud can occur in ecotoxicology just as with any discipline (Enserink 2017; Keith 2015; Marshall 1983) and when exposed, is universally condemned and, in many countries, is career ending. In contrast, the ambiguous, more nuanced issues of science integrity that all of us are likely to experience in our careers require thoughtful consideration, not condemnation. It is toward the latter that we discuss efforts toward remedies from other disciplines to examine similar issues in ecotoxicology, focusing on SETAC.

What is “science” in the context of scientific integrity?

Before we can discuss integrity in ecotoxicology and related environmental science fields, we must first distinguish what is meant by “science” in this context. Broadly speaking, environmental science includes the disciplines of biology, ecology, chemistry, physics, geology, limnology, mineralogy, marine studies, and atmospheric studies; i.e., the study of the natural world and its interconnections. The applications of environmental science extend to agriculture, fisheries management, forestry, natural resource conservation, and chemicals management, all of which have associated multi-billion-dollar industries and vocal environmental advocacy groups. The subdiscipline of environmental toxicology or ecotoxicology, pursued by SETAC scientists, studies in great detail how the natural world is influenced by chemicals, both natural and synthetic, introduced by human endeavors that are largely in pursuit of the production of desired goods and services (food, clean water, plastic products, metals, etc.). Because exposure to chemicals can have negative and sometimes unexpected consequences for people and the environment, a body of regulation has developed over the past century to control the kinds and amounts of allowable chemical exposures. Such regulations necessarily are based on scientific concepts such as Paracelsus’ directive that “the dose makes the poison” and physicochemical properties that influence transport and fate of substances. Because of the complexity, inexactitude, and uncertainty of ecotoxicology and associated sciences, rulemaking often is subject to challenge, leading to accusations of profit over people or the environment or unreasonably restrictive and burdensome requirements. Scientists are called upon to inform disputes based on their knowledge or underlying principals or enter the conversation through self-initiated in-depth literature review and commentary. Only by conscientiously adhering to fundamental principles of the scientific method can environmental scientists maintain their integrity and continue to play a valid role in environmental policy and management.

What is "scientific integrity"?

Impeccable honesty is a fundamental tenet of science. When we read a paper, we might not agree with the conclusions, authors' interpretations of its implications, importance, or many other things, but we have to be confident that the procedures described were indeed followed and all relevant data were shown, not just those fitting the hypothesis. Goodstein (1995) put it well: "*There are, to be sure, minor deceptions in virtually all scientific papers, as there are in all other aspects of human life. For example, scientific papers typically describe investigations as they logically should have been done rather than as they actually were done. False steps, blind alleys and outright mistakes are usually omitted once the results are in and the whole experiment can be seen in proper perspective.*"

Various professional and governmental organizations have established policies and definitions prescribing research integrity, responsible conduct of science, or scientific integrity. These may include broad statements of attributes such as the U.S. National Academy of Science's (NAS) six values that they considered most influential in shaping the norms that constitute research practices and relationships and the integrity of science: objectivity, honesty, openness, accountability, fairness, and stewardship (NAS 2017). More specific "research integrity" guidelines define appropriate expectations of individual researchers and their institutions and may be highly procedural. Protecting the privacy, rights and safety of human research participants and animal welfare with institutional review board clearance requirements is a common element of research integrity guidelines. Academic research integrity guidelines have been established individually or in aggregate by research funders and individual institutions (ARC 2007; Goodstein 1995; NRC-CNRC 2013; NRC 2002; Resnik and Shamoo 2011; Steneck 2006). Research institutions are usually responsible for investigating potential breaches of research integrity by its scientists, although this can create difficult conflicts of interest for the institution (Glanz and Armendariz 2017).

Whether research integrity guidelines should best be defined narrowly or broadly has been an area of controversy. As of 2015, 22 of the world's top 40 research countries had national research conduct policies, all of which included fabrication, falsification, and plagiarism (FFP), with some going further. In this context, "fabrication" is making up data; "falsification" includes manipulating studies or changing or omitting data such that the record does not accurately reflect the actual research; and "plagiarism" includes the appropriation of another person's ideas, methods, results, or words without giving appropriate credit (https://ori.hhs.gov/definition-misconduct). The Research Councils of the UK has a lengthy list of misdeeds including FFP, misrepresentation, breach of duty of care, and improper dealing with allegations of misconduct, with many subcategories (NAS 2017). In contrast, from the 1980s to 2000, the National Science Foundation (US) had defined serious science misconduct broadly to include, "*...fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting and reporting research*" (Goodstein 1995). The controversial part was the catchall phrase "*practices that seriously deviate from those commonly accepted...*" To the stewards of public science funds, such a catchall phrase was preferable to an itemized lists of all potential avenues of mischief, yet it raised the specter of penalizing scientists who strayed too far from orthodox thought (Goodstein 1995). In 2000, this definition of disbarring research misconduct was narrowed to just "*fabrication, falsification, or plagiarism in proposing, performing, or reporting research*" with lesser offenses classified as questionable research practices. Other misconduct was defined as "*forms of unacceptable behavior that are clearly not unique to the conduct of science, although they may occur in the laboratory or research environment*." Yet only FFP research misconduct findings were subject to reporting requirements to federal science funding agencies, with questionable science practices or other misconduct handled locally (NAS 2017; Resnik et al. 2015).

In many countries, there is an active debate about whether a legal definition is appropriate for something that is really an academic judgment rather than a legal one. Denmark recently similarly narrowed its broad definitions of research misconduct to only FFP following high profile cases in which scientists succeeded in having their academic misconduct findings overturned in the courts. Yet if research conduct policies are considered "academic" without legal weight, institutions may have difficulty enforcing polices, such as when deliberate intent is required to be shown and the researcher claims "honest mistake." For instance, the U.S. Office of Research Integrity found that a tenured professor had committed research misconduct by inappropriately altering data in five images from three papers. Yet when the university sought to terminate her, she fought back contesting the university's procedures, and the university ultimately paid her $100,000 USD to leave (Stern 2017). In private research, it is not obvious which scientific integrity concepts have the force of law. In an example from the U.S., testimony of egregious breaches of scientific integrity norms (including faking credentials and selective publication of only favorable results) was disallowed in a court dispute between two private companies because there was no federal law on scientific integrity (Krimsky 2003).

Unfortunately, the "other misconduct" that scientists may commit can reflect that of any work place, such as abuse of power; bullying, sexual coercion, assault, or harassment; misuse of funds;

sabotage; taking advantage of students or subordinates; specious whistleblowing or retaliation against valid whistleblowers; to name a few (e.g., Ghorayshi 2016; Gibbons 2014; http://retractionwatch.com). The exclusion of such malfeasance from "research misconduct" has been questioned. For example, a researcher who failed to meet her study objectives after being sabotaged by a rival argued that she was further penalized by being instructed not to divulge the reason for her study failures to her funders (Enserink 2014). In contrast, institutions often do go beyond the minimum "FFP" definition in their policies (Resnik et al. 2015), which has led to objections of conflation of egregious misconduct such as fraud with failure to comply with administrative requirements that did not compromise data validity (Couzin-Frankel 2017).

The U.S. National Academy of Science (NAS 2017) recently argued that the definitions of research misconduct as fabrication, falsification, or plagiarism were too narrow. In particular, questionable research practices were more than just "questionable," but were clear violations of the fundamental tenets of research and were given a less ambiguous label of "detrimental." Consensus detrimental research practices were:

1. Detrimental authorship practices that may not be considered misconduct, such as honorary authorship, demanding authorship in return for access to previously collected data or materials or denying authorship to those who deserve to be designated as authors. [Here we think it is important to distinguish between pairing a data reuse with an invitation to collaborate and share authorship versus demanding authorship as a condition of data access (Duke and Porter 2013)].
2. Not retaining or making data, code, or other information/materials underlying research results available as specified in institutional or sponsor policies, or standard practices in the field.
3. Neglectful or exploitative supervision in research.
4. Misleading statistical analysis that falls short of falsification.
5. Inadequate institutional policies, procedures, or capacity to foster research integrity and address research misconduct allegations, and deficient implementation of policies and procedures, and
6. Abusive or irresponsible publication practices by journal editors and peer reviewers (NAS 2017).

The term "scientific integrity" is sometimes used synonymously with research integrity. However in recent usage, the term has included insulation of science from political interference, manipulation, or suppression of science (Doremus 2007; Douglas 2014). The term "scientific integrity" has been used in government science policy in the United States. There, scientific integrity guidelines were developed in an overarching sense that includes research integrity at the individual and institutional level but were also intended to protect federal scientists from political interferences. Political officials were not to alter or suppress scientific findings, and transparency was encouraged in the preparation of the government-supported scientific research (Obama 2009; Stein and Eilperin 2010). The scientific integrity guidelines in the US were followed by derivative policies intended to put substance to the transparency provisions, requiring open-access to federally funded research articles and more importantly, requiring archiving and public availability of the underlying raw data (Holdren 2013). These broad policies become more specific and procedural in government science agencies, and expanded to codes of scholarly and scientific conduct such as a list of 19 principles for the U.S. Department of Interior (U.S. Department of Interior 2014).

We expect the vast majority of scientists consider themselves to hold science integrity, as self-defined in terms of honesty, transparency, and objectivity, sticking to the research question and avoiding bias in data interpretation (e.g., Shaw and Satalkar 2018). Yet most scientists will encounter ethically ambiguous situations. For instance, some may feel that they struggle to advance science against a rising tide of administrative requirements accompanied by declining support for science and increasing competition for funding. When does cutting through bureaucratic institutional requirements cross the line from being commendable efficiency to violating research integrity rules? Using grant/project funds for unrelated purchases or conference travel? Should minor misbehaviors such as posting ones' article on a website after signing a publication and copyright transfer agreement with the publisher agreeing not to do so still be considered misbehaviors when done by many? When does cleaning data become cooking data when, for example, anomalous values are suppressed? There are many ethically ambiguous situations in which scientists may consider doing the "right thing" (compliance with all rules) might need to be balanced with doing the "good thing," especially when the welfare of others such as students or subordinates is involved (Johnson and Ecklund 2016).

To us, scientific integrity can be simplified to cultures of personal integrity plus a few profession-specific provisions of transparency and reproducibility. At their roots, these norms are those children are hopefully acculturated to in primary school. *Tell the truth, and tell the whole truth* (no data sanitizing, selective reporting, and report all conflicts), *tell both sides of the story* (avoid bias), *do your own work* (no plagiarism), *read the book, not just the back cover before writing your report* (properly research and cite primary sources), *show your work for full credit (*transparency*)*, *practice makes perfect* (rigor), *share* (publish your work and data in peer-reviewed outlets for collective learning), and *listen* (with humility and collegial fraternity to observations and suggestions of others). Finally, the golden rule "*do unto others as you would have them do unto you*" should resonate throughout the professional interactions of environmental scientists, and especially in peer reviewing and data sharing. When encountering an inevitable science dispute, keep criticisms objective, constructive, and focused on the work and not the worker; do peer reviews of your rivals' work as you would hope to receive reviews of your own, reward and recognize good behavior in science, and so on.

The interested scientist: conflicts of interest, competing interests, and bias

Although we would like to believe that outright fraud or deliberate campaigns to manipulate science are rare in the environmental sciences, at some points in their careers almost every practicing scientist must grapple with questions of conflicting or competing interests and must guard against bias in approaches and interpretation. Conflicts of interest are often narrowly defined to situations where the scientists or their employers stand to gain financially from their work.

The term "conflicts of interest" is commonly defined as financial conflicts. One definition is "a set of conditions in which professional judgment concerning a primary interest tends to be or could be perceived to be unduly influenced by a secondary interest (such as financial gain). More simply, a conflict of interest is any financial arrangement that compromises, has the capacity to compromise, or has the appearance of compromising trust (Krimsky 2003, 2007). The term "competing interests" is often used where non-financial factors compete with objectivity, such as allegiances, personal friendships or dislikes, career advancement, having taken public stances on an issue, political, academic, ideological, or religious affiliations (Nature Editors 2018; PLOS Medicine Editors 2008). Bias in study design or data interpretation may arise from either conflicts or competing interests and can be either overt or unrecognized by the scientist(Suter and Cormier 2015b)

Generally, the concern over conflicting or competing interests in science is that secondary interests such as financial gain or maintaining professional relationships compromise the primary interest of upholding scientific norms such as reporting data accurately and completely, interpreting data appropriately, and acknowledging value judgments or interpretive assumptions (Elliott 2014). Conflict of interest policies may be better developed in the biomedical fields than in the applied environmental sciences because the former often involves human participants, and because of the strong financial ties between academia and the pharmaceutical industry (Tollefson 2015). For instance, if a research team is reporting on the efficacy of a medical device or a pharmaceutical, and they or their employers hold a patent or stand to gain financially from a positive report, then they clearly have a financial conflict of interest (Figure 1).

The mere existence of a potential conflict of interest should not alone throw results in doubt where it is disclosed and acknowledged appropriately. However, although most authors in the environmental sciences routinely disclose funding sources that could be perceived as potential conflicts of interest, major omissions have occurred (Oreskes et al. 2015; Ruff 2015; Tollefson 2015). For instance, the findings of a study on risks of contamination from natural gas extraction from hydraulic fracturing of bedrock were undermined when it came out (apparently unbeknownst to the university) that the research supervisor was being paid 3X his university salary by serving as an advisor to an oil and gas company invested in the practice. The failure to disclose this financial relationship in the publication brought the study's objectivity and credibility into question, independent of its substance (Stokstad 2012). Authors and journals have been criticized for gaming ethical financial disclosure requirements, such as by overly narrow disclosures or disclosing a conflict in the cover letter to the editor accompanying the manuscript (which doesn't get published or shared with reviewers) but not including it in the actual article (Marcus and Oransky 2016).

It should be noted that the severe conflicts of interest that some academic biomedical researchers have created for themselves by setting up business interests to directly and personally profit from their research outcomes (Krimsky 2003) are probably much less of an issue in the environmental sciences. Dual affiliations and the resultant potential for divided loyalties for university researchers have certainly come to light in the environmental sciences, such as if the scientist has a public facing, disinterested, researcher identity but privately has set up spin-off personal, business interests (Fellner 2018; Stokstad 2012). While we are not aware of any systematic review, we think these situations are far less pervasive in the environmental sciences than biomedicine. Rather, in ecotoxicology and environmental chemistry, the more common (and less insidious) concern for authors and institutions to be self-aware of the potential for funding bias through unconscious internalization of the interests of their research sponsors. The informative value of conflict of interest or funding disclosures vary. The shortest (and least informative) statement we have seen was that "*the usual disclaimers apply*" (Descamps 2008), while the detailed disclosures in biomedical literature can go on for pages (Baethge 2013; ICMJE 2016). Funding sources can be obscured by channeling funding through intermediaries, such as a critical review of cancer risks from talcum powder funded by a law firm involved in toxic tort litigation (Muscat and Huncharek 2008). Requirements for highly detailed disclosures risk diminishing their importance to that of the "fine print" cautions in commerce that are seldom read. Much like computer software user terms and conditions that have to be clicked past or the ubiquitous consumer product safety stickers that may be written more to avoid product liability claims than for practical safety, detailed conflict of interest disclosures may reach a point of diminishing returns. Twenty years ago, Goodstein (1995) groused that he was tired of reading disclosure statements that were longer than the methods sections in papers, and they have been expanded upon since. Our view is that in ecotoxicology and the environmental sciences, simple, unambiguous statements of the funding sources that directly or indirectly allowed the work to be completed should generally be sufficient.

Non-financial competing factors may also compete with scientific objectivity. Factors or values such as these are usually termed "competing interests" reserving "conflicts of interest" for financial conflicts (Nature Editors 2018; PLOS Medicine Editors 2008). In our observations, competing interests are rarely if ever mentioned in environmental science publications. Rather, they are often discussed behind the scenes, such in correspondence between an editor and potential reviewers, along the lines of "*yes I would be happy to review this article and believe I can be objective, however, you should know that I used to be a labmate of the PI and we collaborated on an article 3 years ago*." Whether or how competing interests or values affect the assumptions and perspectives of scientists' should be more formally stated is an area of rich debate in the philosophy of science literature (Douglas 2015; Elliott 2016; PLOS Medicine Editors 2008).

We reiterate our belief that the existence of a potential conflicts or competing interests is a ubiquitous part of the environmental science landscape and do not indicate poor science. Most scientists strive to present unbiased data and interpret their data evenhandedly. However, the varied experiences of scientists can influence their perspectives in ways that they may not recognize themselves. The transparency in disclosure reminds the reader to consider perspectives and alternate interpretations when judging the merits of a study, synthesis paper, or risk assessment.

Bias

Many of the published concerns in the environmental science literature come down to cognitive bias (Figure 2). Science is not value free, and personal bias in interpreting science is often related to differing worldviews (Douglas 2015; Elliott 2016; Lackey 2001; Nuzzo 2015). For instance, the collapse of major fisheries that ostensibly had been scientifically managed for sustainable yields helped inspire the Precautionary Principle. This philosophy sought more cautious management and the reversal of the burden of proof for sustainable exploitation of natural resources (Peterman and M'Gonigle 1992). Those with precautionary principle or risk assessment worldviews may interpret the same set of facts very differently. The precautionary principle adherent may emphasize absence of conclusive evidence of safety, and the risk assessment adherent may emphasize absence of conclusive evidence of harm (Fairbrother and Bennett 1999). In such settings, values and biases are interwoven. Even self-disciplined scientists who seek openness and objectivity carry some biases from experiences and acculturation (here meaning how working in different environmental organisations can lead scientists to modify their perception and thinking). Recognizing sources of bias does not imply ill intent, for just the process of acculturation to a particular place of employment can bias perceptions and inclinations (Brain et al. 2016; Suter and Cormier 2015a, b).

Professional societies such as SETAC can serve as a form of acculturation; some of the authors of this essay have been active members of SETAC for much longer than they have been employed by any single employer. Even self-disciplined scientists who seek openness and objectivity carry biases from their experiences. What becomes particularly difficult to self-regulate is the convergence of cognitive bias, a human nature to seek to please one's patron, and the interests of one's employer or client. For instance, studies funded by drug or medical device makers tend to find positive effects that favor the company funding the research (Lexchin et al. 2003; Smith 2006), and the funding effect for studies of chemical toxicity may lean toward finding negative effects (Bero et al. 2016; Krimsky 2003, 2013). However, concordance between a funder's self-interest and research findings does not alone indicate bias. Alternatively the industry-funded researchers could have deeper knowledge of a drug or chemical than non-profit funded academic researcher who might have less extensive experience, the industry-funded work was more thoroughly vetted based on internal research, or the industry-funded scientists might have better ability to obtain the resources and skill to carry out well focused and rigorous research (Krimsky 2013; Macleod 2014). It is doubtful that these influences can be completely separated. To us, disclosure, transparency and balanced external reviews are presently the best pragmatic approach to managing cognitive biases.

Tit for tat, adversarial claims of bias in the scientific literature doubtfully advance the science. Conflicting perspectives can become personalized and intractable. How to know which is more credible? Neither? Both? Food nutrition researchers pointed out examples of selective data interpretations and publication bias in obesity research in relation to sweetened beverage (soft drink) consumption and in the health benefits of breast feeding They termed this distortion of information to further what may be perceived to be righteous ends as "white hat bias" (Cope and Allison 2009). However, the conflicted financial backing from the soft drink industry and from manufacturers of baby formula contributed to counter-criticisms of funding bias (Bes-Rastrollo et al. 2014; Harris and Patrick 2011; Mandrioli et al. 2016). Unresolved in the claims and counter-claims of bias and financial conflicts of interest was what advice was most credible.

In environmental toxicology as well, controversies over the best interpretation of sometimes ambiguous facts can become entrenched and focused on the people holding differing views as much as the evidence behind the different views. Examples include deeply held and personalized disagreements over risks of atrazine to amphibians (Hanson et al. 2018; Hayes 2004; Raloff 2010; Rohr and McCoy 2010; Solomon et al. 2008); sufficiently safe levels of selenium for fish and birds (Renner 2005; Skorupa et al. 2004); and a dispute that was maintained for over 20 years about whether an oil spill resulted in indirect harm to salmon (Burton and Ward 2012).

These intractable, mutual bias criticisms make it very difficult for non-specialist readers to make informed judgements of which is the more credible science.

Suter and Cormier (2015a) noted that conflicting assessments on the same question that have been produced by government agencies, industries, and environmental advocacy groups suggest that biases occur during assessment processes. Sources of bias include personal bias, regulatory capture, advocacy, reliance on volunteer experts, biased stakeholder and peer review processes, literature searches, excluding new science through dependence on standard methods, inappropriate standards of proof, misinterpretation, and ambiguity. Suter and Cormier (2015a) argue for assessors to adopt practices that would increase objectivity, transparency, and clarity of assessments and syntheses.

Some particularly challenging situations in ecotoxicology

Some situations that seem particularly challenging for researcher and institutions to maintain scientific credibility warrant mention. Elliott (2014) argued that scientific findings that are ambiguous or require a good deal of interpretation or are difficult to establish in an obvious and straightforward manner are prone to bias, particularly if strong incentives to influence research findings in ways that damage the credibility of research are present. In environmental toxicology, risk assessments or critical reviews fit that test and can be vulnerable to bias, particularly when funded by sponsors with financial interests in the findings (Suter and Cormier 2015a). This can be heightened by how variability and large uncertainties are handled in environmental toxicology and associated risk assessments and syntheses – for example extrapolation of results from one or more species to protection of wide swaths of our world’s biodiversity; or the difficulty in reproducing field studies; or the variability of chemical exposures across diverse and expansive landscapes and waters. These challenges may lead to differences of opinion on methods for drawing conclusions to support decision-making that, while prone to bias, have, at their root, the need for drawing conclusions in the face of uncertainty.

Costs of large-scale projects to remediate contaminated environments such as sediments contaminated by urban and industrial sources, aged industrial facilities, or large mining operations can be enormous, running to the hundreds of millions of dollars or more (Gustavson et al. 2007; McKinley 2016). In “polluter-pays” schemes, the potential financial liability associated with such projects could imperil the ongoing viability of companies, which in turn would harm the livelihoods of employees, among other social disruptions. In such a setting, the scientists working on behalf of the those who may have to incur the costs of cleanup might understandably be more cautious about the potential for misguided remediation following Type I error (e.g., falsely discovering environmental degradation) than Type II error (failing to discover degradation when in fact it is occurring), when the science is ambiguous. Conversely, the regulatory scientists entrusted to provide scientific advice to protect environmental quality might be obliged to err on the side of precaution and be more accepting of risk of Type I error, especially when it is “other peoples’” money at stake.

While science ethicists and the NAS (Boden and Ozonoff 2008; Elliott 2014; Krimsky 2005; NAS 1992) have emphasized industry funding bias risks, these risks are not unique to industry funding of science. For examples, many countries have provisions for natural resource damage assessment and restoration (NRDAR) to compensate the public for lost opportunities following shipwrecks, oil spills, releases of industrial chemicals, and so on (Boehm and Ginn 2013; Descamps 2008; Flamini et al. 2004; Goldsmith et al. 2014). These assessments rely on science to some degree to establish linkages from the release to harm to the environment. In turn, trustees of natural resources rely on science advisors to assess the extent and scale of injuries (adverse effects) and the monies needed to restore the lost services. In large incidents, the responsible parties will inevitably retain their own science advisors. The resolution of complex situations is resolved either by negotiation or adversarial litigation (Flamini et al. 2004; Goldsmith et al. 2014). This environment produces an atmosphere with strong incentives for plaintiff/trustee science advisors to maximize the magnitude and spatial extent of effects to the environment and to downplay uncertainties or the influence of potential other, non-compensable stressors and vice versa for those scientists retained to help defend against claims. Maintaining objectivity and advancing science in such a work environment would require extraordinary self-discipline by the individual scientists, an institutional environment emphasizing science credibility, and an openness to external, disinterested review (Boden and Ozonoff 2008; Elliott 2014; Wagner 2005).

While at least in some jurisdictions, monies from NRDARs must go to restoring the damaged public natural resources (beyond paying for salaries or consulting fees) and cannot be used to enrich those pursuing the cases. Toxic torts by comparison, pursue damages on behalf of private individuals or groups who consider themselves to have been harmed by exposures to toxic chemicals. Toxic tort cases are adversarial proceedings with the lawyers expected to advocate only for their client, and expert witnesses are paid to present testimony to support just one side. These torts may be highly lucrative for the plaintiff attorneys who select the science testimony. For example, in the Vioxx litigation the share for plaintiff lawyers was about $1.5 billion (32%) of the $4.85 billion settlement (McClellan 2008), and in successful asbestos litigation the average share of payouts going to the victims was only 37% (Elliott 1988). The lures and risks of such immense payouts in toxic torts create strong incentives for biased science. At best, critical reviews or product defense studies conducted for toxic tort science should be regarded with skepticism.

Defense of science and engineering in favor of protecting enterprises reflecting years of devoted work is understandable but becomes dangerous when objectivity is compromised. Case studies such as the Vioxx case, in which the maker of the drug downplayed increased risks of mortality from a successful product in which they were deeply vested (Curfman et al. 2005; McClellan 2008) and the cross-claims of blame between the engineering consultants and the mine operator in the aftermath of the Mount Polley mine tailing dam failure (Topf 2016), remind us that objective science (including recognizing and disclosing uncertainty, and encouraging additional science to narrow that uncertainty) is good business.

Academic – Industry Collaborations

The role of industry funding and concerns of perceived conflicts of interest in academic-industry collaborations have been addressed in literature and are a common element in institutional research integrity policies (Elliott 2014; Resnik and Shamoo 2011). Often through philanthropic foundations, industry may contribute to basic science education and research to strengthen regional universities and further the science literacy of potential workforce and society. Industry may also support applied ecotoxicology and other environmental science research to inform specific scientific questions that affect their business interests. When industry and academic research interests become at least partially congruent, academic scientists may actively seek out such interest and support for their projects and graduate students. Pragmatically, academic-industry collaborations are necessary since public funding alone may be insufficient to support graduate research or to address important questions relevant to industry and society. In the US, about 40% of national research and development is funded by the private sector (NAS 2017). In the US, public funding for university research on the effects of chemicals in the environment has consistently declined since 2000 (Bernhardt et al. 2017; Burton et al. 2017), which implies that without industry-academic collaborations, there would be much less substantive university research. As a university president quipped, "*the only problem with tainted research funding is there t'aint enough of it*" (Krimsky 2003).

Benefits of collaboration run both ways, with expertise from academic and public sectors helping industry find solutions to lessen or avoid contributing to environmental problems (Hopkin 2006). Edwards (2016) lays out several principles for successful, durable industry-academic collaborations, including: establish clear quality criteria and make them public; mandate data sharing; subject work to independent oversight before public release; and enshrine public ownership for all research outputs. Further, effective collaboration between industry and academic scientists requires industry to provide expertise as well as funds. Collaboration with industry scientists engenders a shared desire to succeed and creates a sense of ownership of a project (Edwards 2016). The interchange of science through academic, industry, and government scientists is deeply rooted in SETAC culture, and the favorable views of the authors toward working across sectors is undoubtedly influenced through our history with SETAC. However, industry support to academics or others in support of applied environmental questions may come with inherent conflicts of interest, and critics may consider scientists as collaborators in the pejorative sense of the word (Hopkin 2006). This setting requires vigilance from both industrial research sponsors and recipients to avoid unconscious bias.

While readers might presume situations in which individuals or institutions with strong incentives to influence research findings consistent with their financial interests will do so, it is important not to judge a study solely by its funder, nor to presume the sponsor's preferred outcome. For example, an energy company sponsored a study to see if they could develop a scientific case for relief from costly requirements for meeting dissolved oxygen criteria in a river downstream of its hydroelectric dam. Instead they developed evidence that the existing criteria could impair hatching salmon (Geist et al. 2006). The company scientists easily could have buried the results, which could have been discounted as being from novel techniques. Their path of least resistance would have been to leave the study in the file drawer, rather than going to the trouble of defending novel science and publishing it in the open literature. In the long-view, a reputation of science credibility may be more valuable for companies than short-term project benefits.

Other examples include scientists from mining and metals trade groups publishing studies showing that existing USEPA criteria for zinc and other metals could be under-protective of aquatic species or entire communities (Brix et al. 2011; DeForest and Van Genderen 2012). Conversely, a university quantitative ecologist accepted support from an environmental advocacy group (through university channels) to model the potential population-level effects of elevated selenium from mining on local native trout populations (Van Kirk and Hill 2007). As the advocacy group had been a persistent opponent of the mining operations, officials from the influential mining company apparently presumed that the academics' work would also be biased to favor the advocacy group's positions, and they questioned the researchers' probity (Blumenstyk 2007). In fact, the selenium concentrations projected by these academics to cause detrimental population-level effects were higher than concentrations previously derived by industry-funded consultants who themselves had been on the receiving end of bias implications because they were aligned with corporate interests (Skorupa et al. 2004; Van Kirk and Hill 2007). Unfortunately, these examples are countered by unfavorable examples in which studies were funded as part of deliberate strategies to shape the science to fit business interests. This "tobacco strategy" has been asserted in regard to various substances such as asbestos, benzene, chromium, lead, vinyl chloride, and more (Anderson 2017; Cranor 2008; Krimsky 2003; Michaels 2008; Oreskes and Conway 2011; Sass et al. 2005). As these examples show, judging science and scientists solely by their funding or affiliation is unfair and may lead to misjudgments.

In keeping with the adage to be careful judging a book by its cover or wine by its label, judging science by its funder or by presumed interests or leanings of the scientists can lead to mistaken and unfair perceptions. Brain et al. (2016) pointed out that the career path of environmental scientists is often ambiguous and whether scientists ended up in careers with industry, academic, or government science has more to do with chance and timing of opportunities rather than a particular desire to work in one sector or another. Such is often the case with academic and government scientists who work with industry to jointly fund or investigate a science question of mutual interest (Hopkin 2006). The convergence of scientific interests with financial interests can lead to a good marriage, so long as the parties are principled and forthright with each other. While there may be a perception that research contracts are highly restrictive, in our experiences these agreements establish expectations of academic freedoms. "Interested science" should be viewed with open-minded skepticism, and studies with immense financial implications warrant a higher level of scrutiny than others (Krumholz et al. 2007; Suter and Cormier 2015b; van Kolfschooten 2002). It does not necessarily follow that interested science is wrong or tainted. Ensuring transparency and complete data reporting is one tangible step researchers can take to improve credibility of and perceptions toward industry-academic collaborations.

A scientific society founded on the principles of balancing competing interests

Scientific societies have important roles in promoting scientific integrity and ethical conduct, such as establishing codes of ethics which include disclosure of conflicts of interest, being a focal point for developing and communicating discipline-specific standards to foster research integrity, and providing educational material (AAAS 2000; NAS 2017).

We think the Society of Environmental Toxicology and Chemistry (SETAC) is notable for its directed and sustained efforts to balance competing perspectives in its deliberative processes and other activities. The founding principles and structure of SETAC sets out a tripartisan structure with regulatory, industrial, and academic scientists (Bui et al. 2004; Menzie and Smith 2018). As a result, SETAC now has well developed norms for balancing interests, inclusiveness of differing viewpoints, and neutrality in the reporting. These norms have enabled SETAC to be regarded as a source of consensus-based science with successful partnership or advisory roles in United Nations programs and conventions such as the United Nations Environment Programme's (UNEP) Global Mercury Partnership, Stockholm Convention on persistent organic pollutants, UNEP-SETAC Life Cycle Initiative for reducing hazardous waste as well as informing national-level legislation (Augspurger 2014; Mozur 2012). The intended balanced representation of industry, government, and academia isn't always achievable, for there are also guidelines for gender equity, geographic representation, and of course people have to be willing to volunteer. Further, the tripartisan emphasis underrepresents scientists from environmental advocacy groups. These groups are influential for shaping public debate, policy and law on environmental issues, but their low participation in the Society suggests that they may not be attracted to or feel welcomed by a "hard" scientific society such as SETAC, and meeting costs may be a barrier. Despite these imperfections, the norms of seeking to balance potentially conflicting interests and to provide a safe forum to express differing scientific viewpoints are deeply ingrained in the Society's culture and activities.

Promoting scientific integrity in ecotoxicology

While "scientific integrity" is ultimately a subjective judgment that cannot easily be reduced to review checklists, there are some general points to maintain in ecotoxicology and related science. These include relevance, rigor, reproducibility, objectivity, and transparency.

Relevance

By definition, environmental chemistry and ecotoxicology are concerned with how chemicals, both natural and synthetic, pose a threat or influence the natural world (Johnson et al. 2017). Because of pragmatic and ethical constraints, research in this domain is often done in laboratory environments, testing cultured laboratory organisms or cell lines or other in vitro surrogates for organisms. However, the intent of such research invariably still has some intended relevance to conditions that occur in the environment. We have seen articles in ecotoxicology literature discussing some novel research based on under-tested taxa, underappreciated endpoints, unexpected multiple stressor effects, or unanticipated indirect effects via untested commensal microbes. An article may start out with an introduction on the ecological importance of the novel work, the work is reported, and then the discussion closes arguing that ecological importance of their work, how it should change the thinking in the field, and management implications. Yet to obtain their desired experimental effects, exposure concentrations may have been orders of magnitude higher than those typical in the real world, or exposure routes, chemical forms, or dilution media may be unlike those that the organisms could encounter in nature (Johnson and Sumpter 2016; Mebane and Meyer 2016; Weltje and Sumpter 2017). When authors present such studies with a narrative on the ecological importance of their topic, this may be a form of misrepresentation.

Rigor

Funders, journals, and institutions reward novelty, such as the short-lived discovery of a bacterium that grows with arsenic instead of phosphorus (Alberts 2012). Highly selective journals with article acceptance rates of 10% or less preferentially publish findings that are sensational or at least surprising. These incentives are influential because universities and research institutes often hire and promote scientists based on their record of acquiring grant money and the number of publications times the journal impact factors of the journals published therein (Parker et al. 2016). With finite career opportunities and high network connectivity, the marginal return for being in the top tier of publications may be orders of magnitude higher than an otherwise respectable publication record (Smaldino and McElreath 2016). The editorial quest for novelty has led to publication of questionable articles in elite journals, such as one positing that caterpillars were the results of accidental sex between insects and worms (Borrell 2009). Top tier journals also tend to have higher retraction rates than mid-tier journals, suggesting that rigor has sometimes been compromised in the competition for paradigm shifting results (Nature Editors 2014).

In ecotoxicology, Harris et al (2014) describe 12 basic principles of sound ecotoxicology that should apply to most environmental toxicity studies. These principles range from carefully considering essential aspects of experimental design through to accurately defining the exposure, adequate replication, unbiased analysis and reporting of the results, and repeating experiments that yielded surprising or ambiguous responses. There are ample opportunities for improvement. For example, Harris and Sumpter (2015) asked a very basic question of a sample of studies published in 2013 in three leading ecotoxicological publications: was the concentration of the test chemical actually measured? Of the studies reviewed from *Environmental Toxicology and Chemistry*, 20% failed this basic aspect of experimental credibility, as did 33% and 41% of ecotoxicology studies published in *Aquatic Toxicology*, and *Environmental Science and Technology*, respectively (Harris and Sumpter 2015).

While Harris et al. (2014) emphasized laboratory-based studies, field-based environmental effects studies replace the challenges of the artificiality and questionable relevance of some laboratory-based toxicity testing, with different, messy, real world challenges. Closely related to the 12 principles described by Harris et al, we suggest 8 basic principles relevant to most field-based ecotoxicological studies or environmental effects monitoring.

1. Development of a thorough understanding of the issues to ask informed questions. The available literature should be thoroughly vetted to inform the need for experimentation;
2. A thorough understanding of the questions being posed is an essential prerequisite for designing a robust, reliable, reproducible experiment. Incomplete understanding of the questions leads to vague and often misguided experimentation that undermines the scientific process (Lindenmayer and Likens 2010; Suter et al. 2002);
3. The ability to identify and reliably measure sensitive indicators (Melvin et al. 2009),
4. Careful attention to appropriate reference conditions to avoid potential, actual effects being masked by variability or confounding factors introduced by differences between the reference and test site environments (Arciszewski and Munkittrick 2015;

Mebane et al. 2015). For example, beaches on rocky headlands and protected bays will have very different benthic invertebrate communities, as do flowing rivers and impounded reservoirs. Study designs that attempt to detect pollution effects on communities across such disparate habitats may have very low discriminatory power and by failing to account for natural variability, adverse pollutant effects could be obscured (Buys et al. 2015; Parker and Wiens 2005; Wiens and Parker 1995);

5. Try to study a number of locations that vary in the degree of the factor under investigation, such as chemical pollution, in order to (hopefully) demonstrate a positive relationship between exposure to the environmental factor of interest and the effect of that factor.
6. Time and patience: Just as experimental exposures need to be of appropriate duration for effects of interest to be manifested, environmental monitoring needs to be maintained long enough to pick up true trends if present, or to convincingly argue that trends are not present (Lindenmayer and Likens 2010).
7. Specific definitions of what effects are considered negligible or of concern (Melvin et al. 2009; Munkittrick et al. 2009; Power et al. 1995).
8. Avoid power failures: use a statistical approach appropriate to the question, considering statistical burden of proof issues. For instance, $P>0.05$ in testing for trends or differences between locations does not by itself show the lack of trend or effects (Dixon and Pechmann 2005; Mudge et al. 2012).
9. Transparent reporting with detailed methods and raw data sufficient for others to reproduce the analyses or to further examine the data using alternative analyses (Duke and Porter 2013; McNutt et al. 2016; Schäfer et al. 2013).

Reproducibility

Reproducibility is one indicator of reliable research. However, the inability of researchers to reproduce influential studies of others or their own has garnered enough attention to be called a “reproducibility crisis” (Baker 2016a; Henderson and Thomson 2017). However, not all studies are easily reproduced. Environmental data are often messy and field studies are more often observational than experimental. Large scale, ecologically realistic studies such as long-term, experimental lake studies difficult to do even once; and no one wishes to replicate mishaps such as tailings dam failures or oil spills (Parker and Wiens 2005; Schindler 1998; Wiens and Parker 1995). Such studies require a logical system for causal inference to separate cause-and-effect from serendipitous correlations (Norton et al. 2002; Suter et al. 2002). Even rigorous laboratory studies may be difficult to replicate due to the highly variable nature of biological systems and unanticipated responses to unknown factors. Demands for reproducibility may favor industrial science over academic science. Industry often works within strict Good Laboratory Practice (GLP) rules and with well-studied species tested through standardized protocols (Elliott 2016). Academic science is often framed around education, and grants and graduate student researchers are usually encouraged to go after something new and novel; protocols may be developed as they go, and quality control may be uneven (Baker 2016b; Figure 3). Obstacles to adopting formalized quality management systems such as GLP in small research settings may include costs, lack of resources, lack of mandate, independent cultures, and high turnover. Nevertheless, even if regulatory GLP compliance is not required, small academic research facilities can benefit from embracing core components of GLPs, such as defining responsibilities, maintenance and sanitation of common lab spaces, equipment and materials, well defined experimental protocols, quality control testing, data reviews, audits, and archiving (Bornstein-Forst 2017; Moermond et al. 2016).

Better experimental protocols that are easier to follow is one tangible way to strive for better reproducibility and transferability of both novel and standard experimental methods. Multimedia experimental protocols could be much easier to explain and teach techniques than the conventional, densely worded, printed protocols (Figure 4). The Journal of Visualized Experiments (JoVE) is an innovative peer reviewed, science methods journal in which its articles are a unique blend of the conventional printed article with professionally produced videography. Ecotoxicology methods articles have begun to be published in this format (Calfee et al. 2016; van Iersel et al. 2014). The field would benefit from broader use of new visualization techniques to document new methods and to improve education and training on techniques that need to be highly standardized to be repeatable. At the minimum, with the availability of electronic data repositories and supplemental information in journals, there is no reason why detailed methods including video demonstrations cannot be published.

Reproducing a statistical summary or model run reported in a scientific publication when the underlying data and code are provided and explained is one thing. Reproducing an actual complex experiment is hard and is rarely attempted, unless perhaps the results are novel and have a high regulatory or societal impact. Even under the best of circumstances, such as when the original researchers have the resources to and are diligent enough to repeat an experiment in the same lab with as close to identical methods as they could manage, it can be difficult or impossible to produce the same result twice (Owen et al. 2010). Nosek and Errington (2017) caution that if investigator #2 reports that the results of study #1 could not be reproduced, that does not indicate which is more credible: result #1, #2, neither, or both. Further, much of the “reproducibility” debate in the natural sciences is focused on cell biology or human behavior (psychology) experiments, which may be more tractable to reproducibility studies than messy environmental observational or experimental studies. Especially with complex biological testing such as multi-generation tests, a green thumb husbandry factor may bring together art and science to environmental chemistry and toxicology. Subtle methods differences, strain differences or stochastic events can be so puzzling that investigators are left thinking demons must have snuck into their study and interfered with one treatment but not others (Hurlbert 1984). (We presume that Hurlbert’s (1984) suggestions for exorcisms or human sacrifice for troubleshooting suspected demonic intrusions, might run afoul of some contemporary institutional review board policies.)

Still, reproducibility is a core tenet of science and successful reproduction adds confidence in the credibility of novel findings. Divergent but individually credible results may further advance the science by illuminating important aspects missed in the initial study (Owen et al. 2010). If for instance, an investigator were to find a novel, major adverse effect of a class of chemicals to a previously untested taxonomic group, then other equally diligent investigators should be able produce similar effects in other research settings, even if the test conditions were only similar. A standalone paper from the 1970s that a snail was anomalously sensitive to Pb was skeptically regarded. Over 30 years later, this open-minded skepticism led to follow-on studies from a new generation of scientists that not only affirmed the anomalous early report of sensitivity but also led to important advances in comparative physiology and underlying mechanisms of toxicity (Brix et al. 2012). Similarly, early reports that freshwater mussels and other mollusks were unusually sensitive to ammonia were not widely persuasive. After repeated studies across multiple laboratories and species showed similar findings, the issue gained traction with standardized method development, inter-laboratory round robin testing, and attention by environmental managers (Farris and Hassel 2006; USEPA 2013).

Individual investigators may not always have the opportunities for self-replication, but best practices call for repeating what one can (Harris and Sumpter 2015). In field studies, multiple measures of exposure, multiple years of field data, and so on give credence to findings. We recognize that all science has practical resource limits and we are not going as far as arguing that novel findings from small sample studies should never be published. Rather, the appropriate conclusion from such studies is along the lines of "if these findings turn out to be repeatable, they could be an important development." In our view, novel, major findings that are supported only by a one-off study are best regarded as tentative.

Transparency

Transparency in reporting research, including all the relevant underlying data that were relied upon in the paper, has become a critical element of integrity in science. Science's claim to self-correction and overall reliability is based on the ability of researchers to replicate the results of published studies (Nosek and 39 co-authors 2015). Studies cannot be replicated or even reconstructed if scientists will not share additional data, information, or materials from published studies, and we believe that upholding such ethical norms is every scientist's responsibility. The embrace of the principle of transparent reporting has been uneven across disciplines, and the field of ecotoxicology has certainly not distinguished itself as a leader in this regard (McNutt et al. 2016; Meyer and Francisco 2013; Parker et al. 2016; Schäfer et al. 2013; Womack 2015).

Researchers in ecotoxicology and environmental chemistry have long only presented highly reduced data summaries. The only "data" included in some publications are crowded figures and tables with results of statistical outputs, such as F- values, effects concentration point estimates

(EC50, EC10, etc.), or no-and lowest-observed effects concentrations (NOECs, LOECs). These derived values are not data. Such data-poor publications essentially represent an implicit claim by the researcher to "*trust us, we know what we're doing, our interpretation of the data is the only appropriate interpretation, you don't need to see what you don't see, and besides it's our data to share as we see fit.*" Such attitudes reflect the norm in scientific publishing prior to the early 2000s, in which strict page limits and word limits precluded authors "wasting" space publishing data tables. With the provisions for electronic supplemental material beginning in the 2000s, and dedicated data repositories becoming widely available at low or no costs to authors in the 2010s, these reasons for opaque publication are no longer justified. Researchers who choose not to transparently report the actual data underlying their scientific findings may have other reasons for doing so. They may be concerned about others scooping them on their own data (McNutt 2016), although counterintuitively, publishing data may actually help establish priority and reduce scooping concerns (Laine 2017). Other less charitable reasons why researchers might resist publishing data include that they haven't devoted the needed time to organize their data in a coherent fashion that is interpretable by others, because reported results might not be able to be reconstructed from the underlying data, they are not keen to facilitate alternate statistical analyses or interpretations of their data, that they wish to publish unfalsifiable findings, or because there's simply less there than they led readers to believe (Smith and Roberts 2016).

Data sharing may still be regarded more as an imposition from science funders to be complied with rather than as a universal principle embraced by those conducting and publishing scientific research (Burwell et al. 2013; Collins and Verdier 2017; European Commission 2016; Holdren 2013; Nelson 2009; Nosek and 39 co-authors 2015). There are many pragmatic obstacles to effective data sharing, such as the expertise, extra work, and costs to researchers to organize, serve, and preserve their data in a comprehensible manner, privacy and anonymity concerns for environmental data collected from private property, about human subjects, and balancing intellectual property concerns. Some environmental science research is intended to be confidential, such as private sector economic geology, agricultural chemical product development, and innumerable other corporate research efforts which are intended to develop products and recoup research and development investments. However, in our view, researchers on such ventures cannot have it both ways, by publishing some outcomes in the peer reviewed literature, but withholding the supporting data as private. A recent corporate initiative to make available traditionally protected crop safety information is noteworthy in this regard (https://cropscience-transparency.bayer.com/).

Most environmental science journals have policies encouraging and facilitating data sharing. SETAC journals are probably typical in requiring a statement by the authors' whether and how the data underlying their analyses are available, with an admonition that authors should share upon request. A passable statement may be something as feeble as "*Contact the Corresponding Author for data availability*."

The strongest data disclosure policy for journals publishing in the environmental sciences is probably that developed for the Public Library of Science (PLoS) family of journals. "*PLoS journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception*" (PLOS 2014). Exceptions are limited to privacy or vulnerability concerns such as data on human research subjects that could not be fully anonymized, locations of archeological, fossil, or endangered species, that could be exploited or damaged, or safety and security considerations. Penalties for authors who fail to comply include rejection, or if they decline to provide data for an already published article, the editors could flag their article with a cautionary correction or even retract it (PLOS 2014). Whether PLoS's stand requiring authors to make available all data underlying their findings will lead other journals to stiffen their resolve, or whether the comparatively lax policies of competing journals will undermine PLoS and other open-science advocates remains to be seen (Davis 2016; Nosek and 39 co-authors 2015).

Implicit to such requirements is the assumption that common understandings of what constitutes "raw data" will be contextual. Often, when researchers ask for raw data, what they really want are detailed and curated data summaries. Some "data points" such as a streamflow measurement or a chemical concentration in a medium are actually derived values, and the true raw data behind a data point includes survey data, unprocessed sensor readings, spectral outputs, and such. Unless the study involves methods comparisons or forensic data audits, usually the researcher just wants the resultant derived values at a level of detail sufficient to reconstruct and further analyze the original detail.

While the notion that investigators should preserve and share underlying data is simple, the reality of doing so is much more complicated and challenging. To us, it is a priority to strongly encourage, for without data, the credibility of science cannot be evaluated. Some research has shown the willingness and ability for authors to share data declines significantly with time, and having a weak data availability policy is only marginally better than having no policy at all (Vines et al. 2014). Computer servers get replaced, directories flushed, inactive files get dumped in office moves, and investigators move on, retire, and eventually die.

Rather than mandates, one simple incentive to improving openness in reporting has been for journals to award prominent open data "badges" for articles verified as being supported by available, correct, usable, and complete data. By showing an open data badge on the issue table of contents, article web page, and including a "verified open data" statement in the bibliographic indexing metadata, articles without such badge endorsement may be seen as incomplete. Over time, this might shift the norm toward open preservation and sharing. In at least one journal, this approach appeared to markedly improve the sharing and preservation of data through linked, independent repositories (Kidwell et al. 2016).

Critical Reviews and Literature Syntheses

In ecotoxicology, published literature can roughly be broken down into two categories: original research and the review article. The original research article usually is based upon field observations, laboratory experiments, modelling, or blended approaches. Generalizing original articles through reviews and syntheses are critical parts of the ecotoxicology and most environmental science literature. Critical reviews, risk assessments, environmental quality standards, are based on syntheses of the literature, and not on individual studies. Synthesis articles have rather distinct scientific integrity problems from the original research article. Decisions must be made on how studies were located, results categorized, and a host of data manipulation and analyses decisions need to be made. These decisions and associated biases may be deliberate and clearly explained or the analyst may not even recognize that they have made a decision. In some cases we suspect analysts obscured their decisions. Systematic review methodology is now being used also for chemical assessments in which case data synthesis may be highly structured, with criteria clearly defined upfront for data inclusion and search strategies (Hobbs et al. 2005; Whaley et al. 2016). Other situations may follow the wending path of the present article: discussions among the authors "*have you seen so-and-so?*", and readings that led to other relevant material through forward and backward citing, along with by some specific subject searches. This path led to much relevant and thoughtful material across many disciplines. But it was hardly systematic or reproducible.

Literature searches from different sources can yield very different results. For example, using a 2007 original research article on population modeling of selenium toxicity to trout (Van Kirk and Hill 2007), four leading bibliographic indexing services were searched for articles citing that study. Web of Science (WoS), Elsevier's Scopus, Digital Science's Dimensions, and Google Scholar found 7, 10, 15, and 22 citing publications respectively. Scopus found all articles found by WoS, plus articles WoS missed in *Human and Ecological Risk Assessment* and *IEAM*. Google Scholar found all articles found by Scopus and WoS, plus articles in *Ecotoxicology Modeling*, *Water Resources Research*, 3 government reports, 2 books, a thesis, a conference proceeding, a duplicate, and 2 ambiguous citations. It follows from this 3-fold difference in valid citations that a critical review of published literature on a topic or a regulatory assessment could miss relevant science if the assessors relied too heavily on a single search provider.

This simple example was from the "current era" of science, which began by 1996 or so, depending on which bibliographic indexing service scholars are using. Web sites for WoS and Scopus respectively report their indexing databases are reliable from 1971 and 1996 forward.

Relying exclusively on bibliographic index searching may omit important, relevant older research.

Thus, we have the indexing bias problem in meta-analyses and assessment (that not indexed won't be retrieved), and the related problem of reviewing the secondary source but citing the original. We have seen assessments that omitted seminal research published before the current digital era, which may reflect indexing bias. Ecotoxicology syntheses often rely on variations of species-sensitivity distributions, which may provide more explanations of statistical characteristics of the datasets, data extrapolations, transformations, normalizations, than on where the data came from in the first place. We have seen micrograms and milligrams mixed up, and rankings that mistakenly commingled endpoints such as time to death in hours with effects concentrations. Some of these issues are undoubtedly related to the online availability of well curated databases such as ECETOC Aquatic Toxicity (EAT) Database from the European Center for Ecotoxicology and Toxicology of Chemicals or the U.S. Environmental Protection Agency's EcoTox databases. These compiled databases are valuable resources but reliance on secondary, compilations deprive the original authors of credit via citations. At least for publicly funded science, citations may be a way that authors demonstrate the value of their work to the scientific community, and thus build the case for further funding. Further, reliance on secondary sources is a good way to introduce or repeat inaccuracies (Rekdal 2014). We echo previous calls for better training and rigor when conducting and reporting secondary analyses of ecotoxicology and related literature. Practices from other fields, such as the Cochrane systematic review approach and guidelines for the ethical reuse of data could be adapted to the ecotoxicology practice (Duke and Porter 2013; Roberts et al. 2006; Suter and Cormier 2015a).

Environmental Chemistry

Environmental chemistry has different scientific integrity challenges than the biological side of ecotoxicology. Unlike the situation in the biological side of ecotoxicology where serious questions have been aired about the reproducibility of some of the published research (Scott 2012, 2018; Sumpter et al. 2014; Sumpter et al. 2016), analytical environmental chemistry does not appear to suffer from such problems to the same extent. The likely reason for this is that quality assurance mechanisms are routinely incorporated into analytical projects involving the measurement of environmental concentrations of chemicals, thus ensuring that the results are accurate. These include the use of high quality standards, which are widely available, the use of high specification instruments, and general guidelines proposed by national and international institutions. That combination enables recovery rates to be determined, preferably at different concentration ranges, for intra- and inter-day precision to be assessed, detection limits to be quantified, and matrix effects (interference from other substances) to be investigated. These quality assurance procedures are adopted routinely, are always checked by reviewers of analytical papers, and ensure quality is maintained.

However, the reporting and interpretation of environmental chemistry has common pitfalls, particularly in analyses from large datasets or compiled databases, and in citation practices. For example, metadata specifying fundamental details may be missing or misunderstood, such as whether concentrations of metals or other elements in water are from filtered or unfiltered samples or if they reflect the total mass of the element or only one speciation state (Sprague et al. 2017). Aquatic metals concentrations declined from mg/L levels in reports from the 1980s to µg/L or sub µg/L levels by the late 1990s. This remarkable, widespread decline was not due to better pollution controls or global geochemical change, but to improved recognition and control of ubiquitous contamination in field and laboratory sampling and analysis methods. There are ongoing debates over the most appropriate sampling and analysis methods for inorganic water quality constituents particularly for environments that are expensive and difficult to sample representatively, such as large rivers. (Horowitz 2013). Such sampling biases and analytical method differences may be substantial enough to confound analyses.

Organic environmental chemistry datasets have similar pitfalls that can confound subsequent reviews and secondary analyses. For example, Kolpin et al (2002a) published a summary of a survey of different pharmaceuticals, hormones and other organic contaminants from 139 streams. This highly influential paper showed that some organic contaminants were widespread in streams and contributed to heightened concern and research interest in their potential health and environmental risks. The paper is presently the most highly cited paper ever published in the journal *Environmental Science & Technology* (1st of 46,011 papers published, with 5,104 citations in the Scopus database as of 16 August 2018). However, reported concentrations of at least 1 of the 95 chemicals reported, 17α-ethinyl estradiol (EE2), were questioned because the median and maximum concentrations of 73 and 831 ng/L, respectively were about 10 to 1000 times higher than those from other surveys or analyses (Ericson et al. 2002; Hannah et al. 2009). Kolpin et al. (2002b) responded that upon further inspection, they had discovered that their maximum reported EE2 value of 831 ng/L was indeed incorrect owing to analytical interferences. They further explained that they had defined “median” in a peculiar way, as the median of detected values in streams, not in its usual meaning as a central tendency of all values. Because EE2 was undetectable in 94% of streams sampled, the median of detected values was skewed far above the median in all streams (<5 ng/L). However, despite the discovery of the mistakes, no correction was issued for the original publication. The Kolpin et al. (2002b) acknowledgment of the mistaken values was buried among the other 5,103 citing papers and the subtle, peculiar definition of a “median” was likely overlooked by most readers. As of August 2018, at least 50 citing papers were identified that re-reported and perpetuated the exorbitant and mistaken 831 ng/L maximum EE2 value for U.S. streams.

Thus, it is easy for authors to easily misinterpret or to perpetuate erroneous relevant values from the literature. The problem of citing unreliable maximum values would be avoided if authors simply cited extreme statistics, such as percentile concentrations (e.g., the 10th to 90th, 5th to 95th, 1st to 99th) instead of ranges (Weltje and Sumpter 2017). Whereas a single extreme value defines the range, extreme percentiles are more representative of severe conditions that organisms may actually encounter and will be more stable and are far less vulnerable to be mistaken. For instance, Santore et al. (2018, their figure 9) elegantly summarized about 29,000 paired reports from aggregated data sources of dissolved and total aluminum (Al) in freshwater. Logically, the dissolved fraction of trace metals in water can be no greater than the total, although in practice results don't always come out that way especially when the two values are close. Factors such as differences in sample digestion, differences between instruments, or slight differences in technique may introduce subtle analytical biases and impede reproducibility (Paul et al. 2016). In the Santore et al. (2018) comparison, at least 150 of the 29,000 Al pairs show the dissolved fraction is greater than the whole. While such logically impossible values should usually simply indicate that close to 100% of the total Al was present in dissolved form, some are obviously impossible values with the dissolved fraction 2-3 orders of magnitude higher than the total concentration. Should an imprudent analyst uncritically report on ranges of dissolved and total Al, they would report nonsense results. Simply backing off to the 99th or 95th percentile for a large dataset such as this one would still reflect the extremes of environmental conditions but be somewhat insulated from dubious, single values.

Advocacy

Science is the enterprise for answering questions and making predictions about the how the universe works. Science can inform issues, but science can never answer "should" questions. For example, science cannot tell societies whether they should restrict chemical uses and releases, whether natural preserves should be set aside from human exploitation, or whether biodiversity should be protected. These are among the myriad value judgements that societies must make, and while science can support societies in making these choices through predictions founded upon a body of knowledge, there are never "scientifically correct" answers to questions of human values, morals, and ethics (Snyder and Hooper-Bui 2018). Scientists are humans, and like all people, hold ethical and moral values which drive assumptions which may not be explicitly stated if even recognized. For example, the notion of "environmental protection" in the environmental toxicology field is rooted in societal norms, statues, and international agreements with goals of minimizing harm (a human concept) from activities such as extraction, manufacture, use, and disposal of chemical products. Scientists in the field develop informed opinions toward the "should" questions relating to their experiences, which leads to questions of whether and how scientists advocate for "should" questions.

The underpinnings of science are that researchers have no vested interest in the results of their observations, that they objectively record and analyze these results, and that they fairly report the outcomes in the peer-reviewed literature. Advocacy can compromise these underpinnings, at the cost of scientists' credibility (Fenn and Milton 1997). Scientists tend to be passionate about their science, which has led to controversy over the role that scientists should play in related public policy debates. While we think most scientists would agree that advocacy for science having a role in environmental policy debates is appropriate, there is likely much less agreement whether it is appropriate for scientists to advocate for particular outcomes in policy debates. If the policy debate turns on questions of science central to a scientist's particular area of study, probably no one is better positioned than that scientist to lay out the evidence for or against a particular course of action. If the scientist is regarded as a neutral and informed voice, their advice may be valued by all sides in a policy dispute (Sedlak 2016). However, if the scientist's experience or analyses leads them to the strong conviction that one policy direction is more correct and should be adopted, then they are no longer a neutral broker and have become an advocate.

When questions of science are central to adversarial adjudicated proceedings, the protagonists controlling the proceedings are often lawyers. The lawyers are expected to advocate for their client's interest; not for objective science. The lawyers retain consulting scientists as expert witnesses to support their side of the case. The lawyers' will presumably seek out scientists whose research findings and views will increase their chance of winning. In the close, intense working environment of a team preparing for a complex, science-based legal strategy, it is easy for scientists to get caught up in the enthusiasm of a "team spirit" with a loss of detached impartiality and objectivity. Scientists who begin to function as "hired guns" focused on team wins are no longer scientists but advocates (Christensen and Klauda 1988).

Policy advocacy is potentially problematic because it may compromise use of research findings in policy and management deliberations if the information is not viewed as credible by all sides (Scott et al. 2007). In some situations, advocacy is beyond reproach, such as a university scientist who uncovered a lead poisoned community water system. Simply reporting the findings to the responsible officials likely would have been ineffective, if the ineptitude or indifference of those same responsible officials contributed to the situation in the first place (Sedlak 2016). However, not all situations are so clear cut, and reasonable people who share similar motivations, skills, and agree that researchers should do the right thing may not agree on what that is. Deliberations on major environmental issues are complex and science may only be one element of the deliberations. Developing and providing technical and scientific information to inform policy deliberations in an objective and relevant way is a formidable challenge (Meyer et al. 2010; Nelson and Vucetich 2009).

Institutional constraints aside, how scientists balance these competing issues and choose when or whether to engage in advocacy is a deeply personal choice and is situational (Meyer et al. 2010; Sedlak 2016). However, just as science journals discourage comingling original research results and commentary, scientists should keep science and advocacy distinct in their publications and speaking. In particular, we argue that scientists should be watchful for stealth policy advocacy. Stealth advocacy is the use of value-laden language in scientific writing that assumes a policy preference (Lackey 2007; Pielke 2007). Rather than openly disclosing assumed values or policy preferences, biases may be unconsciously (or deliberately) cloaked through normative science. Normative science is science developed, presented, or interpreted all based on an assumed, usually unstated, preference for a particular policy or class of policy choices. This covert advocacy may be reflected in word choices, and such advocacy is not always apparent even to the advocate. For instance, value-laden words such as *stressors, impacted, degraded, improved, good,* and *poor* may be used to describe habitats or other environmental features. Less value-laden words would be *factors, exposed*, *altered, changed, increased,* or *decreased.* The use of normative science is potentially insidious because the tacit, usually unstated, preference for a particular policy or class of policy choices is not perceptibly normative to policy makers or even to many scientists (Lackey 2007).

Criticisms of normative science can be excessive, as taken literally, the entire discipline of conservation biology could be considered too normative. Similarly, the mission statement of SETAC "*to support the development of principles and practices for protection, enhancement and management of sustainable environmental quality and ecosystem integrity*" could be too much for some. Science is normative, with topics and study questions influenced by normative treaties and regulations. Areas of study or techniques once considered appropriate areas of science inquiry such as craniometry, eugenics, or experimentation on human subjects without informed consent are no longer considered to be within the norms of ethical science. Within environmental toxicology, pressure to reduce the use of animal testing might be an example of normative science.

Our point is not to argue for or against scientists engaging in overt policy advocacy, which is a personal decision, but for clarity and transparency. Just as original results, opinion, judgements and speculation should not be blended in a scientific paper, science and advocacy need some separation (Scott and Rachlow 2011). Covert advocacy is a form of bias. Environmental scientists should clearly differentiate between research findings and policy advocacy based upon those findings.

Weaponizing scientific integrity and transparency

We recognize that "scientific integrity" discussions can easily be diminished to going down the path carved by "sound science" strategic initiatives, which often boiled down to campaigns to call "*my science good science and your science junk*" (Doremus 2007; Kapustka 2016; McGarity 2003; Oreskes and Conway 2011). The goal may be to recast policy, ideological, or economic disputes as doubt or created conflicts in science. In countries with a tort-based, adversarial legal system for resolving injuries or damages, science-based information becomes just another tool for dueling experts, who often have primary responsibility for advocating for the interests of their client (Wagner 2005). Research integrity policies or requirements for data transparency can be used as weapons to bury public university or government scientists with vexatious, intrusive, and costly demands for records such as raw laboratory notebooks, instrument calibration records, emails between coauthors, working drafts, and peer comments and responses. Such demands can be effective tools for interfering with the work of public-sector scientists, including academics in public institutions (Folta 2015; Halpern and Mann 2015; Kloor 2015; Kollipara 2015; Lewandowsky and Bishop 2016), or academics in private institutions but who receive research support from public sources (Hey and Chalmers 2010; Shrader-Frechette 2012). For example, Deborah Swackhamer, an environmental chemist at the University of Minnesota, was targeted under state sunshine laws with legal demands for raw unpublished data, class notes, purchase records, telephone records, and more from a 15-year period. Ironically, the identity of those seeking the information was shielded behind the law firm communicating the demands (Halpern 2015). Some scientists have learned to use transparency laws against their peers in the highly competitive arena of grant funding. Through freedom of information demands for competing grant proposals, scientists have been able to obtain details on competitors' new research direction, preliminary results, and cost structure. For those targeted scientists, such information gathering may be seen as research espionage under the rubric of transparency (Carey and Woodward 2017).

The sunshine laws enacted in many jurisdictions were intended to illuminate the business of government officials and were doubtfully intended by their crafters to sweep up university professors. Nevertheless, some see scientists are fair targets of such tactics, as inspections of their erstwhile private communications have uncovered undisclosed conflicts of interest or bias (e.g., Russell et al. 2010). Privately funded research is generally shielded from such practices (Brain et al. 2016; Wagner and Michaels 2004). Researchers at private institutions may however be subject to baseless litigation to intimidate scientists and deter others by inflicting long and costly legal processes, disruption, and threats of personal financial liability. Such harassing lawsuits have been employed often enough to get a name, SLAPP suits for Strategic Litigation Against Public Participation (Johnson 2007; Nature Medicine Editors 2017; Robbins 2017). While legal, such strategies represent detrimental practices cloaked in the vernacular of transparent science (Johnson 2007; Levy and Johns 2016; McGarity and Wagner 2008; Wagner and Michaels 2004).

Education as the way forward

It is one thing to realize that there is a problem, but quite another to find an effective solution to that problem. The unethical behaviors discussed above are primarily a consequence of the perverse incentives under which scientists currently operate. These incentives are publications and grants. In the case of publications, the number of these seems to be much more important than their quality. This is probably because assessing the quality of a scientific article is not easy; there is no established, widely accepted way of doing this. The 'status' of the journal in which an article is published, which is most often taken to be the impact factor of that journal, also is considered an important factor. Hence scientists strive to get their papers published in journals with high impact factors; and may act unethically to do so. In the case of grants, the more, and bigger, they are, the better, as far as institutions are concerned. These incentives, particularly those concerning publications ('the more the merrier'), are probably responsible for the many lapses in integrity currently obvious and prevalent in ecotoxicology. Moving from the present situation to a significantly more ethical one in which integrity is central to any endeavor in ecotoxicology will not be easy to accomplish, nor will it be achieved quickly.

Education of ecotoxicologists, both young and old, is the only way forward towards better integrity in our discipline. That education can be delivered in a variety of ways, the two most obvious and practical being (1) the publication of articles in journals in which integrity and ethics are discussed, and (2) courses run by scientific societies such as SETAC. This article is an example of a very direct attempt at highlighting integrity issues in our field, with the hope that by making ecotoxicologists aware of these unethical practices they will change their behavior and act more ethically. Other, less direct, attempts have involved the publication of papers covering suggestions for how to improve the quality of ecotoxicology research, from the planning stage (Harris et al. 2014) to the publication stage (e.g., Hanson et al. 2017; Moermond et al. 2016). However, it seems unlikely that published papers alone will have a significant influence on the quality of ecotoxicology research because few scientists will be aware of them, and even fewer will read them carefully and subsequently act on the advice in them.

Although there has been some public discussion about what training and skills the ecotoxicologists of the future will require (Harris et al. 2017), this crucial aspect of producing better ecotoxicologists, capable of doing better, and hence more useful, research has rarely been addressed. Yet there are undoubtedly things that could be done to better educate the ecotoxicologists of the future. A radical proposal would be to require aspiring ecotoxicologists to pass examinations before they are allowed to practice ecotoxicology, either as researchers or regulators. Many professions do insist that its practitioners pass examinations before they are allowed to practice: doctors, dentists, accountants and lawyers are obvious examples. This strategy ensures that practitioners are adequately trained. As a first step towards the goal of ensuring all ecotoxicologists are appropriately trained, specific courses could be introduced, and attendance become mandatory. Courses on topics such as experimental design, statistical analysis, data presentation and how to write a scientific paper could be designed easily. In fact, many research organizations and some industries already run 'in house' courses on these topics. It would be equally feasible to design a course on integrity in ecotoxicology research. In fact, as the issue of integrity (or, more accurately, the lack of it) has gained in prominence in the last few years, some organizations have responded by running training courses for their young scientists on integrity and ethical behavior in research. SETAC could offer such training courses and does so to a limited extent already. Another possibility would be for consultancy companies to develop and run these training courses for clients, who could be universities, research organizations or industrial companies. Consultancy companies that specialized exclusively in providing training could be established; this has happened already in many other professions.

In summary, identifying that there are problems with the way ecotoxicologists are trained currently about integrity issues in their discipline is only the first step. Much better education of ecotoxicologists (both those starting their careers and those well-established already) is desperately needed. Such education will need to be provided in a range of formats, to maximize its chances of succeeding. But if ecotoxicology is to be taken seriously as a profession, change and improve it must. The environment cannot be protected by poor quality ecotoxicology.

Promoting scientific integrity in environmental toxicology

Scientific integrity is harnessed by high quality environmental research characterized by rigor, relevance, reproducibility, and objectivity. Our review suggested several conclusions, tangible actions and less tangible directions that professional societies such as SETAC could do to encourage scientists, their supporting institutions, and science journals to maintain and improve a culture of science integrity. Scientific integrity is reinforced through full transparency exemplified by full disclosures of potential conflicting and competing interests that could contribute to bias, and by making all data and observations readily accessible. Specifically:

1. Scientific integrity in ecotoxicology and the environmental sciences cannot be ensured by impeccable policies or checklists. It is an attitude to be embraced, maintained, and enforced through the support, guidance and approval of one's peers through a community of practices.
2. Reliability, rigor, relevance and reproducibility of science are more important than novel advances, if those advances neglect these "four Rs."
3. Increased attention to a culture of quality management training and transparency could improve the confidence in published findings.
4. Studies that are not supported by primary data released through data repositories or detailed supporting information are not fully credible.
5. Journal publishers and editors could strongly encourage the complete presentation of supporting data, with prominent labeling on the journal and article front matter indicating

whether data are available. They should caution authors at the outset that the inability to produce data upon request could be cause for retraction.

6. One practical step investigators can take toward improving reproducibility of experiments would be to produce detailed video illustration of their methods.
7. As a community, be aware of and disclose potential conflicting or competing interests that could contribute to, or be perceived as, bias and not tolerate extreme conflicts or bias.
8. Discourage judging science by its funder; rather, open-minded skepticism is applicable when the funder has a stake in the outcome of a study.
9. Scientists, like all people, have moral and ethical assumptions, based upon their values. These should not be intermixed with their interpretations and reporting of science. If scientists' values lead them to cross the lines from analysis to advocacy, they need to be particularly careful about distinguishing between science, values, assumptions, and opinion.
10. Professional societies such as SETAC have an important role in fostering respectful evidence-based dialog, in meetings and correspondence on published works.
11. Professional societies such as SETAC could support a standing training seminar on principles of scientific integrity, the transparent conduct of science and best practices for peer review in conjunction with its annual meetings.
12. Professional societies such as SETAC have a valuable role in facilitating balanced, expert reviews of controversial science topics, such as has been done with their Pellston Workshops series and resulting publications.

Supplemental Information

SI 1. Peer review process

Acknowledgements

We thank the four anonymous peer reviewers and the two institutional reviewers for their constructive criticisms and insights. This commentary and perspectives followed discussions by the Society of Environmental Toxicology and Chemistry (SETAC) ad hoc subcommittee on Scientific Integrity (Fairbrother 2016). Author roles: CAM wrote most of the paper, with contributions by JPS and AF. All other authors participated in discussions, reviewed and edited the drafts, supported the recommendations, and approved of its publication. Specific required disclaimers: The views expressed in this article are those of the authors and do not necessarily represent the views or policies of SETAC, the U.S. Environmental Protection Agency, or the U.S. Fish and Wildlife Service. Affiliations are given for identification, but do not necessarily imply endorsement. Author affiliations reflect a variety of academic, business, and governmental employment, and all authors have competing interests from our previous experiences and work environments, but hopefully these were somewhat balanced out. No specific funding was provided for the writing of this manuscript and no financial conflicts of interest were declared by any author.

References

AAAS. 2000. The Role and Activities of Scientific Societies in Promoting Research Integrity. April 10-11, 2000. American Association for the Advancement of Science (AAAS) and the 24 pp. https://www.aaas.org/page/scientific-integrity.

Alberts, B. 2012. Editor's note [A bacterium that can grow using arsenic instead of phosphorus]. *Science*. 334(6034): 1149. https://doi.org/10.1126/science.1208877

Anderson, K., 2017, Trust Falls — Are We In a New Phase of Corporate Research? (August 15, 2017), *Scholarly Kitchen* at https://scholarlykitchen.sspnet.org/2017/08/15/trust-falls-new-phase-corporate-research/.

ARC. 2007. Australian Code for the Responsible Conduct of Research. Australian Research Council (ARC). http://www.arc.gov.au/research-integrity.

Arciszewski, T.J. and K.R. Munkittrick. 2015. Development of an adaptive monitoring framework for long-term programs: An example using indicators of fish health. *Integrated Environmental Assessment and Management*. 11(4): 701-718. https://doi.org/10.1002/ieam.1636

Atkinson, R.L. and I. Macdonald. 2010. White hat bias: The need for authors to have the spin stop with them. *International Journal of Obesity*. 34(1): 83-83. https://doi.org/10.1038/ijo.2009.269

Augspurger, T. 2014. SETAC Hosts Successful TSCA [Toxic Substances Control Act] Risk Assessment Science Seminar for Congressional Staff (6 November 2014). *SETAC Globe*. 15(11)

Baba, A., D.M. Cook, T.O. McGarity, and L.A. Bero. 2005. Legislating "Sound Science": The Role of the Tobacco Industry. *American Journal of Public Health*. 95(S1): S20-S27. https://doi.org/10.2105/AJPH.2004.050963

Baethge, C. 2013. The effect of a conflict of interest disclosure form using closed questions on the number of positive conflicts of interest declared – a controlled study. *PeerJ*. 1: e128. https://doi.org/10.7717/peerj.128

Baker, M. 2016a. 1,500 scientists lift the lid on reproducibility. *Nature*. Nature. https://doi.org/10.1038/533452a

Baker, M. 2016b. How quality control could save your science. *Nature*. 529: 456–458. https://doi.org/10.1038/529456a

Bernhardt, E.S., E.J. Rosi, and M.O. Gessner. 2017. Synthetic chemicals as agents of global change. *Frontiers in Ecology and the Environment*. 15(2): 84-90. https://doi.org/10.1002/fee.1450

Bero, L., A. Anglemyer, H. Vesterinen, and D. Krauth. 2016. The relationship between study sponsorship, risks of bias, and research outcomes in atrazine exposure studies conducted in non-human animals: Systematic review and meta-analysis. *Environment International*. 92-93: 597-604. https://doi.org/10.1016/j.envint.2015.10.011

Bes-Rastrollo, M., M.B. Schulze, M. Ruiz-Canela, and M.A. Martinez-Gonzalez. 2014. Financial Conflicts of Interest and Reporting Bias Regarding the Association between Sugar-Sweetened Beverages and Weight Gain: A Systematic Review of Systematic Reviews. *PLOS Medicine*. 10(12): e1001578. https://doi.org/10.1371/journal.pmed.1001578

Blumenstyk, G. 2007. When Research Criticizes an Industry. *Chronicle of Higher Education*. 54(4): A21

Boden, L.I. and D. Ozonoff. 2008. Litigation-Generated Science: Why Should We Care? *Environmental Health Perspectives*. 116(1): 117-122

Boehm, P.D. and T.C. Ginn. 2013. The Science of Natural Resource Damage Assessments. *Environmental Claims Journal*. 25(3): 185-225. https://doi.org/10.1080/10406026.2013.785910

Bohannon, J. 2013. Who's Afraid of Peer Review? *Science*. 342(6154): 60-65. https://doi.org/10.1126/science.342.6154.60

Boone, M.D., C.A. Bishop, L.A. Boswell, R.D. Brodman, J. Burger, C. Davidson, M. Gochfeld, J.T. Hoverman, L.A. Neuman-Lee, R.A. Relyea, J.R. Rohr, C. Salice, R.D. Semlitsch, D. Sparling, and S. Weir. 2014. Pesticide Regulation amid the Influence of Industry. *BioScience*. 64(10): 917-922. https://doi.org/10.1093/biosci/biu138

Bornstein-Forst, S.M. 2017. Establishing Good Laboratory Practice at Small Colleges and Universities. *Journal of Microbiology & Biology Education*. 18(1): 18.1.10. https://doi.org/10.1128/jmbe.v18i1.1222

Borrell, B. 2009. National Academy as *National Enquirer*? PNAS Publishes Theory That Caterpillars Originated from Interspecies Sex. *Scientific American*. August 24, 2009

Brain, R., J. Stavely, and L. Ortego. 2016. In Response: Resolving the perception of bias in a discipline founded on objectivity—A perspective from industry. *Environmental Toxicology and Chemistry*. 35(5): 1070-1072. https://doi.org/10.1002/etc.3357

Brix, K.V., D.K. DeForest, and W.J. Adams. 2011. The sensitivity of aquatic insects to divalent metals: A comparative analysis of laboratory and field data. *Science of the Total Environment*. 409(20): 4187-4197. https://doi.org/10.1016/j.scitotenv.2011.06.061

Brix, K.V., A.J. Esbaugh, K.M. Munley, and M. Grosell. 2012. Investigations into the mechanism of lead toxicity to the freshwater pulmonate snail, *Lymnaea stagnalis*. *Aquatic Toxicology*. 106–107(0): 147-156. https://doi.org/10.1016/j.aquatox.2011.11.007

Bui, C., R. Parrish, M. Meredith, and J. Giddings, editors. 2004. *A History of SETAC Part 1: The Beginning*. SETAC Press, Pensacola, FL, 36 pp

Burton, G.A., R. Di Giulio, D. Costello, and J.R. Rohr. 2017. Slipping through the Cracks: Why is the U.S. Environmental Protection Agency Not Funding Extramural Research on Chemicals in Our Environment? *Environmental Science & Technology*. 51(2): 755-756. https://doi.org/10.1021/acs.est.6b05877

Burton, G.A. and C.H. Ward. 2012. Editorial preface [letters to the editor on the Exxon Valdez oil spill and pink salmon]. *Environmental Toxicology and Chemistry*. 31(3): 469-469. https://doi.org/10.1002/etc.1741

Burwell, S.M., S. VanRoekel, T. Park, and D.J. Mancini. 2013. Open Data Policy-Managing Information as an Asset. U.S. Office of Management and Budget M-13-13. https://obamawhitehouse.archives.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf [Accessed 15 September 2018].

Buys, D.J., A.R. Stojak, W. Stiteler, and T.F. Baker. 2015. Ecological risk assessment for residual coal fly ash at Watts Bar Reservoir, Tennessee: Limited alteration of riverine-reservoir benthic invertebrate community following dredging of ash-contaminated sediment. *Integrated Environmental Assessment and Management*. 11(1): 43-55. https://doi.org/10.1002/ieam.1577

Calfee, R.D., H.J. Puglis, E.E. Little, W.G. Brumbaugh, and C.A. Mebane. 2016. Quantifying Fish Swimming Behavior in Response to Acute Exposure of Aqueous Copper Using Computer Assisted Video and Digital Image Analysis. *Journal of Visualized Experiments*. (108): e53477. https://doi.org/doi:10.3791/53477

Callaway, E. 2015. Faked peer reviews prompt 64 retractions (18 August 2015). *Nature*. http://dx.doi.org/10.1038/nature.2015.18202. https://doi.org/10.1038/nature.2015.18202

Campbell, T. and J. Friesen. 2015. Why people "fly from facts": Research shows the appeal of untestable beliefs and how they lead to a polarized society (March 3, 2015). *Scientific American*.

Carey, T.L. and A. Woodward, 2017, These Scientists Got To See Their Competitors' Research Through Public Records Requests (September 2, 2017), *BuzzFeed News* at https://www.buzzfeednews.com/article/teresalcarey/when-scientists-foia.

Christensen, S.W. and R.J. Klauda. 1988. Two scientists in the courtroom: what they didn't teach us in graduate school. *American Fisheries Society Monograph*. 4: 307-315

Collins, S.L. and J.M. Verdier. 2017. The Coming Era of Open Data. *BioScience*. 67(3): 191-192. https://doi.org/10.1093/biosci/bix023

Cope, M.B. and D.B. Allison. 2009. White hat bias: examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. *International Journal of Obesity*. 34(1): 84-88

Couzin-Frankel, J. 2017. Firing of veteran NIH scientist prompts protests over publication ban (February 22, 2017). *Science*. https://doi.org/10.1126/science.aal0808

Cranor, C.F. 2008. The Tobacco Strategy Entrenched. *Science*. 321(5894): 1296-1297. https://doi.org/10.1126/science.1162339

Curfman, G.D., S. Morrissey, and J.M. Drazen. 2005. Expression of Concern: Bombardier et al., "Comparison of Upper Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients with Rheumatoid Arthritis," N Engl J Med 2000;343:1520-8. *New England Journal of Medicine*. 353(26): 2813-2814. https://doi.org/doi:10.1056/NEJMe058314

Davis, P., 2016, *Scientific Reports* On Track To Become Largest Journal In The World (August 23, 2016), *Scholarly Kitchen* at https://scholarlykitchen.sspnet.org/2016/08/23/scientific-reports-on-track-to-become-largest-journal-in-the-world/.

DeForest, D.K. and E.J. Van Genderen. 2012. Application of U.S. EPA guidelines in a bioavailability-based assessment of ambient water quality criteria for zinc in freshwater. *Environmental Toxicology and Chemistry*. 31(6): 1264–1272. https://doi.org/10.1002/etc.1810

Descamps, H. 2008. Natural resource damage assessment (NRDA) under the European Directive on Environmental Liability: a comparative legal point of view. *Océanis*. 32(3/4): 439-462

Dixon, P.M. and J.H.K. Pechmann. 2005. A statistical test to show negligible trend. *Ecology*. 86(7): 1751–1756. https://doi.org/10.1890/04-1343

Doremus, H. 2007. Scientific and Political Integrity in Environmental Policy. *Texas Law Journal*. 86: 1601-1653

Douglas, H. 2015. Politics and Science: Untangling Values, Ideologies, and Reasons. *The ANNALS of the American Academy of Political and Social Science*. 658: 296-306 https://doi.org/10.1177/0002716214557237

Douglas, H.E. 2014. Scientific Integrity in a politicized world. Pages 254-268 *in* P. Schroeder-Heister, G. Heinzmann, W. Hodges, and P. E. Bour, editors. *Logic, Methodology and Philosophy of Science. Proceedings of the Fourteenth International Congress (Nancy, France)*.

Duke, C.S. and J.H. Porter. 2013. The ethics of data sharing and reuse in biology. *BioScience*. 63(6): 483-489. https://doi.org/10.1525/bio.2013.63.6.10

Edwards, A. 2016. Reproducibility: Team up with industry. *Nature*. 531(301): 299-301. https://doi.org/10.1038/531299a

Edwards, M.A. and S. Roy. 2016. Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*. 34(1): 51-61. https://doi.org/10.1089/ees.2016.0223

Elliott, E.D. 1988. The Future of Toxic Torts: Of Chemophobia, Risk as a Compensable Injury and Hybrid Compensation Systems. *Houston Law Review*. 25: 781-799

Elliott, K.C. 2014. Financial conflicts of interest and criteria for research credibility. *Erkenntnis*. 79(5): 917-937. https://doi.org/10.1007/s10670-013-9536-2

Elliott, K.C. 2016. Standardized Study Designs, Value Judgments, and Financial Conflicts of Interest in Research. *Perspectives on Science*. 24(5): 529-551. https://doi.org/10.1162/POSC_a_00222

Enserink, M. 2014. Sabotaged Scientist Sues Yale and Her Lab Chief. *Science*. 343(6175): 1065-1066. https://doi.org/10.1126/science.343.6175.1065

Enserink, M. 2017. A groundbreaking study on the dangers of 'microplastics' may be unraveling (March 21, 2017). *Science*. https://doi.org/10.1126/science.aal0939

Ericson, J.F., R. Laenge, and D.E. Sullivan. 2002. Comment on "Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999−2000: A National Reconnaissance". *Environmental Science & Technology*. 36(18): 4005-4006. https://doi.org/10.1021/es0200903

European Commission, 2016, Open Access to scientific information | Digital Single Market (October 28, 2016), at https://ec.europa.eu/digital-single-market/en/open-access-scientific-information.

Fairbrother, A. 2016. New Subcommittee on Scientific Integrity (18 February 2016). *SETAC Globe*. 17(2)

Fairbrother, A. and R.S. Bennett. 1999. Ecological Risk Assessment and the Precautionary Principle. *Human and Ecological Risk Assessment: An International Journal*. 5(5): 943-949. https://doi.org/10.1080/10807039991289220

Fanelli, D. 2009. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*. 4(5): e5738. https://doi.org/10.1371/journal.pone.0005738

Fanelli, D. 2012. Negative results are disappearing from most disciplines and countries. *Scientometrics*. 90(3): 891-904. https://doi.org/10.1007/s11192-011-0494-7

Farris, J.L. and J.H.V. Hassel, editors. 2006. *Freshwater Bivalve Ecotoxicology*. SETAC and CRC Press, 408 pp

Fellner, C. 2018. Toxic Secrets: Professor 'bragged about burying bad science' on 3M chemicals (18 June 2018). Sydney Morning Herald.

Fenn, D.B. and N.M. Milton. 1997. Advocacy and the Scientist. *Fisheries*. 22(11): 4

Flamini, F., N. Hanley, and W.D. Shaw. 2004. Settlement versus Litigation in Environmental Damage Cases. Working Paper. Business School - Economics, University of Glasgow. http://EconPapers.repec.org/RePEc:gla:glaewp:2003_4.

Folta, K.M., 2015, Transparency Weaponized Against Scientists (August 16, 2015): Science 2.0, at http://www.science20.com/kevin_folta/transparency_weaponized_against_scientists-156873.

Fraser, H., T.H. Parker, S. Nakagawa, A. Barnett, and F. Fidler. 2018. Questionable Research Practices in Ecology and Evolution. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0200303

Gardner, M. 1989. Water with memory? The dilution affair. *The Skeptical Inquirer*. 13: 132-141

Geist, D.R., C.S. Abernethy, K.D. Hand, V.I. Cullinan, J.A. Chandler, and P.A. Groves. 2006. Survival, development, and growth of fall Chinook salmon embryos, alevin, and fry exposed to a variable thermal and dissolved oxygen regime. *Transactions of the American Fisheries Society*. 135(6): 1462–1477. https://doi.org/10.1577/T05-294.1

Ghorayshi, A. 2016. "He thinks he's untouchable"-Sexual Harassment Case Exposes Renowned Ebola Scientist (June 29, 2016). BuzzFeedNews. https://www.buzzfeed.com/azeenghorayshi/michael-katze-investigation.

Gibbons, A. 2014. Sexual harassment is common in scientific fieldwork (July 16, 2014). *Science*.

Glanz, J. and A. Armendariz, 2017, Years of Ethics Charges, but Star Cancer Researcher Gets a Pass (March 8, 2017): New York Times, at https://www.nytimes.com/2017/03/08/science/cancer-carlo-croce.html.

Gleick, P.H. and 252 co-authors. 2010. Climate Change and the Integrity of Science. *Science*. 328(5979): 689-690. https://doi.org/10.1126/science.328.5979.689

Goldsmith, B.J., T.K. Waikem, and T. Franey. 2014. Environmental Damage Liability Regimes Concerning Oil Spills - A Global Review and Comparison. *International Oil Spill Conference Proceedings*. 2014(1): 2172-2192. https://doi.org/10.7901/2169-3358-2014.1.2172

Goodstein, D. 1995. Conduct and Misconduct in Science. *Annals of the New York Academy of Sciences*. 775(1): 31-38. https://doi.org/10.1111/j.1749-6632.1996.tb23124.x

Gustavson, K.E., L.W. Barnthouse, C.L. Brierley, E.H. Clark, II, and C.H. Ward. 2007. Superfund and mining megasites. *Environmental Science and Technology*. 41(8): 2667-2672. https://doi.org/10.1021/es0725091

Halpern, M. 2015. Freedom to Bully: How Laws Intended to Free Information Are Used to Harass Researchers (2015). Union of Concerned Scientists. 20 pp. https://www.ucsusa.org/center-science-and-democracy/protecting-scientists-harassment/freedom-bully-how-laws.

Halpern, M. and M. Mann. 2015. Transparency versus harassment. *Science*. 348(6234): 479-479. https://doi.org/10.1126/science.aac4245

Hannah, R., V.J. D'Aco, P.D. Anderson, M.E. Buzby, D.J. Caldwell, V.L. Cunningham, J.F. Ericson, A.C. Johnson, N.J. Parke, J.H. Samuelian, and J.P. Sumpter. 2009. Exposure assessment of 17α-ethinylestradiol in surface waters of the United States and Europe. *Environmental Toxicology and Chemistry*. 28(12): 2725-2732. https://doi.org/10.1897/08-622.1

Hanson, M.L., L.E. Deeth, and R.S. Prosser. 2018. Evidence of citation bias in the pesticide ecotoxicology literature. *Ecotoxicology*. https://doi.org/10.1007/s10646-018-1918-4

Hanson, M.L., B.A. Wolff, J.W. Green, M. Kivi, G.H. Panter, M.S.J. Warne, M. Ågerstrand, and J.P. Sumpter. 2017. How we can make ecotoxicology more valuable to environmental protection. *Science of the Total Environment*. 578: 228-235. https://doi.org/10.1016/j.scitotenv.2016.07.160

Harris, C.A., A.P. Scott, A.C. Johnson, G.H. Panter, D. Sheahan, M. Roberts, and J.P. Sumpter. 2014. Principles of Sound Ecotoxicology. *Environmental Science & Technology*. 48(6): 3100-3111. https://doi.org/10.1021/es4047507

Harris, C.A. and J.P. Sumpter. 2015. Could the Quality of Published Ecotoxicological Research Be Better? *Environmental Science & Technology*. 49(16): 9495-9496. https://doi.org/10.1021/acs.est.5b01465

Harris, D. and M. Patrick, 2011, Is 'Big Food's' Big Money Influencing the Science of Nutrition? (June 21,2011): ABC News, at http://abcnews.go.com/US/big-food-money-accused-influencing-science/story?id=13845186.

Harris, M.J., D.B. Huggett, J.P. Staveley, and J.P. Sumpter. 2017. What training and skills will the ecotoxicologists of the future require? *Integrated Environmental Assessment and Management*. 13(4): 580-584. https://doi.org/10.1002/ieam.1877

Hayes, T.B. 2004. There Is No Denying This: Defusing the Confusion about Atrazine. *BioScience*. 54(12): 1138-1149. https://doi.org/10.1641/0006-3568(2004)054[1138:TINDTD]2.0.CO;2

Henderson, D. and K. Thomson. 2017. Why Should Scientific Results Be Reproducible? (NOVA, WGBH, 15 minute video, released 19 January 2017). WGBH, NOVA. http://www.pbs.org/wgbh/nova/next/body/reproducibility-explainer/.

Hey, E. and I. Chalmers. 2010. Mis-investigating alleged research misconduct can cause widespread, unpredictable damage. *Journal of the Royal Society of Medicine*. 103(4): 133-138. https://doi.org/10.1258/jrsm.2010.09k045

Hobbs, D.A., M.S.J. Warne, and S.J. Markich. 2005. Evaluation of criteria used to assess the quality of aquatic toxicity data. *Integrated Environmental Assessment and Management*. 1(3): 174–180. https://doi.org/10.1897/2004-003R.1

Holdren, J.P. 2013. Increasing Access to the Results of Federally Funded Scientific Research. Memorandum for the heads of executive departments and agencies (February 22, 2013). Executive Office of the President, Office of Science and Technology Policy. , Washington, D.C. 6 pp.

Hopkin, M. 2006. Ecology: Caught between shores. *Nature*. 440(7081): 144-145. https://doi.org/10.1038/440144a

Horowitz, A.J. 2013. A Review of Selected Inorganic Surface Water Quality-Monitoring Practices: Are We Really Measuring What We Think, and If So, Are We Doing It Right? *Environmental Science & Technology*. 47(6): 2471-2486. https://doi.org/10.1021/es304058q

Huff, D. and I. Geis. 1954. How to lie with statistics. W.W. Norton, New York and London. (1993 reissue). 142 pp

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*. 54(2): 187-211. https://doi.org/10.2307/1942661

Hutchings, J.A. 1997. Is scientific inquiry incompatible with government information control? *Canadian Journal of Fisheries and Aquatic Sciences*. 54(5): 1198-1210. https://doi.org/10.1139/f97-051

ICMJE. 2016. ICMJE Form for Disclosure of Potential Conflicts of Interest. International Committee of Medical Journal Editors. 3 pp. http://icmje.org/conflicts-of-interest/.

Ioannidis, J.P.A. 2005. Why Most Published Research Findings Are False. *PLoS Med*. 2(8): e124. https://doi.org/10.1371/journal.pmed.0020124

Johnson, A.C., R.L. Donnachie, J.P. Sumpter, M.D. Jürgens, C. Moeckel, and M.G. Pereira. 2017. An alternative approach to risk rank chemicals on the threat they pose to the aquatic environment. *Science of the Total Environment*. 599-600: 1372-1381. https://doi.org/10.1016/j.scitotenv.2017.05.039

Johnson, A.C. and J.P. Sumpter. 2016. Are we going about chemical risk assessment for the aquatic environment the wrong way? *Environmental Toxicology and Chemistry*. 35(7): 1609-1616. https://doi.org/10.1002/etc.3441

Johnson, B.L. 2007. Intimidation of Scientists: A HERA Experience. *Human and Ecological Risk Assessment: An International Journal*. 13(3): 475-477. https://doi.org/10.1080/10807030701340870

Johnson, D.R. and E.H. Ecklund. 2016. Ethical ambiguity in science. *Science and Engineering Ethics*. 22(4): 989-1005. https://doi.org/10.1007/s11948-015-9682-9

Kapustka, L. 2016. Words matter. *Integrated Environmental Assessment and Management*. 12(3): 592-593. https://doi.org/10.1002/ieam.1767

Keith, R., 2015, 13th retraction issued for Jesús Ángel Lemus (October 9, 2015), *Retraction Watch* at http://retractionwatch.com/2015/10/09/13th-retraction-issued-for-jesus-angel-lemus/.

Kidwell, M.C., L.B. Lazarević, E. Baranski, T.E. Hardwicke, S. Piechowski, L.-S. Falkenberg, C. Kennett, A. Slowik, C. Sonnleitner, C. Hess-Holden, T.M. Errington, S. Fiedler, and B.A. Nosek. 2016. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLOS Biology*. 14(5): e1002456. https://doi.org/10.1371/journal.pbio.1002456

Kloor, K. 2015. Agricultural researchers rattled by demands for documents from group opposed to GM foods. *Science Insider*. . https://doi.org/10.1126/science.aaa7846

Kollipara, P. 2015. Open records laws becoming vehicle for harassing academic researchers, report warns (Feb. 13, 2015). *Science*. https://doi.org/10.1126/science.aaa7856

Kolpin, D.W., E.T. Furlong, M.T. Meyer, E.M. Thurman, S.D. Zaugg, L.B. Barber, and H.T. Buxton. 2002a. Pharmaceuticals, hormones, and other organic wastewater contaminants in U.S. streams, 1999-2000: a national reconnaissance. *Environmental Science and Technology*. 36(6): 1202-1211. https://doi.org/10.1021/es011055j

Kolpin, D.W., E.T. Furlong, M.T. Meyer, E.M. Thurman, S.D. Zaugg, L.B. Barber, and H.T. Buxton. 2002b. Response to Comment on "Pharmaceuticals, Hormones, and Other Organic Wastewater Contaminants in U.S. Streams, 1999−2000: A National Reconnaissance". *Environmental Science & Technology*. 36(18): 4007-4008. https://doi.org/10.1021/es020136s

Krimsky, S. 2003. Science in the Private Interest: Has the Lure of Profits Corrupted Biomedical Research? Bowman and Littlefield Publishers. 265 pp

Krimsky, S. 2005. The Funding Effect in Science and Its Implications for the Judiciary. *Journal of Law and Policy*. 13(1): 43-68

Krimsky, S. 2007. When Conflict-of-Interest Is a Factor in Scientific Misconduct. *Medicine and Law*. 26(3): 447-463

Krimsky, S. 2013. Do financial conflicts of interest bias research?: An inquiry into the "funding effect" hypothesis. *Science, Technology, & Human Values*. 38(4): 566-587. https://doi.org/10.1177/0162243912456271

Krumholz, H.M., J.S. Ross, A.H. Presler, and D.S. Egilman. 2007. What have we learnt from Vioxx? *BMJ : British Medical Journal*. 334(7585): 120-123. https://doi.org/10.1136/bmj.39024.487720.68

Lackey, R.T. 2001. Values, Policy, and Ecosystem Health. *BioScience*. 51(6): 437-443. https://doi.org/10.1641/0006-3568(2001)051[0437:VPAEH]2.0.CO;2

Lackey, R.T. 2007. Science, scientists, and policy advocacy. *Conservation Biology*. 21(1): 12-17. https://doi.org/10.1111/j.1523-1739.2006.00639.x

Laine, H. 2017. Afraid of Scooping; Case Study on Researcher Strategies against Fear of Scooping in the Context of Open Science. *Data Science Journal*. 16: 29. https://doi.org/10.5334/dsj-2017-029

Levy, K.E. and D.M. Johns. 2016. When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society*. 3(1). https://doi.org/10.1177/2053951715621568

Lewandowsky, S. and D. Bishop. 2016. Research integrity: Don't let transparency damage science. *Nature*. 529: 459-461. https://doi.org/10.1038/529459a

Lexchin, J., L.A. Bero, B. Djulbegovic, and O. Clark. 2003. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 326(7400): 1167-1170. https://doi.org/10.1136/bmj.326.7400.1167

Lindenmayer, D.B. and G.E. Likens. 2010. The science and application of ecological monitoring. *Biological Conservation*. 143(6): 1317-1328. https://doi.org/10.1016/j.biocon.2010.02.013

Macleod, M. 2014. Some salt with your statin, Professor? *PLOS Biology*. 12(1): e1001768. https://doi.org/10.1371/journal.pbio.1001768

Mandrioli, D., C.E. Kearns, and L.A. Bero. 2016. Relationship between Research Outcomes and Risk of Bias, Study Sponsorship, and Author Financial Conflicts of Interest in Reviews of the Effects of Artificially Sweetened Beverages on Weight Outcomes: A Systematic Review of Reviews. *PLoS ONE*. 11(9): e0162198. https://doi.org/10.1371/journal.pone.0162198

Marcus, A. and I. Oransky. 2016. CRISPR controversy reveals how badly journals handle conflicts of interest (January 21, 2016). *STAT*. https://www.statnews.com/2016/01/21/crispr-conflicts-of-interest/

Marshall, E. 1983. The murky world of toxicity testing. *Science*. 220(4602): 1130-1132. https://doi.org/10.1126/science.6857237

Martinson, B.C., M.S. Anderson, and R. de Vries. 2005. Scientists behaving badly. *Nature*. 435(7043): 737-738. https://doi.org/10.1038/435737a

McClellan, F. 2008. The Vioxx Litigation: A Critical Look at Trial Tactics, the Tort System, and the Roles of Lawyers in Mass Tort Litigation. *DePaul Law Review*. 57: 509–538

McGarity, T.O. 2003. Our Science Is Sound Science and Their Science Is Junk Science: Science-Based Strategies for Avoiding Accountability and Responsibility for Risk-Producing Products and Activities. *Kansas Law Review*. 897-937

McGarity, T.O. and W.E. Wagner. 2008. Bending Science: How Special Interests Corrupt Public Health Research. Harvard University Press, Cambridge, MA

McKinley, J. 2016. G.E. Spent Years Cleaning Up the Hudson. Was It Enough? , New York Times. https://www.nytimes.com/2016/09/09/nyregion/general-electric-pcbs-hudson-river.html.

McNutt, M. 2016. #IAmAResearchParasite. *Science*. 351(6277): 1005-1005. https://doi.org/10.1126/science.aaf4701

McNutt, M., K. Lehnert, B. Hanson, B.A. Nosek, A.M. Ellison, and J.L. King. 2016. Liberating field science samples and data. *Science*. 351(6277): 1024-1026. https://doi.org/10.1126/science.aad7048

Mebane, C.A., R.J. Eakins, B.G. Fraser, and W.J. Adams. 2015. Recovery of a mining-damaged stream ecosystem. *Elementa: Science of the Anthropocene*. 3(1): 000042. https://doi.org/10.12952/journal.elementa.000042

Mebane, C.A. and J.S. Meyer. 2016. Environmental Toxicology without Chemistry and Publications without Discourse: Linked Impediments to Better Science. *Environmental Toxicology and Chemistry*. 35(6): 1335–1336. https://doi.org/10.1002/etc.3418

Melvin, S.D., K.R. Munkittrick, T. Bosker, and D.L. MacLatchy. 2009. Detectable effect size and bioassay power of mummichog (*Fundulus heteroclitus*) and fathead minnow (*Pimephales promelas*) adult reproductive tests. *Environmental Toxicology and Chemistry*. 28(11): 2416–2425. https://doi.org/10.1897/08-601.1

Menzie, C. and R. Smith. 2018. Scientific Integrity Must Rise Above Partisanship (9 August 2018). SETAC Globe. https://globe.setac.org/scientific-integrity-must-rise-above-partisanship/.

Meyer, J.L., P.C. Frumhoff, S.P. Hamburg, and C. de la Rosa. 2010. Above the din but in the fray: environmental scientists as effective advocates. *Frontiers in Ecology and the Environment*. 8(6): 299-305. https://doi.org/10.1890/090143

Meyer, J.N. and A.B. Francisco. 2013. A call for fuller reporting of toxicity test data. *Integrated Environmental Assessment and Management*. 9(2): 347-348. https://doi.org/10.1002/ieam.1406

Michaels, D. 2008. Doubt is Their Product: How Industry's Assault on Science Threatens Your Health. Oxford University Press. 384 pp

Moermond, C.T.A., R. Kase, M. Korkaric, and M. Ågerstrand. 2016. CRED: Criteria for reporting and evaluating ecotoxicity data. *Environmental Toxicology and Chemistry*. 35(5): 1297-1309. https://doi.org/10.1002/etc.3259

Mozur, M.C. 2012. A global professional society. *Integrated Environmental Assessment and Management*. 8(1): 1-1. https://doi.org/10.1002/ieam.1270

Mudge, J.F., T.J. Barrett, K.R. Munkittrick, and J.E. Houlahan. 2012. Negative Consequences of Using α = 0.05 for Environmental Monitoring Decisions: A Case Study from a Decade of Canada's Environmental Effects Monitoring Program. *Environmental Science & Technology*. 46(17): 9249-9255. https://doi.org/10.1021/es301320n

Munkittrick, K.R., C.J. Arens, R.B. Lowell, and G.P. Kaminski. 2009. A review of potential methods for determining critical effect size for designing environmental monitoring programs. *Environmental Toxicology and Chemistry*. 28(7): 1361-1371. https://doi.org/10.1897/08-376.1

Muscat, J.E. and M.S. Huncharek. 2008. Perineal Talc Use and Ovarian Cancer: A Critical Review. *European journal of cancer prevention*. 17(2): 139-146. https://doi.org/10.1097/CEJ.0b013e32811080ef

NAS. 1992. Responsible Science: Ensuring the Integrity of the Research Process. National Academy of Sciences, National Academy Press, Washington, D.C. https://doi.org/10.17226/1864.

NAS. 2017. Fostering Integrity in Research. National Academy of Sciences, National Academy Press, Washington, D.C. 284 pp. https://doi.org/10.17226/21896.

Nature Editors. 2014. Why high-profile journals have more retractions (17 September 2014). *Nature*. https://doi.org/10.1038/nature.2014.15951

Nature Editors. 2018. Nature journals tighten rules on non-financial conflicts. *Nature*. 554(6). https://doi.org/10.1038/d41586-018-01420-8

Nature Medicine Editors. 2017. Take science off the stand. *Nature Medicine*. 23(3): 265-265. https://doi.org/10.1038/nm.4303

Nelson, B. 2009. Data sharing: Empty archives. *Nature*. 461: 160-163. https://doi.org/10.1038/461160a

Nelson, M.P. and J.A. Vucetich. 2009. On Advocacy by Environmental Scientists: What, Whether, Why, and How. *Conservation Biology*. 23(5): 1090-1101. https://doi.org/10.1111/j.1523-1739.2009.01250.x

Nichols, T. 2017. The Death of Expertise: The Campaign Against Established Knowledge and Why it Matters. Oxford University Press. 252 pp

Norton, S.B., S.M. Cormier, and G.W. Suter, II. 2002. The easiest person to fool. *Environmental Toxicology and Chemistry*. 21(6): 1099-1100. https://doi.org/10.1002/etc.5620210601

Nosek, B.A. and 39 co-authors. 2015. Promoting an open research culture. *Science*. 348(6242): 1422-1425. https://doi.org/10.1126/science.aab2374

Nosek, B.A. and T.M. Errington. 2017. Making sense of replications. *eLife*. 6: e23383. https://doi.org/10.7554/eLife.23383

NRC-CNRC. 2013. NRC's Research Integrity Policy. National Research Council Canada. 28 pp. http://www.nrc-cnrc.gc.ca/eng/about/policies/research_integrity/.

NRC. 2002. Integrity in Scientific Research: Creating an Environment That Promotes Responsible Conduct. National Research Council (US), National Academy Press, Washington, D.C.

Nuzzo, R. 2015. How scientists fool themselves – and how they can stop. *Nature*. 526: 182–185 (08 October 2015). https://doi.org/10.1038/526182a

Obama, B. 2009. Scientific Integrity (Presidential documents, Memorandum of March 9, 2009). *Federal Register*. 74(46): 10671-10672

Ogden, L.E. 2016. Nine years of censorship. *Nature*. 533 26–28. https://doi.org/10.1038/533026a

Oreskes, N., D. Carlat, M.E. Mann, P.D. Thacker, and F.S. vom Saal. 2015. Viewpoint: Why Disclosure Matters. *Environmental Science & Technology*. 49(13): 7527-7528. https://doi.org/10.1021/acs.est.5b02726

Oreskes, N. and E.M. Conway. 2011. Merchants of doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming Bloomsbury Press. 355 pp

Owen, S.F., D.B. Huggett, T.H. Hutchinson, M.J. Hetheridge, P. McCormack, L.B. Kinter, J.F. Ericson, L.A. Constantine, and J.P. Sumpter. 2010. The value of repeating studies and multiple controls: replicated 28-day growth studies of rainbow trout exposed to clofibric acid. *Environmental Toxicology and Chemistry*. 29(12): 2831-2839. https://doi.org/10.1002/etc.351

Parker, K.R. and J.A. Wiens. 2005. Assessing recovery following environmental accidents: environmental variation, ecological assumptions, and strategies. *Ecological Applications*. 15(6): 2037-2051. https://doi.org/10.1890/04-1723

Parker, T.H., W. Forstmeier, J. Koricheva, F. Fidler, J.D. Hadfield, Y.E. Chee, C.D. Kelly, J. Gurevitch, and S. Nakagawa. 2016. Transparency in Ecology and Evolution: Real Problems, Real Solutions. *Trends in Ecology & Evolution*. 31(9): 711-719. https://doi.org/10.1016/j.tree.2016.07.002

Paul, A.P., J.R. Garbarino, L.D. Olsen, M.R. Rosen, C.A. Mebane, and T.M. Struzeski. 2016. Potential sources of analytical bias and error in selected trace element data-quality analyses. U. S. Geological Survey Scientific Investigations Report 2016-5135, Reston, VA. 68 pp. http://pubs.er.usgs.gov/publication/sir20165135.

Peterman, R.M. and M. M'Gonigle. 1992. Statistical power analysis and the precautionary principle. *Marine Pollution Bulletin*. 24(5): 231-234. https://doi.org/10.1016/0025-326X(92)90559-O

Pielke, R.A., Jr. 2007. The Honest Broker. Cambridge University Press. 188 pp

PLOS, 2014, Data Availability, accessed March 2, 2017 at http://journals.plos.org/plosone/s/data-availability.

PLOS Medicine Editors. 2008. Making Sense of Non-Financial Competing Interests. *PLoS Med*. 5(9): e199. https://doi.org/10.1371/journal.pmed.0050199

Power, M., G. Power, and D.G. Dixon. 1995. Detection and decision-making in environmental effects monitoring. *Environmental Management*. 19(5): 629-639. https://doi.org/10.1007/BF02471945

Raloff, J. 2010. Atrazine paper's challenge: Who's responsible for accuracy? *Science News*. May 6, 2010

Rekdal, O.B. 2014. Academic urban legends. *Social Studies of Science*. 44(4): 638-654. https://doi.org/10.1177/0306312714535679

Renner, R. 2005. Proposed selenium standard under attack. *Environmental Science & Technology*. 39(6): 125A-126A. https://doi.org/10.1021/es053215n

Resnik, D.B., T. Neal, A. Raymond, and G.E. Kissling. 2015. Research Misconduct Definitions Adopted by U.S. Research Institutions. *Accountability in Research*. 22(1): 14-21. https://doi.org/10.1080/08989621.2014.891943

Resnik, D.B. and A.E. Shamoo. 2011. The Singapore Statement on Research Integrity. *Accountability in Research*. 18(2): 71-75. https://doi.org/10.1080/08989621.2011.557296

Robbins, R. 2017. A supplement maker tried to silence this Harvard doctor — and put academic freedom on trial (January 10, 2017). STAT. https://www.statnews.com/2017/01/10/supplement-harvard-pieter-cohen/.

Roberts, P.D., G.B. Stewart, and A.S. Pullin. 2006. Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation*. 132(4): 409-423. https://doi.org/10.1016/j.biocon.2006.04.034

Rohr, J.R. and K.A. McCoy. 2010. Preserving environmental health and scientific credibility: a practical guide to reducing conflicts of interest. *Conservation Letters*. 3(3): 143-150. https://doi.org/10.1111/j.1755-263X.2010.00114.x

Ross, T. 2017. Crime in the cancer lab (January 27, 2017). New York Times. https://www.nytimes.com/2017/01/28/opinion/sunday/a-crime-in-the-cancer-lab.html?_r=1.

Ruff, K. 2015. Scientific journals and conflict of interest disclosure: what progress has been made? *Environmental Health*. 14(1): 1-8. https://doi.org/10.1186/s12940-015-0035-6

Russell, M., G. Boulton, P. Clarke, D. Eyton, and J. Norton. 2010. The Independent Climate Change Email Review. 160 pp. http://www.cce-review.org/.

Santore, R.C., A.C. Ryan, F. Kroglund, H.C. Teien, P.H. Rodriguez, W.A. Stubblefield, A.S. Cardwell, W.J. Adams, and E. Nordheim. 2018. Development and application of a Biotic Ligand Model for predicting the toxicity of dissolved and precipitated aluminum. *Environmental Toxicology and Chemistry*. 37(1): 70-79. https://doi.org/10.1002/etc.4020

Sass, J.B., B. Castleman, and D. Wallinga. 2005. Vinyl Chloride: A Case Study of Data Suppression and Misrepresentation. *Environmental Health Perspectives*. 113(7): 809-812

Schäfer, R.B., M. Bundschuh, A. Focks, and P.C. von der Ohe. 2013. To the Editor [authors should make all their raw data accessible]. *Environmental Toxicology and Chemistry*. 32(4): 734-735. https://doi.org/10.1002/etc.2140

Schiermeier, Q. 2012. Arsenic-loving bacterium needs phosphorus after all. *Nature*. https://doi.org/10.1038/nature.2012.10971

Schindler, D.W. 1998. Replication versus realism: The need for ecosystem-scale experiments. *Ecosystems*. 1(4): 323-334. https://doi.org/10.1007/s100219900026

Scott, A.P. 2012. Do mollusks use vertebrate sex steroids as reproductive hormones? Part I: Critical appraisal of the evidence for the presence, biosynthesis and uptake of steroids. *Steroids*. 77(13): 1450-1468. https://doi.org/10.1016/j.steroids.2012.08.009

Scott, A.P. 2018. Is there any value in measuring vertebrate steroids in invertebrates? *General and Comparative Endocrinology*. 265: 77-82. https://doi.org/10.1016/j.ygcen.2018.04.005

Scott, J.M. and J.L. Rachlow. 2011. Refocusing the Debate about Advocacy. *Conservation Biology*. 25(1): 1-3. https://doi.org/10.1111/j.1523-1739.2010.01629.x

Scott, J.M., J.L. Rachlow, R.T. Lackey, A.B. Pidgorna, J.L. Aycrigg, G.R. Feldman, L.K. Svancara, D.A. Rupp, D.I. Stanish, and R.K. Steinhorst. 2007. Policy Advocacy in Science: Prevalence, Perspectives, and Implications for Conservation Biologists. *Conservation Biology*. 21(1): 29-35. https://doi.org/10.1111/j.1523-1739.2006.00641.x

Sedlak, D. 2016. Crossing The Imaginary Line. *Environmental Science & Technology*. 50(18): 9803-9804. https://doi.org/10.1021/acs.est.6b04432

Shaw, D. and P. Satalkar. 2018. Researchers' interpretations of research integrity: A qualitative study. *Accountability in Research*. 1-15. https://doi.org/10.1080/08989621.2017.1413940

Shrader-Frechette, K. 2012. Research Integrity and Conflicts of Interest: The Case of Unethical Research-Misconduct Charges Filed by Edward Calabrese. *Accountability in Research*. 19(4): 220-242. https://doi.org/10.1080/08989621.2012.700882

Skorupa, J.P., T.S. Presser, S.J. Hamilton, A.D. Lemly, and B.E. Sample. 2004. EPA's Draft Tissue-Based Selenium Criterion: A Technical Review. U.S. Fish and Wildlife Service, U.S. Geological Survey, U.S. Forest Service and CH2M Hill. 35 pp. https://wwwrcamnl.wr.usgs.gov/Selenium/library.htm

Smaldino, P.E. and R. McElreath. 2016. The natural selection of bad science. *Royal Society Open Science*. 3(9). https://doi.org/10.1098/rsos.160384

Smith, R. 2006. Conflicts of interest: how money clouds objectivity. *Journal of the Royal Society of Medicine*. 99(6): 292-297

Smith, R. and I. Roberts. 2016. Time for sharing data to become routine: the seven excuses for not doing so are all invalid [version 1; referees: 2 approved, 1 approved with reservations]. *F1000Research*. 5(781). https://doi.org/10.12688/f1000research.8422.1

Snyder, B.F. and L.M. Hooper-Bui. 2018. A Teachable Moment: The Relevance of Ethics and the Limits of Science. *BioScience*. bix165-bix165. https://doi.org/10.1093/biosci/bix165

Solomon, K.R., J.A. Carr, L.H. Du Preez, J.P. Giesy, R.J. Kendall, E.E. Smith, and G.J. Van Der Kraak. 2008. Effects of Atrazine on Fish, Amphibians, and Aquatic Reptiles: A Critical Review. *Critical Reviews in Toxicology*. 38(9): 721-772. https://doi.org/10.1080/10408440802116496

Sprague, L.A., G.P. Oelsner, and D.M. Argue. 2017. Challenges with secondary use of multi-source water-quality data in the United States. *Water Research*. 110: 252-261. https://doi.org/10.1016/j.watres.2016.12.024

Stedeford, T. 2007. Prior restraint and censorship: acknowledged occupational hazards for government scientists. *William and Mary Environmental Law and Policy Review*. 31(3): 725–745.

Stein, R. and J. Eilperin. 2010. Obama administration issues guidelines designed to ensure 'scientific integrity' (December 17, 2010). *Washington Post*.

Steneck, N.H. 2006. Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*. 12(1): 53-74. https://doi.org/10.1007/PL00022268

Stern, V. 2017. Updated: Why would a university pay a scientist found guilty of misconduct to leave? (2 October 2017). *Science*. https://doi.org/10.1126/science.aan7182

Stokstad, E. 2012. Fracking Report Criticized for Apparent Conflict of Interest (July 24, 2012). Science. http://www.sciencemag.org/news/2012/07/fracking-report-criticized-apparent-conflict-interest.

Sumner, P., S. Vivian-Griffiths, J. Boivin, A. Williams, C.A. Venetis, A. Davies, J. Ogden, L. Whelan, B. Hughes, B. Dalton, F. Boy, and C.D. Chambers. 2014. The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ : British Medical Journal*. 349. https://doi.org/10.1136/bmj.g7015

Sumpter, J.P., R.L. Donnachie, and A.C. Johnson. 2014. The apparently very variable potency of the anti-depressant fluoxetine. *Aquatic Toxicology*. 151: 57-60. https://doi.org/10.1016/j.aquatox.2013.12.010

Sumpter, J.P., A.P. Scott, and I. Katsiadaki. 2016. Comments on Niemuth, N.J. and Klaper, R.D. 2015. Emerging wastewater contaminant metformin causes intersex and reduced fecundity in fish. Chemosphere 135, 38–45. *Chemosphere*. 165: 566-569. https://doi.org/10.1016/j.chemosphere.2016.08.049

Suter, G.W., II and S.M. Cormier. 2015a. Bias in the development of health and ecological assessments and potential solutions. *Human and Ecological Risk Assessment*. 22(1): 99-115. https://doi.org/10.1080/10807039.2015.1056062

Suter, G.W., II and S.M. Cormier. 2015b. The Problem of Biased Data and Potential Solutions for Health and Environmental Assessments. *Human and Ecological Risk Assessment*. 21(7): 1736-1752. https://doi.org/10.1080/10807039.2014.974499

Suter, G.W., II., S.B. Norton, and S.M. Cormier. 2002. A method for inferring the causes of observed impairments in aquatic ecosystems. *Environmental Toxicology and Chemistry*. 21(6): 1101-1111. https://doi.org/10.1002/etc.5620210602

Tollefson, J. 2015. Earth science wrestles with conflict-of-interest policies. *Nature*. 522: 403-404. https://doi.org/10.1038/522403a

Topf, A., 2016, Imperial Metals sues engineering companies over Mount Polley disaster (July 10, 2016): Mining.com, accessed February 7, 2017 at http://www.mining.com/imperial-metals-sues-engineering-companies-mount-polley-disaster/.

U.S. Department of Interior, 2014, Code of Scientific and Scholarly Conduct (poster), at https://www.doi.gov/scientificintegrity/upload/DOI-Code-of-Scientific-and-Scholarly-Conduct-Poster-December-2014.pdf.

USEPA. 2013. Aquatic Life Ambient Water Quality Criteria For Ammonia – Freshwater 2013. EPA 822-R-13-001. 255 pp. http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/ammonia/index.cfm.

van Iersel, S., E.M. Swart, Y. Nakadera, N.M. van Straalen, and J.M. Koene. 2014. Effect of Male Accessory Gland Products on Egg Laying in Gastropod Molluscs. *Journal of Visualized Experiments*. 88: e51698. https://doi.org/10.3791/51698

Van Kirk, R.W. and S.L. Hill. 2007. Demographic model predicts trout population response to selenium based on individual-level toxicity. *Ecological Modelling*. 206(3-4): 407-420. https://doi.org/10.1016/j.ecolmodel.2007.04.003

van Kolfschooten, F. 2002. Conflicts of interest: Can you believe what you read? *Nature*. 416(6879): 360-363

Van Noorden, R. 2014. Publishers withdraw more than 120 gibberish papers (24 February 2014). *Nature*. https://doi.org/10.1038/nature.2014.14763

Vines, T.H., A.Y.K. Albert, R.L. Andrew, F. Débarre, D.G. Bock, M.T. Franklin, K.J. Gilbert, J.-S. Moore, S. Renaut, and D.J. Rennison. 2014. The availability of research data declines rapidly with article age. *Current Biology*. 24(1): 94-97. https://doi.org/http://dx.doi.org/10.1016/j.cub.2013.11.014

Wadman, M. 1997. $100m payout after drug data withheld. *Nature*. 388(6644): 703-703

Wagner, W.E. 2005. The perils of relying on interested parties to evaluate scientific quality. *American Journal of Public Health*. 95(S1): S99-S106. https://doi.org/10.2105/AJPH.2004.044792

Wagner, W.E. and D. Michaels. 2004. Equal Treatment for Regulatory Science: Extending the Controls Governing the Quality of Public Research to Private Research. *American Journal of Law & Medicine*. 30(2-3): 119-154. https://doi.org/10.1177/009885880403000202

Weltje, L. and J.P. Sumpter. 2017. What makes a concentration environmentally relevant? Critique and a Proposal. *Environmental Science & Technology*. 51: 11520−11521. https://doi.org/10.1021/acs.est.7b04673

Whaley, P., C. Halsall, M. Ågerstrand, E. Aiassa, D. Benford, G. Bilotta, D. Coggon, C. Collins, C. Dempsey, R. Duarte-Davidson, R. FitzGerald, M. Galay-Burgos, D. Gee, S. Hoffmann, J. Lam, T. Lasserson, L. Levy, S. Lipworth, S.M. Ross, O. Martin, C. Meads, M. Meyer-Baron, J. Miller, C. Pease, A. Rooney, A. Sapiets, G. Stewart, and D. Taylor. 2016. Implementing systematic review techniques in chemical risk assessment: Challenges, opportunities and recommendations. *Environment International*. 92-93: 556-564. https://doi.org/10.1016/j.envint.2015.11.002

Wiens, J.A. and K.R. Parker. 1995. Analyzing the effects of accidental environmental impacts: approaches and assumptions. *Ecological Applications*. 5(4): 1069-1083. https://doi.org/10.2307/2269355

Wise, J. 1997. Research suppressed for seven years by drug company. *BMJ : British Medical Journal*. 314(7088): 1145-1145. https://doi.org/10.1136/bmj.314.7088.1145

Womack, R.P. 2015. Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics. *PLoS ONE*. 10(12): e0143460. https://doi.org/10.1371/journal.pone.0143460

Young, N.S., J.P.A. Ioannidis, and O. Al-Ubaydli. 2008. Why Current Publication Practices May Distort Science. *PLOS Medicine*. 5(10): e201. https://doi.org/10.1371/journal.pmed.0050201

Figure 1. Conflicts of interest in science arise when secondary interests such as financial gain or maintaining professional relationships compromise the primary interest of upholding scientific norms such as the objective design, conduct, and interpretation of studies and the open sharing of scientific discoveries to advance our collective learning (© Benita Epstein, used with permission.)

423x417mm (72 x 72 DPI)

Figure 2. Confirmation bias is the tendency to seek and interpret evidence in a way that confirms preexisting beliefs and gives less consideration to alternative hypotheses (© Benita Epstein, used with permission]).

423x353mm (72 x 72 DPI)

Figure 3. Large environmental chemistry and toxicology laboratories that use standard methods to produce results that may be submitted to regulatory agencies usually have a well-established quality management structure. Quality management in academic research laboratories focused on novel methods may be more ad hoc, especially if the research work force is dominated by transient scientists, such as students or those on short-term postgraduate appointments (Credit: S.Harris, sciencecartoonsplus.com).

216x240mm (300 x 300 DPI)

Figure 4. The brief methods descriptions in journal articles are seldom sufficient to be reproducible by others. Step-by-step video documentation of experimental protocols can be published as video articles, uploaded to online repositories, or published as supplemental information. Video protocols are underutilized in environmental toxicology (Credit: S. Harris, Sciencecartoonsplus.com).

184x242mm (300 x 300 DPI)

Integrated Environmental
Assessment and Management

Scientific Integrity Issues in Environmental Toxicology: improving research reproducibility, credibility, and transparency

Journal:	*Integrated Environmental Assessment and Management*
Manuscript ID	IEAM-2018-029-CR
Wiley - Manuscript type:	Critical Review
Date Submitted by the Author:	27-Feb-2018
Complete List of Authors:	Mebane, Christopher; National Oceanic and Atmospheric Administration, Sumpter , John; Brunel University Fairbrother, Anne; Exponent, EcoSciences Augspurger, Tom; USFWS Canfield, Timothy; US Environmental Protection Agency, National Risk Management Research Laboratory, Ground Water and Ecosystems Restoration Division Goodfellow, William; Exponent Inc Guiney, Pat; Guiney LeHuray, Anne; Chemical Management Associates LLC Maltby, Lorraine; University of Sheffield, Department of Animal and Plant Sciences Mayfield, David; Gradient, Toxicology Ortego, Lisa; Bayer CropScience LP McLaughlin, Mike; CSIRO/University of Adelaide, Schlekat, Tamar; Society of Environmental Toxicology and Chemistry Scroggins, Rick; Environment Canada, Verslycke, Tim; Gradient,
Keywords:	scientific integrity, reproducibility, bias, transparency
Abstract:	High profile reports of detrimental scientific practices leading to retractions in the scientific literature contribute to lack of trust in scientific experts. While the bulk of these have been in the biomedical literature, environmental sciences and ecotoxicology are not excepted from questionable practices. While we believe that egregious misconduct such as fraud is rare and when uncovered is universally condemned, we are more concerned with more commonly encountered issues, such as poor reliability and bias. These issues may be nuanced, and require thoughtful consideration rather than condemnation. We review a range of topics including conflicts of interests, competing interests, particularly challenging situations, reproducibility, publication, and other biases, quality management and other attributes of ecotoxicological studies that enhance or detract from scientific credibility. We propose a conceptual framework for considering scientific integrity as an extension of personal integrity, encouraging a self-correcting culture of scientific rigor, appropriate transparency, and objective review.

Scientific Integrity Issues in Environmental Toxicology: improving research reproducibility, credibility, and transparency

Christopher A. Mebane[1], John P. Sumpter[2], Anne Fairbrother[3], Thomas P. Augspurger[4], Timothy J. Canfield[5], William L. Goodfellow[6], Patrick D. Guiney[7], Anne LeHuray[8], Lorraine Maltby[9], David B. Mayfield[10], Michael J. McLaughlin[11], Lisa Ortego[12], Tamar Schlekat[13], Richard P. Scroggins[14], Tim A. Verslycke[15]

[1]U.S. Geological Survey, Boise, ID USA; [2]Brunel University London, UK; [3]Exponent Inc., Bellevue, WA, USA; [4]U.S. Fish and Wildlife Service, Raleigh, NC USA; [5]U.S. Environmental Protection Agency, Ada, OK USA; [6]Exponent Inc., Alexandria, VA USA; [7]University of Wisconsin, Madison, WI USA; [8]Chemical Management Associates LLC, Alexandria VA, USA; [9]University of Sheffield, Sheffield, UK; [10]Gradient, Seattle, WA, USA; [11]University of Adelaide, Adelaide, South Australia, Australia; [12]Bayer, Crop Science Division, Research Triangle Park, NC, USA; [13]SETAC, Pensacola, FL USA; [14] Environment and Climate Change Canada, Ottawa, ON, Canada; [15]Gradient, Cambridge, MA, USA

Abstract

High profile reports of detrimental scientific practices leading to retractions in the scientific literature contribute to lack of trust in scientific experts. While the bulk of these have been in the biomedical literature, environmental sciences and ecotoxicology are not excepted from questionable practices. While we believe that egregious misconduct such as fraud is rare and when uncovered is universally condemned, we are more concerned with more commonly encountered issues, such as poor reliability and bias. These issues may be nuanced, and require thoughtful consideration rather than condemnation. We review a range of topics including conflicts of interests, competing interests, particularly challenging situations, reproducibility, publication, and other biases, quality management and other attributes of ecotoxicological studies that enhance or detract from scientific credibility. We propose a conceptual framework for considering scientific integrity as an extension of personal integrity, encouraging a self-correcting culture of scientific rigor, appropriate transparency, and objective review.

Introduction

Highly polarized recent elections in Europe and North America have shown that large segments of society are distrustful of scientific and other experts. Some have suggested that we are in a post-modern public culture in which reality is defined by the observer and objective facts do not change peoples' minds, and those that conflict with one's beliefs are justifiably questionable (Campbell and Friesen 2015). Science and scientists have been central to these debates, and the boundaries of science, policy and politics may be indistinct. In a social climate skeptical of science, the easy availability of numerous reports of dubious scientific practices gives fodder to skeptics. Because environmental regulations on use of chemicals and waste management rely heavily on the disciplines of ecotoxicology and chemistry, the integrity of the science is of utmost importance. Here we discuss scientific integrity in the applied environmental sciences, with a focus on ecotoxicology and how the role and culture of the Society of Environmental Toxicology and Chemistry (SETAC) may influence such issues.

Science has long endured questionable science practices and a skeptical public. Socrates is remembered as an early proponent of evidence-based inquiry but who suffered career-ending bad reviews from his peers. Galileo's criticisms of prevailing beliefs resulted in his issuing a public retraction of his seminal work. In contrast, science "discoveries" such as Piltdown Man, canals on Mars, cold fusion, Archaeoraptor, arsenic-life, and many others have not stood the test of time. By 1954, Huff and Geist (1954) illustrated how the presentation of scientific data could be manipulated to become completely misleading yet accurate. Are things worse now? Recent articles in both the scientific literature and popular print and broadcast venues paint a bleak picture of the status of science. One does not have to search hard to find plenty of published concerns about the credibility of science. These include overstated and unreliable results (Harris and Sumpter 2015; Henderson and Thomson 2017; Ioannidis 2005), conflicts of interest (Boone et al. 2014; McGarity and Wagner 2012; Oreskes et al. 2015; Stokstad 2012; Tollefson 2015), profound bias (Atkinson and Macdonald 2010; Bes-Rastrollo et al. 2014; Suter and Cormier 2015a, b), suppression of results to protect financial interests (Wadman 1997; Wise 1997), deliberate misinformation campaigns as a public relations strategy for financial or ideological aims (Baba et al. 2005; Gleick and 252 co-authors 2010; McGarity and Wagner 2012; Oreskes and Conway 2010), political interference with or suppression of results from government scientists (Hutchings 1997; Ogden 2016; Stedeford 2007), self-promotion and sabotage of rivals in hypercompetitive settings (Edwards and Roy 2016; Martinson et al. 2005; Ross 2017), publication bias, peer review and authorship games (Callaway 2015; Fanelli 2012; Young et al. 2008), overhyped institutional press releases that are incommensurate with the actual science behind them (Cope and Allison 2009; Sumner et al. 2014), dodgy journals (Bohannon 2013), dodgy conferences (Van Noorden 2014), and that beer both prevents and causes cancer (Oransky 2015)[1].

Such published concerns reasonably raise doubts about science and scientists, and could even lead some to conclude that the contemporary system of science is broken. In writing this commentary, we attempt to address some prominent science integrity concerns in the context of

---

[1] Throughout this commentary, citations are intended to be representative, without the "e.g." qualifier, which would otherwise be needed in nearly every instance.

environmental toxicology. In our view, there is ample room for improvement within our discipline, but environmental science is not broken and some criticisms are overstated. In writing this commentary, we do not pretend to have some moral high ground that sets us apart from our peers or arguments that will overturn insidious pressures on scientists and funders for impressive results. Thus our recommendations are pragmatic, not dogmatic. Our goal is nudge practices and pressures on scientists to advance the science, while maintain and improving credibility through transparency, ongoing review, and self-correction.

Many of the prominent science integrity controversies have been in the high stakes biomedical discipline, and in response that discipline probably has done more self-evaluation and taken more steps toward best practices than most other disciplines. Results of self-reported, anonymous surveys of scientists, mostly in the biomedical fields, have not been reassuring. In a 2002 survey of early and mid-career scientists, 0.3% admitted to falsification of data, 6% to a failure to present conflicting evidence, and 16% to changing of study design, methodology or results in response to funder pressure (Martinson et al. 2005). A subsequent meta-analysis of surveys suggested problems were more common, with close to 2% of scientists admitting to having been involved in serious misconduct, and over 70% reported they personally knew of colleagues committing less severe detrimental research practices (Fanelli 2009). Serious misconduct such as fraud can occur in ecotoxicology just as with any disciplines (Enserink 2017; Keith 2015) and when exposed, is universally condemned and in most countries is career ending. In contrast, the ambiguous, more nuanced issues of science integrity that all of us are likely to experience in our careers require thoughtful consideration, not condemnation. It is toward the latter that we discuss efforts toward remedies from other disciplines to examine similar issues in ecotoxicology, focusing on SETAC.

What is "science" in the context of scientific integrity?

Before we can discuss integrity in ecotoxicology and related environmental science fields, we must first distinguish what is meant by "science" in this context. Broadly speaking, environmental science includes the disciplines of biology, ecology, chemistry, physics, geology, limnology, mineralogy, marine studies, and atmospheric studies; i.e., the study of the natural world and its interconnections. The applications of environmental science extend to agriculture, fisheries management, forestry, natural resource conservation, and chemicals management, all of which have associated multi-billion dollar industries and vocal environmental advocacy groups. The subdiscipline of environmental toxicology or ecotoxicology, pursued by SETAC scientists, studies in great detail how the natural world is influenced by chemicals, both natural and synthetic, introduced by human endeavors that are largely in pursuit of the production of desired goods and services (food, clean water, plastic products, metals, etc.). Because exposure to chemicals can have negative and sometimes unexpected consequences for people and the environment, a body of regulation has developed over the past century to control the kinds and amounts of allowable chemical exposures. Such regulations necessarily are based on scientific concepts such as Paracelsus' directive that "the dose makes the poison" (Kolok 2016) and physicochemical properties that influence transport and fate of substances. Because of the complexity, inexactitude, and uncertainty of ecotoxicology and associated sciences, rulemaking often is subject to challenge, leading to accusations of profit over people or the environment or unreasonably restrictive and burdensome requirements. Scientists are called upon to inform disputes based on their knowledge or underlying principals or enter the conversation through self-initiated in-depth literature review and commentary. While regulators and the public assume such scientific advice is unbiased and policy-neutral, there are concerns of science being normative with biased interpretations as scientists engage in controversial policy debates (Lackey 2007). Only by conscientiously adhering to fundamental principles of the scientific method can environmental scientists maintain their integrity and continue to play a valid role in environmental policy and management.

What is "scientific integrity"?

Impeccable honesty is a fundamental tenet of science. When we read a paper, we might not agree with the conclusions, authors' interpretations of its implications, importance, or many other things, but we have to be confident that the procedures described were indeed followed and all relevant data were shown, not just those fitting the hypothesis. Goodstein (1995) put it well. "*There are, to be sure, minor deceptions in virtually all scientific papers, as there are in all other aspects of human life. For example, scientific papers typically describe investigations as they logically should have been done rather than as they actually were done. False steps, blind alleys and outright mistakes are usually omitted once the results are in and the whole experiment can be seen in proper perspective.*"

Various professional and governmental organizations have established policies and definitions prescribing research integrity, responsible conduct of science, or scientific integrity. These may include broad statements of attributes such as the U.S. National Academy of Science's (NAS) six values that they considered most influential in shaping the norms that constitute research practices and relationships and the integrity of science: objectivity, honesty, openness, accountability, fairness, and stewardship (NAS 2017). More specific "research integrity" guidelines define appropriate expectations of individual researchers and their institutions and may be highly procedural. Protecting the privacy, rights and safety of human research participants and animal welfare with institutional review board clearance requirements is a common element of research integrity guidelines. Academic research integrity guidelines have been established individually or in aggregate by research funders and individual institutions (ARC 2007; Goodstein 1995; NRC-CNRC 2013; NRC 2002; Resnik and Shamoo 2011; Steneck 2006). Research institutions are usually responsible for investigating potential breaches of research integrity by its scientists, although this can create difficult conflicts of interest for the institution (Glanz and Armendariz 2017).

Whether research integrity guidelines should best be defined narrowly or broadly has been an area of controversy. As of 2015, 22 of the world's top 40 research countries had national research conduct policies, all of which included fabrication, falsification, and plagiarism (FFP), with some going further. In this context, "fabrication" is making up data; "falsification" includes manipulating studies or changing or omitting data such that the record does not accurately reflect the actual research; and "plagiarism" includes the appropriation of another person's ideas, methods, results, or words without giving appropriate credit (https://ori.hhs.gov/definition-misconduct). The Research Councils of the UK has a lengthy list of misdeeds including FFP, misrepresentation, breach of duty of care, and improper dealing with allegations of misconduct, with many subcategories (NAS 2017). In contrast, from the 1980s to 2000, the National Science Foundation (US) had defined serious science misconduct broadly to include, "*...fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting and reporting research*" (Goodstein 1995). The controversial part was the catchall phrase "*practices that seriously deviate from those commonly accepted...*" To the stewards of public science funds, such a catchall phrase was preferable to an itemized lists of all potential avenues of mischief, yet it raised the specter of penalizing scientists who strayed too far from orthodox thought (Goodstein

1995). In 2000, this definition of disbarring research misconduct was narrowed to just "*fabrication, falsification, or plagiarism in proposing, performing, or reporting research*" with lesser offenses classified as questionable research practices. Other misconduct was defined as "*forms of unacceptable behavior that are clearly not unique to the conduct of science, although they may occur in the laboratory or research environment*." Yet only FFP research misconduct findings were subject to reporting requirements to federal science funding agencies, with questionable science practices or other misconduct handled locally (NAS 2017; Resnik et al. 2015). In many countries, there is an active debate about whether a legal definition is appropriate for something that is really an academic judgment rather than a legal one. Denmark recently similarly narrowed its broad definitions of research misconduct to only FFP following high profile cases in which scientists succeeded in having their academic misconduct findings overturned in the courts. Yet if research conduct policies are considered "academic" without legal weight, institutions may have difficulty enforcing polices, such as when deliberate intent is required to be shown and the researcher claims "honest mistake" .For instance, the U.S. Office of Research Integrity found that a tenured professor had committed research misconduct by inappropriately altering data in five images from three papers. Yet when the university sought to terminate her, she fought back hiring a lawyer to contest the university's procedures, and the university ultimately paid her $100,000 USD to leave (Stern 2017).

Unfortunately, the "other misconduct" that scientists may commit can reflect that of any work place, such as abuse of power; bullying, sexual coercion, assault, or harassment; misuse of funds; sabotage; taking advantage of students or subordinates; specious whistleblowing; or retaliation against valid whistleblowers; to name a few (e.g., Ghorayshi 2016; Gibbons 2014; http://retractionwatch.com). The exclusion of such malfeasance from "research misconduct" has been questioned. For example, a researcher who failed to meet her study objectives after being sabotaged by a rival argued that she was further penalized by being instructed not to divulge the reason for her study failures to her funders (Enserink 2014). In contrast, institutions often do go beyond the minimum "FFP" definition in their policies (Resnik et al. 2015), which has led to objections of conflation of egregious misconduct such as fraud with failure to comply with administrative requirements that did not compromise data validity (Couzin-Frankel 2017).

The U.S. National Academy of Science (NAS 2017) recently argued that the definitions of research misconduct as fabrication, falsification, or plagiarism were too narrow. In particular, questionable research practices were more than just "questionable," but were clear violations of the fundamental tenets of research and were given a less ambiguous label of "detrimental." Consensus detrimental research practices were:

1. Detrimental authorship practices that may not be considered misconduct, such as honorary authorship, demanding authorship in return for access to previously collected data or materials, or denying authorship to those who deserve to be designated as authors.
2. Not retaining or making data, code, or other information/materials underlying research results available as specified in institutional or sponsor policies, or standard practices in the field.
3. Neglectful or exploitative supervision in research.
4. Misleading statistical analysis that falls short of falsification.

5. Inadequate institutional policies, procedures, or capacity to foster research integrity and address research misconduct allegations, and deficient implementation of policies and procedures, and
6. Abusive or irresponsible publication practices by journal editors and peer reviewers (NAS 2017).

The term “scientific integrity” is sometimes used synonymously with research integrity. However in recent usage, the term has included insulation of science from political interference, manipulation, or suppression of science (Doremus 2007; Douglas 2014). The term “scientific integrity” has been used in government science policy in the United States. There, scientific integrity guidelines were developed in an overarching sense that includes research integrity at the individual and institutional level, but were also intended to protect federal scientists from political interferences. Political officials were not to alter or suppress scientific findings, and transparency was encouraged in the preparation of the government-supported scientific research (Obama 2009; Stein and Eilperin 2010). The scientific integrity guidelines in the US were followed by derivative policies intended to put substance to the transparency provisions, requiring open-access to federally funded research articles and more importantly, requiring archiving and public availability of the underlying raw data (Holdren 2013). These broad policies become more specific and procedural in government science agencies, and expanded to codes of scholarly and scientific conduct such as a list of 19 principles for the U.S. Department of Interior (U.S. Department of Interior 2014).

We expect the vast majority of scientists consider themselves to hold science integrity, as self-defined in terms of honesty, transparency, and objectivity, sticking to the research question and avoiding bias in data interpretation (e.g., Shaw and Satalkar 2018). Yet most scientists will encounter ethically ambiguous situations. For instance, some may feel that they struggle to advance science against a rising tide of administrative requirements accompanied by declining support for science and increasing competition for funding. When does cutting through bureaucratic institutional requirements cross the line from being commendable efficiency to violating research integrity rules? Using grant/project funds for unrelated conference travel? Should minor misbehaviors such as posting ones' article on a website after signing a publication and copyright transfer agreement with the publisher agreeing not to do so still be considered misbehaviors when done by many? When does cleaning data become cooking data when, for example, anomalous values are suppressed? There are many ethically ambiguous situations in which scientists may consider doing the "right thing" (compliance with all rules) might need to be balanced with doing the "good thing," especially when the welfare of others such as students or subordinates is involved (Johnson and Ecklund 2016).

To us, scientific integrity can be simplified to personal integrity plus a few profession-specific provisions of transparency and reproducibility. At their roots, these norms are those children are hopefully exposed to in primary school. *Tell the truth, and tell the whole truth* (no data sanitizing, selective reporting, and report all conflicts), *tell both sides of the story* (avoid bias), *do your own work* (no plagiarism), *read the book, not just the back cover before writing your report* (properly research and cite primary sources), *show your work for full credit (*transparency*)*, *practice makes perfect* (reproducibility), *share* (publish your work and data in peer-reviewed outlets for collective learning), and *listen* (with humility and collegial fraternity to observations and suggestions of others). Finally, the golden rule "*do unto others as you would have them do unto you*" should resonate throughout the professional interactions of environmental scientists, and especially in peer reviewing and data sharing. When encountering an inevitable science dispute, keep criticisms objective, constructive, and focused on the work and not the worker; do peer reviews of your rivals' work as you would hope to receive reviews of your own, reward and recognize good behavior in science, and so on.

The interested scientist: conflicts of interest, competing interests, and bias

Although as far as we know, outright fraud or deliberate campaigns to manipulate science are rare in the environmental sciences, at some points in their careers almost every practicing scientist has to grapple with questions of conflicting or competing interests and must guard against bias in approaches and interpretation. Conflicts of interest refer to those where the scientists stands to gain financially from their work. Competing interests are areas where non-financial factors compete with objectivity, such as personal friendships or dislikes, having taken public stances on an issue, political, academic, ideological, or religious affiliations (Nature Editors 2018; PLOS Medicine Editors 2008). Bias in study design or data interpretation may arise from either conflicts or competing interests and can be either overt or unrecognized by the scientist (Lackey 2007).

Generally, the concern over conflicting or competing interests in science is that secondary interests such as financial gain or maintaining professional relationships compromise the primary interest of upholding scientific norms such as reporting data accurately and completely, interpreting data appropriately, and acknowledging value judgments or interpretive assumptions (Elliott 2014). Conflict of interest policies may be better developed in the biomedical fields than in the applied environmental sciences because the former often involves human participants, and because of the strong financial ties between academia and the pharmaceutical industry (Tollefson 2015). For instance, if a research team is reporting on the efficacy of a medical device or a pharmaceutical, and they or their employers hold a patent or stand to gain financially from a positive report, then they clearly have a financial conflict of interest.

Figure 1. Conflicts of interest in science arise when secondary interests such as financial gain or maintaining professional relationships compromise the primary interest of upholding scientific norms such as the objective design, conduct, and interpretation of studies and the open sharing of scientific discoveries to advance our collective learning (© Benita Epstein, used with permission.)

The mere existence of a potential conflict of interest should not throw results in doubt where it is disclosed and acknowledged appropriately. However, although most authors in the environmental sciences routinely disclose funding sources that could be perceived as potential conflicts of interest, major omissions have occurred (Oreskes et al. 2015; Ruff 2015; Tollefson 2015). For instance, the findings of a study on risks of contamination from natural gas extraction from hydraulic fracturing of bedrock were undermined when it came out that (unbeknownst to the university) the research supervisor was being paid 3X his university salary by serving as an advisor to an oil and gas company invested in the practice. The failure to disclose this financial relationship in the publication brought the study's objectivity and credibility into question, independent of its substance (Stokstad 2012). Authors and journals have been criticized for gaming ethical financial disclosure requirements, such as by overly narrow disclosures or disclosing a conflict in the cover letter to the editor accompanying the manuscript (which doesn't get published) but not including it in the actual article (Marcus and Oransky 2016).

The detail of conflict of interest disclosures vary. The shortest (and least informative) statement we have seen was that "*the usual disclaimers apply*" (Descamps 2008), while the detailed disclosures expected in biomedical literature can go on for pages (Baethge 2013; ICMJE 2016). Requirements for highly detailed disclosures risk diminishing their importance to that of the "fine print" cautions the writers would just as soon the readers not take the time to carefully read. Much like computer software user agreements or the ubiquitous consumer product safety stickers that may be written more to avoid product liability than for consumer safety, detailed conflict of interest disclosures may reach a point of diminishing returns. Twenty years ago, Goodstein (1995) groused that he was tired of reading disclosure statements that were longer than the methods sections in papers, and they have been expanded upon since. Similarly, our view is that in ecotoxicology and the environmental sciences, simple, unambiguous statements of the funding sources that allowed the work to be completed should generally be sufficient.

Non-financial competing factors may also compete with scientific objectivity. Factors or values such as these are usually termed "competing interests" reserving "conflicts of interest" for financial conflicts (Nature Editors 2018; PLOS Medicine Editors 2008). In our experience, competing interests are rarely if ever mentioned in environmental science publications. Rather, they are often discussed behind the scenes, such in correspondence between an editor and potential reviewers, along the lines of "*yes I would be happy to review this article and believe I can be objective, however, you should know that I used to be a labmate of the PI and we collaborated on an article 3 years ago*." Whether or how competing interests or values affect the assumptions and perspectives of scientists' should be more formally stated is an area of rich debate in the philosophy of science literature (Douglas 2015; Elliott 2016; PLOS Medicine Editors 2008).

We reiterate our belief that the existence of a potential conflicts or competing interests is a ubiquitous part of the environmental science landscape and do not indicate poor science. Most scientists strive to present unbiased data and interpret their data evenhandedly. However, the varied experiences of scientists can influence their perspectives in ways that they may not recognize themselves. The transparency in disclosure reminds the reader to consider perspectives and alternate interpretations when judging the merits of a study, synthesis paper, or risk assessment.

Bias

Many of the published concerns in the environmental science literature come down to cognitive bias. Science is not value free, and personal bias in interpreting science is often related to differing worldviews (Douglas 2015; Elliott 2016; Lackey 2001; Nuzzo 2015). For instance, the collapse of major fisheries that ostensibly had been scientifically managed for sustainable yields helped inspire the Precautionary Principle. This philosophy sought more cautious management and the reversal of the burden of proof for sustainable exploitation of natural resources to put it on industry not management agencies (Peterman and M'Gonigle 1992). Those with precautionary principle or risk assessment worldviews may interpret the same set of facts very differently. The precautionary principle adherent may emphasize absence of conclusive evidence of safety, and the risk assessment adherent may emphasize absence of conclusive evidence of harm (Fairbrother and Bennett 1999). In such settings, values and biases are interwoven. Even self-disciplined scientists who seek openness and objectivity carry some biases from experiences and acculturation (here meaning how working in different environmental organisations can lead scientists to modify their thinking). Recognizing sources of bias does not imply ill intent, for just the process of acculturation to a particular place of employment can bias perceptions and inclinations (Brain et al. 2016; Suter and Cormier 2015a, b).

Professional societies such as SETAC can serve as a form of acculturation; some of the authors of this essay have been active members of SETAC for much longer than they have been employed by any single employer. Even self-disciplined scientists who seek openness and objectivity carry biases from their experiences. What becomes particularly difficult to self-regulate is the convergence of cognitive bias, a human nature to seek to please one's patron, and the interests of one's employer or client. For instance, studies funded by drug or medical device makers tend to favor the company funding the research (Lexchin et al. 2003; Smith 2006). That might reflect the self-interest and bias of the sponsor, or the researchers' intimate knowledge and their ability to obtain the resources and skill to carry out well focused and rigorous research (Macleod 2014). These influences doubtfully can be completely separated. To us, disclosure, transparency and balanced external reviews are presently the best pragmatic approach to managing cognitive biases.

Tit for tat, adversarial claims of bias in the scientific literature doubtfully advance the science. Conflicting perspectives can become personalized and intractable. How to know which is more credible? Neither? Both? Food nutrition researchers pointed out examples of selective data interpretations and publication bias in obesity research in relation to sweetened beverage (soft drink) consumption and in the health benefits of breast feeding They termed this distortion of information to further what may be perceived to be righteous ends as "white hat bias" (Cope and Allison 2009). However, their financial backing from the soft drink industry and from manufacturers of baby formula contributed to criticisms of their own objectivity (Bes-Rastrollo et al. 2014; Harris and Patrick 2011). Unresolved in the claims and counter-claims of bias and financial conflicts of interest was what advice was most credible.

In environmental toxicology as well, controversies over the best interpretation of sometimes ambiguous facts can become entrenched and focused on the people holding differing views as much as the evidence behind the different views. Examples include disagreements over risks of atrazine to amphibians (Aviv 2014; Hayes 2004; Solomon et al. 2008); sufficiently safe levels of selenium for fish and birds (Renner 2005; Skorupa et al. 2004); and 20 years on, disputes over indirect effects of oil spills on salmon (Burton and Ward 2012). These intractable, mutual bias criticisms make it very difficult for non-specialist readers to make informed judgements of which is the more credible science.

Suter and Cormier (2015a) noted that conflicting assessments on the same question that have been produced by government agencies, industries, and environmental advocacy groups suggest that biases occur during assessment processes. Sources of bias include personal bias, regulatory capture, advocacy, reliance on volunteer experts, biased stakeholder and peer review processes, literature searches, excluding new science through dependence on standard methods, inappropriate standards of proof, misinterpretation, and ambiguity. Assessors can adopt practices to increase objectivity, transparency, and clarity (Suter and Cormier 2015a).

Some particularly challenging situations in ecotoxicology – Some situations that seem particularly challenging for researcher and institutions to maintain scientific credibility warrant mention. Elliott (2014) argued that scientific findings that are ambiguous or require a good deal of interpretation or are difficult to establish in an obvious and straightforward manner are prone to bias, particularly if strong incentives to influence research findings in ways that damage the credibility of research are present. In environmental toxicology, risk assessments or critical reviews fit that test and can be vulnerable to bias, particularly when funded by sponsors with financial interests in the findings (Suter and Cormier 2015a). This can be heightened by how variability and large uncertainties are handled in environmental toxicology and associated risk assessments and syntheses -- for example extrapolation of results from one or more species to protection of wide swaths of our world's biodiversity; or the difficulty in reproducing field studies; or the variability of chemical exposures across diverse and expansive landscapes and waters. These challenges may lead to differences of opinion on methods for drawing conclusions to support decision-making that, while prone to bias, have, at their root, the need for drawing conclusions in the face of uncertainty.

Costs of large-scale projects to remediate contaminated environments such as sediments contaminated by urban and industrial sources, aged industrial facilities, or large mining operations can be extremely expensive, running to the hundreds of millions of dollars. In "polluter-pays" schemes, the potential financial liability associated with such a finding could imperil the ongoing viability of companies, which in turn would affect the livelihoods of employees, among other social disruptions. In such a setting, the scientists working on behalf of the those who may have to incur the costs of cleanup might understandably be more cautious about the potential for misguided remediation following Type I error (e.g., falsely discovering environmental degradation) than Type II error (failing to discover degradation when in fact it is occurring), when the science is ambiguous. Conversely, the regulatory scientists entrusted to provide scientific advice to protect environmental quality might be obliged to err on the side of precaution, and be more accepting of risk of Type I error, especially when it is other peoples' money at stake.

While science ethicists and the NAS (Boden and Ozonoff 2008; Elliott 2014; Krimsky 2005; NAS 1992) may emphasize industry funding as a pressure for bias, these pressures are not unique to industry funding of science. Natural resource damage assessment (NRDA) is an example in the environmental toxicology field where government-funded science has strong incentives to generate biased science or assessments. NRDAs compensate for harm to natural resources from oil spills or poorly managed industrial activities have been good for the environment, but as practiced in many countries have parallels to the legal and science tactics of product liability torts (Descamps 2008). The potential for recovering fees or paying substantial penalties can provide financial incentives for lawyers, regulators, or resource trustees and interested parties to pursue cases (Murray et al. 1999). This environment produces an atmosphere with strong incentives for plaintiff/trustee science advisors to exaggerate the magnitude and spatial extent of effects to the environment and to downplay uncertainties or the influence of potential other, non-compensable stressors and vice versa for those scientists retained to help defend against claims. Maintaining objectivity and advancing science in such a work environment would require extraordinary self-discipline by the individual scientists, an institutional environment emphasizing science first, and an openness to external, disinterested review (Boden and Ozonoff 2008; Elliott 2014; Wagner 2005).

Defense of science and engineering in favor of protecting enterprises reflecting years of devoted work are understandable, but is dangerous when objectivity is compromised. Case studies such as the Vioxx case, in which the maker of the drug downplayed increased risks of mortality from a successful product in which they were deeply vested (Curfman et al. 2005; McClellan 2008) and the cross-claims of blame in the aftermath of the Mount Polley mine tailing dam failure (Topf 2016), remind us that objective science (including recognizing and disclosing uncertainty, and encouraging additional science to narrow that uncertainty) is good business.

Academic – Industry Collaborations

The role of industry funding and concerns of perceived conflicts of interest in academic-industry collaborations have been addressed in literature and are a common element in institutional research integrity policies (Elliott 2014; Resnik and Shamoo 2011). Often through philanthropic foundations, industry may contribute to basic science education and research to strengthen regional universities and further the science literacy of potential workforce and society. Industry may also support applied ecotoxicology and other environmental science research to inform specific scientific questions that affect their business interests. When industry and academic research interests become at least partially congruent, academic scientists may actively seek out such interest and support for their projects and graduate students. Pragmatically, academic-industry collaborations are necessary since public funding alone may be insufficient to support graduate research or to address important questions relevant to industry and society. For instance in the US, about 40% of national research and development is funded by the private sector (NAS 2017). In the US, public funding for university research on the effects of chemicals in the environment has consistently declined since 2000 (Bernhardt et al. 2017;

Burton et al. 2017), which implies that without industry-academic collaborations, there would be much less substantive university research in the field.

Benefits of collaboration run both ways, with expertise from academic and public sectors helping industry find solutions to lessen or avoid contributing to environmental problems (Hopkin 2006). The interchange of science through academic, industry, and government scientists is deeply rooted in SETAC culture, and the favorable views of the authors toward working across sectors is undoubtedly influenced through our history with SETAC. However, industry support to academics or others in support of applied environmental questions may come with inherent conflicts of interest, and critics may consider scientists as collaborators in the pejorative sense of the word (Hopkin 2006). This setting requires vigilance from both industrial research sponsors and recipients to avoid bias.

While readers might presume situations in which individuals or institutions with strong incentives to influence research findings consistent with their financial interests will do so, it is important not to judge a study solely by its funder, nor to presume the sponsor’s preferred outcome. For example, an energy company sponsored a study to see if they could develop a scientific case for relief from costly requirements for meeting dissolved oxygen criteria in a river downstream of its hydroelectric dam. Instead they developed evidence that the existing criteria could impair hatching salmon (Geist et al. 2006). The company scientists easily could have buried the results, which could have been discounted as being from novel techniques. Their path of least resistance would have been to leave the study in the file drawer, rather than going to the trouble of defending novel science and publishing it in the open literature. In the long-view, a reputation of science credibility may be more valuable for companies than short-term project benefits.

Other examples include scientists from mining and metals trade groups publishing studies showing that existing USEPA criteria for zinc and other metals could be under-protective of aquatic species or entire communities (Brix et al. 2011; DeForest and Van Genderen 2012). Conversely, a university quantitative ecologist accepted support from an environmental advocacy group (through university channels) to model the potential population-level effects of elevated selenium from mining on local native trout populations (Van Kirk and Hill 2007). As the advocacy group had been a persistent opponent of the mining operations, officials from the influential mining company apparently presumed that the academics' work would also be biased to favor the advocacy group's positions, and they questioned the researchers' probity (Blumenstyk 2007). In fact, the selenium concentrations projected by these academics to cause detrimental population-level effects were higher than concentrations previously derived by industry-funded consultants who themselves had been on the receiving end of bias implications because they were aligned with corporate interests (Skorupa et al. 2004; Van Kirk and Hill 2007). As these examples show, judging science and scientists solely by their funding may be unfair and lead to misjudgments.

In keeping with the adage to be careful judging a book by its cover or wine by its label, judging science by its funder or by presumed interests or leanings of the scientists can lead to mistaken and unfair perceptions. Brain et al. (2016) pointed out that the career path of environmental scientists is often ambiguous and whether scientists ended up in careers with industry, academic, or government science has more to do with chance and timing of opportunities rather than a particular desire to work in one sector or another. Such is often the case with academic and government scientists who work with industry to jointly fund or investigate a science question of mutual interest (Hopkin 2006). The convergence of scientific interests with financial interests can lead to a good marriage, so long as the parties are principled and forthright with each other. “Interested” science” should be viewed with open-minded skepticism, and studies with immense financial implications warrant a higher level of scrutiny than others (Krumholz et al. 2007; Suter and Cormier 2015b; van Kolfschooten 2002). It does not necessarily follow that interested science is wrong or tainted.

"I already wrote the paper. That's why it's so hard to get the right data."

Figure 2. Confirmation bias is the tendency to seek and interpret evidence in a way that confirms preexisting beliefs, and gives less consideration to alternative hypotheses (© Benita Epstein, used with permission]).

A scientific society founded on the principles of balancing competing interests

Scientific societies have important roles in promoting scientific integrity and ethical conduct, such as establishing codes of ethics which include disclosure of conflicts of interest, being a focal point for developing and communicating discipline-specific standards to foster research integrity, and providing educational material (AAAS 2000; NAS 2017).

We think the Society of Environmental Toxicology and Chemistry (SETAC) is notable for its directed and sustained efforts to balance competing perspectives in its deliberative processes and other activities. The founding principles and structure of SETAC sets out a tripartisan structure with regulatory, industrial, and academic scientists (Bui et al. 2004). As a result, SETAC now has well developed norms for balancing interests, inclusiveness of differing viewpoints, and neutrality in the reporting. These norms have enabled SETAC to be regarded as a source of consensus-based science with successful partnership or advisory roles in United Nations programs and conventions such as the United Nations Environment Programme's (UNEP) Global Mercury Partnership, Stockholm Convention on persistent organic pollutants, UNEP-SETAC Life Cycle Initiative for reducing hazardous waste as well as informing national-level legislation (Augspurger 2014; Mozur 2012). The intended balanced representation of industry, government, and academia isn't always achievable, for there are also guidelines for gender equity, geographic representation, and of course people have to be willing to volunteer. Further, the tripartisan emphasis underrepresents scientists from environmental advocacy groups. These groups are influential for shaping public debate, policy and law on environmental issues, but their low participation in the Society suggests that they may not be attracted to or feel welcomed by a "hard" scientific society such as SETAC. Despite these imperfections, the norms of seeking to balance potentially conflicting interests and to provide a safe forum to express differing scientific viewpoints are deeply ingrained in the Society's culture and activities.

Promoting scientific integrity in ecotoxicology

While “scientific integrity” is ultimately a subjective judgment that cannot easily be reduced to review checklists, there are some general points to maintain in ecotoxicology and related science. These include relevance, rigor, reproducibility, objectivity, and transparency.

Relevance

By definition, environmental chemistry and ecotoxicology is concerned with how chemicals, both natural and synthetic, pose a threat or influence the natural world (Johnson et al. 2017). Because of pragmatic and ethical constraints, research in this domain is often done in laboratory environments, testing cultured laboratory organisms or cell lines or other in vitro surrogates for organisms. However, the intent of such research invariably still has some intended relevance to conditions that occur in the environment. We have seen articles in ecotoxicology literature discussing some novel research based on under-tested taxa, underappreciated endpoints, unexpected multiple stressor effects, or unanticipated indirect effects via untested commensal microbes. An article may start out with an introduction on the ecological importance of the novel work, the work is reported, and then the discussion closes arguing that ecological importance of their work, how it should change the thinking in the field, and management implications. Yet to obtain their desired experimental effects, exposure concentrations may have been orders of magnitude higher than those typical in the real world, or exposure routes, chemical forms, or dilution media may be unlike those that they organisms could encounter in nature (Johnson and Sumpter 2016; Mebane and Meyer 2016). When authors present such studies with a narrative on the ecological importance of their topic, this may be a form of misrepresentation.

Rigor

Funders, journals, and institutions reward novelty, such as the short-lived discovery of a bacterium that grows with arsenic instead of phosphorus (Alberts 2012). Highly selective journals with article acceptance rates of 10% or less preferentially publish findings that are surprising. These incentives are influential because universities and research institutes often hire and promote scientists based on their record of acquiring grant money and the number of publications times the journal impact factors of the journals published therein (Parker et al. 2016). With finite career opportunities and high network connectivity, the marginal return for being in the top tier of publications may be orders of magnitude higher than an otherwise respectable publication record (Smaldino and McElreath 2016). The editorial quest for novelty has led to publication of questionable articles in elite journals, such as one positing that caterpillars were the results of accidental sex between insects and worms (Borrell 2009). Top tier journals also tend to have higher retraction rates than mid-tier journals, suggesting that rigor has sometimes been compromised in the competition for paradigm shifting results (Nature Editors 2014).

In ecotoxicology, Harris et al (2014) describe 12 basic principles of sound ecotoxicology that should apply to most environmental toxicity studies. These principles range from carefully considering essential aspects of experimental design through to accurately defining the exposure, adequate replication, unbiased analysis and reporting of the results, and repeating experiments that yielded surprising or ambiguous responses. There are ample opportunities for improvement. For example, Harris and Sumpter (2015) asked a very basic question of a sample of studies published in 2013 in three leading ecotoxicological publications: was the concentration of the test chemical actually measured? Of the studies reviewed from *Environmental Toxicology and*

Chemistry, 20% failed this basic aspect of experimental credibility, as did 33% and 41% of ecotoxicology studies published in *Aquatic Toxicology*, and *Environmental Science and Technology*, respectively (Harris and Sumpter 2015).

While Harris et al. (2014) emphasized laboratory-based studies, field-based environmental effects studies replace the challenges of the artificiality and questionable relevance of some laboratory-based toxicity testing, with different, messy, real world challenges. Closely related to the 12 principles described by Harris et al, we suggest 8 basic principles relevant to most field-based ecotoxicological studies or environmental effects monitoring.

1. The study design is grounded in a good understanding of the test questions (Lindenmayer and Likens 2010; Suter et al. 2002);
2. The ability to identify and reliably measure sensitive indicators (Melvin et al. 2009),
3. Careful attention to appropriate reference conditions to avoid potential, actual effects being masked by variability or confounding factors introduced by differences between the reference and test site environments (Arciszewski and Munkittrick 2015; Mebane et al. 2015). For example, beaches on rocky headlands and protected bays will have very different benthic invertebrate communities, as do flowing rivers and impounded reservoirs. Study designs that attempt to detect pollution effects on communities across such disparate habitats may have very low discriminatory power and by failing to account for natural variability, adverse pollutant effects could be obscured (Buys et al. 2015; Parker and Wiens 2005; Wiens and Parker 1995);
4. Try to study a number of locations that vary in the degree of the factor under investigation, such as chemical pollution, in order to (hopefully) demonstrate a

positive relationship between exposure to the environmental factor of interest and the effect of that factor.

5. Time and patience. Just as experimental exposures need to be of appropriate duration for effects of interest to be manifested, environmental monitoring needs to be maintained long enough to pick up true trends if present, or to convincingly argue that trends are not present (Lindenmayer and Likens 2010; Melvin et al. 2009).
6. Specific definitions of what effects are considered negligible or of concern (Munkittrick et al. 2009; Power et al. 1995).
7. Avoid power failures: use a statistical approach appropriate to the question, considering statistical burden of proof issues. For instance, $P>0.05$ in testing for trends or differences between locations does not by itself show the lack of trend or effects (Dixon and Pechmann 2005; Mudge et al. 2012).
8. Transparent reporting with detailed methods and raw data sufficient for others to reproduce the analyses or to further examine the data using alternative analyses (Duke and Porter 2013; McNutt et al. 2016; Schäfer et al. 2013).

Reproducibility

Reproducibility is one indicator of reliable research. However, the inability of researchers to reproduce influential studies of others or their own has garnered enough attention to be called a "reproducibility crisis" (Baker 2016a; Henderson and Thomson 2017). However, not all studies are easily reproduced. Environmental data are often messy, field studies are more often observational than experimental, large scale, ecologically realistic studies such as long-term, experimental lake studies difficult to do even once, and no one wishes to replicate mishaps such as tailings dam failures or oil spills (Parker and Wiens 2005; Schindler 1998; Wiens and Parker

1995). Such studies require a logical system for causal inference to separate cause-and-effect from serendipitous correlations (Norton et al. 2002; Suter et al. 2002). Even rigorous laboratory studies may be difficult to replicate due to the highly variable nature of biological systems and unanticipated responses to unknown factors. Demands for reproducibility may favor industrial science over academic science. Industry often works within strict Good Laboratory Practice (GLP) rules and with well-studied species tested through standardized protocols (Elliott 2016). Academic science is often framed around education, and grants and graduate student research are usually required to go after something new and novel; protocols may be developed as they go, and quality control may be uneven (Baker 2016b). Obstacles to adopting formalized quality management systems such as GLP in small research settings may include costs, lack of resources, lack of mandate, independent cultures, and high turnover. Nevertheless, even if regulatory GLP compliance is not required, small academic research facilities can benefit from embracing core components of GLPs, such as defining responsibilities, maintenance and sanitation of common lab spaces, equipment and materials, well defined experimental protocols, quality control testing, data reviews, audits, and archiving (Bornstein-Forst 2017).

Better experimental protocols that are easier to follow is one tangible way to strive for better reproducibility and transferability of both novel and standard experimental methods. Multimedia experimental protocols could be much easier to explain and teach techniques than the conventional, densely worded, printed protocols. The Journal of Visualized Experiments (JoVE) is an innovative peer reviewed, science methods journal in which its articles are a unique blend of the conventional printed article with professionally produced videography. Ecotoxicology methods articles have begun to be published in this format (Calfee et al. 2016; van Iersel et al. 2014). The field would benefit from better exploiting new visualization techniques to document new methods and to improve education and training on techniques that need to be highly standardized to be repeatable. At the minimum, with the availability of electronic data repositories and supplemental information in journals, there is no reason why detailed methods cannot be published.

Figure 3. Large environmental chemistry and toxicology laboratories that use standard methods to produce results that may be submitted to regulatory agencies usually have a well-established quality management structure. Quality management in academic research laboratories focused on novel methods may be more ad hoc, especially if the research work force is dominated by transient scientists, such as students or those on short-term postgraduate appointments (Credit: Sidney Harris, sciencecartoonsplus.com).

Reproducing a statistical summary or model run reported in a scientific publication when the underlying data and code are provided and explained is one thing. Reproducing an actual complex experiment is hard and is rarely attempted, unless perhaps the results are novel and have a high regulatory or societal impact. Even under the best of circumstances, such as when the original researchers are diligent enough to repeat an experiment in the same lab with as close to identical methods as they could manage, it can be difficult or impossible to produce the same result twice (Owen et al. 2010). Nosek and Errington (2017) caution that if investigator #2 reports that the results of study #1 could not be reproduced, that does not indicate which is more credible: result #1, #2, neither, or both. Further, much of the "reproducibility" debate in the natural sciences is focused on cell biology or human behavior (psychology) experiments, which may be more tractable to reproducibility studies than messy environmental observational or experimental studies. Especially with complex biological testing such as multi-generation tests, a green thumb husbandry factor may bring together art and science to environmental chemistry and toxicology. Subtle methods differences, strain differences or stochastic events can be so puzzling that investigators are left thinking demons must have snuck into their study and interfered with one treatment but not others (Hurlbert 1984). (We note that Hurlbert's (1984) suggestions for exorcisms or human sacrifice for troubleshooting suspected demonic intrusions, might run afoul of contemporary institutional review board policies.)

Still, reproducibility is a core tenet of science and successful reproduction adds confidence in the credibility of novel findings. Divergent but individually credible results may further advance the science by illuminating important aspects missed in the initial study (Owen et al. 2010). If for instance, an investigator were to find a novel, major adverse effect of a class of chemicals to a previously untested taxonomic group, then other equally diligent investigators should be able produce similar effects in other research settings, even if the test conditions were only similar. A standalone paper from the 1970s that a snail was anomalously sensitive to Pb was skeptically regarded. Over 30 years later, this open-minded skepticism led to follow-on studies from a new generation of scientists that not only affirmed the unusual early report of sensitivity but also led to important advances in comparative physiology and underlying mechanisms of toxicity (Brix et al. 2012). Similarly, early reports that freshwater mussels and other mollusks were unusually sensitive to ammonia were not widely persuasive. After repeated studies across multiple laboratories and species showed similar findings, the issue gained traction with standardized method development, inter-laboratory round robin testing, and attention by environmental managers (Farris and Hassel 2006; USEPA 2013).

Individual investigators may not always have the opportunities for self-replication, but best practices call for repeating what one can (Harris and Sumpter 2015). In field studies, multiple measures of exposure, multiple years of field data, and so on give credence to findings. We recognize that all science has practical resource limits and we are not going as far as arguing that novel findings from small sample studies should never be published. Rather, the appropriate conclusion from such studies is along the lines of "if these findings turn out to be repeatable, they could be an important development." In our view, novel, major findings that are supported only by a one-off study are best regarded as tentative.

Figure 4. The brief methods descriptions in journal articles are seldom sufficient to be reproducible by others. Step-by-step video documentation of experimental protocols can be published as video articles, uploaded to online repositories, or published as supplemental information. Video protocols are underutilized in environmental toxicology (Credit: Sidney Harris, Sciencecartoonsplus.com).

Transparency

Transparency in reporting research, including all the relevant underlying data that were relied upon in the paper, has become a critical element of integrity in science. Science's claim to self-correction and overall reliability is based on the ability of researchers to replicate the results of published studies (Nosek and 39 co-authors 2015). Studies cannot be replicated if scientists will not share additional data, information, or materials from published studies, and we believe that upholding such ethical norms is every scientist's responsibility. The embrace of the principle of transparent reporting has been uneven across disciplines, and the field of ecotoxicology has certainly not distinguished itself as a leader in this regard (McNutt et al. 2016; Meyer and Francisco 2013; Parker et al. 2016; Schäfer et al. 2013; Womack 2015).

Researchers in ecotoxicology and environmental chemistry have long only presented highly reduced data summaries. The only "data" included in some publications are crowded figures and tables with results of statistical outputs, such as F- values, effects concentration point estimates (EC50, EC10, etc.), or no-and lowest-observed effects concentrations (NOECs, LOECs). These derived values are not data. Such data-poor publications essentially represent an implicit claim by the researcher to "*trust us, we know what we're doing, our interpretation of the data is the only appropriate interpretation, you don't need to see what you don't see, and besides it's our data to share as we see fit.*" Such attitudes reflect the norm in scientific publishing prior to the early 2000s, in which strict page limits and word limits precluded authors "wasting" space publishing data tables. With the provisions for electronic supplemental material beginning in the 2000s, and dedicated data repositories becoming widely available at low or no costs to authors in the 2010s, these reasons for opaque publication are no longer justified. Researchers who choose not to transparently report the actual data underlying their scientific findings may have other reasons for doing so. They may be concerned about others scooping them on their own data (McNutt 2016), although counterintuitively, publishing data may actually help establish priority and reduce scooping concerns (Laine 2017). Other less charitable reasons why researchers might resist publishing data include that they haven't devoted the needed time to organize their data in a coherent fashion that is interpretable by others, because reported results might not be replicable from the underlying data, they are not keen to facilitate alternate statistical analyses or interpretations of their data, that they wish to publish unfalsifiable findings, or because there's simply less there than they led readers to believe (Smith and Roberts 2016).

Data sharing may still be regarded more as an imposition from science funders to be complied with rather than as a universal principle embraced by those conducting and publishing scientific research (Collins and Verdier 2017; European Commission 2016; Holdren 2013; Nelson 2009; Nosek and 39 co-authors 2015). There are many pragmatic obstacles to effective data sharing, such as the expertise, extra work, and costs to researchers to organize, serve, and preserve their data in a comprehensible manner, privacy and anonymity concerns for environmental data collected from private property, about human subjects, and balancing intellectual property concerns. Some environmental science research is intended to be secret, such as mining and economic geology, agricultural chemical product development, and innumerable other corporate research efforts which are intended to develop products and recoup investments[2]. However, in our view, researchers on such ventures cannot have it both ways, by publishing some outcomes in the peer reviewed literature, but withholding the supporting data as private.

Most environmental science journals have policies encouraging and facilitating data sharing. SETAC journals are probably typical in requiring a statement by the authors' whether and how the data underlying their analyses are available, with an admonition that authors should share upon request. A passable statement may be something as weak as "*Contact the Corresponding Author for data availability*."

The strongest data disclosure policy for journals publishing in the environmental sciences is probably that developed for the Public Library of Science (PLoS) family of journals. "*PLoS*

---

[2] (see however, a recent corporate initiative to make available traditionally protected crop safety information https://cropscience-transparency.bayer.com/) .

journals require authors to make all data underlying the findings described in their manuscript fully available without restriction, with rare exception" (PLOS 2014). Exceptions are limited to privacy or vulnerability concerns such as data on human research subjects that could not be fully anonymized, locations of archeological, fossil, or endangered species, that could be exploited or damaged, or safety and security considerations. Penalties for authors who fail to comply include rejection, or if they decline to provide data for an already published article, the editors could flag their article with a cautionary correction or even retract it (PLOS 2014). Whether PLoS's stand requiring authors to make available all data underlying their findings will lead other journals to stiffen their resolve, or whether the comparatively lax policies of competing journals will undermine PLoS and other open-science advocates remains to be seen (Davis 2016; Nosek and 39 co-authors 2015).

The reality of moving toward transparent data availability and preservation is thus more challenging and complicated than the notion that it should be done. To us is it a priority to strongly encourage, for without data, the credibility of science cannot be evaluated. Some research has shown the willingness and ability for authors to share data declines significantly with time, and having a weak data availability policy is only marginally better than having no policy at all (Vines et al. 2014).

Rather than mandates, one simple incentive to improving openness in reporting has been for journals to award prominent open data "badges" for articles verified as being supported by available, correct, usable, and complete data. By showing an open data badge on the issue table of contents, article web page, and including a "verified open data" statement in the bibliographic indexing metadata, articles without such badge endorsement may be seen as incomplete. Over time, this might shift the norm toward open preservation and sharing. In at least one journal, this approach appeared to markedly improve the sharing and preservation of data through linked, independent repositories (Kidwell et al. 2016).

Critical Reviews and Literature Syntheses

In ecotoxicology, published literature can roughly be broken down into two categories: original research and the review article. The original research article usually is based upon field observations, laboratory experiments, modelling, or blended approaches. Generalizing original articles through reviews and syntheses are critical parts of the ecotoxicology and most environmental science literature. Critical reviews, risk assessments, environmental quality standards, are based on syntheses of the literature, and not on individual studies. Synthesis articles have rather distinct scientific integrity problems from the original research article. Decisions must be made on how studies were located, results categorized, and a host of data manipulation and analyses decisions need to be made. These decisions and associated biases may be deliberate and clearly explained or the analyst may not even recognize that they have made a decision. In some cases we suspect analysts obscured their decisions. In some cases, data synthesis may be highly structured, with clearly defined criteria for data inclusion (Hobbs et al. 2005), and search strategies. Others may follow the wending path of the present article: discussions among the authors "*have you read so-and-so?*", and readings that led to other relevant material through forward and backward citing, along with by some specific subject searches. This path led to much relevant and thoughtful material across many disciplines. But it was hardly systematic or reproducible.

Literature searches from different sources can yield very different results. For example, using a 2007 original research article on population modeling of selenium toxicity to trout (Van Kirk and Hill 2007), four leading bibliographic indexing services were searched for articles citing that study. Web of Science (WoS), Elsevier's Scopus, Digital Science's Dimensions, and Google Scholar found 7, 10, 15, and 22 citing publications respectively. Scopus found all articles found by WoS, plus articles in *Human and Ecological Risk Assessment* and *IEAM*. Google Scholar found all articles found by Scopus plus articles in *Ecotoxicology Modeling*, *Water Resources Research*, 3 government reports, 2 books, a thesis, a conference proceeding, a duplicate, and 2 ambiguous citations from grey regulatory documents. It follows from this 3 fold difference in valid citations to an article that a critical review of published literature on a topic or a regulatory assessment could miss relevant science if the assessors relied too heavily on a single search provider.

This simple example was from the current era of science, which began by 1996 or so, depending on which bibliographic indexing service scholars are using. Web sites for WoS and Scopus respectively report their indexing databases are reliable from 1971 and 1996 forward. Relying exclusively on bibliographic index searching may omit important, relevant older research.

Thus we have the indexing bias problem in meta-analyses and assessment (that not indexed won't be retrieved), and the related problem of reviewing the secondary source but citing the original. We have seen assessments that omitted seminal research published before the current digital era, which may reflect indexing bias. Ecotoxicology syntheses often rely on variations of species-sensitivity distributions, which may provide more explanations of statistical characteristics of the datasets, data extrapolations, transformations, normalizations, than on where the data came from in the first place. We have seen micrograms and milligrams mixed up, and statistical rankings that commingled endpoints such as time to death in hours with effects concentrations. Some of these issues are undoubtedly related to the online availability of well curated databases such as ECETOC Aquatic Toxicity (EAT) Database from the European Center for Ecotoxicology and Toxicology of Chemicals or the U.S. Environmental Protection Agency's EcoTox databases. These compiled databases are valuable resources but reliance on secondary, compilations deprive the original authors of credit via citations. At least for publicly funded science, citations may be a way that authors demonstrate the value of their work to the scientific community, and thus build the case for further funding. Further, reliance on secondary sources is a good way to introduce or repeat inaccuracies (Rekdal 2014). We echo previous calls for better training and rigor when conducting and reporting secondary analyses of ecotoxicology and related literature. Practices from other fields, such as the Cochrane systematic review approach and guidelines for the ethical reuse of data could be adapted to the ecotoxicology practices (Duke and Porter 2013; Roberts et al. 2006; Suter and Cormier 2015a).

Advocacy

Science is the enterprise for answering questions and making predictions about the how the universe works, but science can never answer "should" questions. For example, science cannot tell societies whether they should restrict chemical uses and releases, whether natural preserves should be set aside from human exploitation, or whether biodiversity should be protected. These are among the myriad value judgements that societies must make, and while science can support societies in making these choices through predictions founded upon a body of knowledge, there are never "scientifically correct" answers to questions of human values, morals, and ethics (Snyder and Hooper-Bui 2018). Scientists are humans, and like all people, hold ethical and moral values which drive assumptions which may not be explicitly stated if even recognized. For example, the notion of "environmental protection" in the environmental toxicology field is rooted in societal norms, statues, and international agreements with goals of minimizing harm (a human concept) from activities such as extraction, manufacture, use, and disposal of chemical products. Scientists in the field develop informed opinions toward the "should" questions relating to their experiences, which leads to questions of whether and how scientists advocate for "should" questions.

The underpinnings of science are that researchers have no vested interest in the results of their observations, that they objectively record and analyze these results, and that they fairly report the outcomes in the peer-reviewed literature. Advocacy can compromise these underpinnings, at the cost of scientists' credibility (Fenn and Milton 1997). Scientists tend to be passionate about their science, which has led to controversy over the role that scientists should play in related public policy debates. While we think most scientists would agree that advocacy for science having a role in environmental policy debates is appropriate, there is likely much less agreement whether it is appropriate for scientists to advocate for particular outcomes in policy debates. If the policy debate turns on questions of science central to a scientist's particular area of study, probably no one is better positioned than that scientist to lay out the evidence for or against a particular course of action. If the scientist is regarded as a neutral and informed voice, their advice may be valued by all sides in a policy dispute (Sedlak 2016). However, if the scientist's experience or analyses leads them to the strong conviction that one policy direction is more correct and should be adopted, then they are no longer a neutral broker and have become an advocate.

Policy advocacy is potentially problematic because it may compromise use of research findings in policy and management deliberations if the information is not viewed as credible by all sides (Scott et al. 2007). In some situations, advocacy is beyond reproach, such as a university scientist who uncovered a lead poisoned community water system. Simply reporting his findings to the responsible officials would have been ineffective, if the ineptitude or indifference of those same responsible officials contributed to the situation in the first place (Sedlak 2016). However, not all situations are so clear cut, and reasonable people who share similar motivations, skills, and agree that researchers should do the right thing may not agree on what that is. Deliberations on major environmental issues are complex and science may only be one element of the deliberations. Developing and providing technical and scientific information to inform policy deliberations in an objective and relevant way is formidable challenge that is easily undermined when scientists meld their own policy preferences into their scientific advice (Lackey 2007).

Institutional constraints aside, how scientists balance these competing issues and choose when or whether to engage in advocacy is a deeply personal choice and is situational. However, just as science journals discourage comingling original research results and commentary, scientists should keep science and advocacy distinct in their publications and speaking. In particular, we argue that scientists should be watchful for stealth policy advocacy. Stealth advocacy is the use of value-laden language in scientific writing that assumes a policy preference (Lackey 2007). Rather than openly disclosing assumed values or policy preferences, biases may be unconsciously (or deliberately) cloaked through normative science. Normative science is science developed, presented, or interpreted all based on an assumed, usually unstated, preference for a particular policy or class of policy choices. This covert advocacy may be reflected in word choices, and such advocacy is not always apparent even to the advocate. For instance, value-laden words such as *impacted, degraded, improved, good,* and *poor* may be used to describe habitats or other environmental features. Less value-laden words would be *exposed*, *altered, changed, increased,* or *decreased.* The use of normative science is potentially insidious because the tacit, usually unstated, preference for a particular policy or class of policy choices is not perceptibly normative to policy makers or even to many scientists (Lackey 2007). Criticisms of normative science can be too extreme, as taken literally, the entire discipline of conservation biology could be considered too normative. Similarly, the mission statement of SETAC "*to support the development of principles and practices for protection, enhancement and management of sustainable environmental quality and ecosystem integrity*" could be too much for some. Science is normative. Areas of study or techniques once considered appropriate areas of science inquiry such as craniometry, eugenics, or experimentation on human subjects without informed consent are no longer considered to be within the norms of ethical science. Within environmental toxicology, pressure to reduce the use of animal testing might be an example of normative science.

Our point is not to argue for or against scientists engaging in overt policy advocacy, which is a personal decision, but for clarity and transparency. Just as original results, opinion, judgements and speculation should not be blended in a scientific paper, science and advocacy need some separation (Scott and Rachlow 2011). Covert advocacy is a form of bias. Environmental scientists should clearly differentiate between research findings and policy advocacy based upon those findings.

Weaponizing scientific integrity

We recognize that "scientific integrity" discussions could easily be diminished to going down the path carved by "sound science" strategic initiatives, which often boiled down to campaigns to call "*my science good science and your science junk*" (Doremus 2007; Kapustka 2016; McGarity 2003). The goal may be to recast policy, ideological, or economic disputes as doubt or created conflicts in science. In countries with a tort-based, adversarial legal system for resolving injuries or damages, science-based information becomes just another tool for dueling experts, who often have primary responsibility for advocating for the interests of their client (Wagner 2005).

Research integrity policies or requirements for data transparency can be used as weapons to bury public university or government scientists with vexatious, intrusive, and costly demands for records such as raw laboratory notebooks, instrument calibration records, emails between coauthors, working drafts, and peer comments and responses. Such demands can be effective tools for interfering with the work of public-sector scientists, including academics in public institutions (Folta 2015; Halpern and Mann 2015; Kloor 2015; Kollipara 2015; Lewandowsky and Bishop 2016), or academics in private institutions but who receive research support from public sources (Hey and Chalmers 2010; Shrader-Frechette 2012). Privately funded research is generally shielded from such practices (Brain et al. 2016; Wagner and Michaels 2004). Researchers at private institutions may however be subject to baseless litigation to intimidate scientists and deter others by inflicting long and costly legal processes, disruption, and threats of personal financial liability. Such harassing lawsuits have been employed often enough to get a name, SLAPP (Strategic Litigation Against Public Participation) suits (Johnson 2007; Nature Medicine Editors 2017). While legal, such strategies represent detrimental practices cloaked in the vernacular of science (Johnson 2007; Levy and Johns 2016; McGarity and Wagner 2012; Wagner and Michaels 2004).

Promoting scientific integrity in environmental toxicology

Scientific integrity is harnessed by high quality environmental research characterized by rigor, relevance, reproducibility, and objectivity. Our review suggested several conclusions, tangible actions and less tangible directions that professional societies such as SETAC could do to encourage scientists, their supporting institutions, and science journals to maintain and improve science integrity. Scientific integrity is reinforced through full transparency exemplified by full disclosures of potential conflicting and competing interests that could contribute to bias, and by making all data and observations readily accessible. Specifically:

1. Scientific integrity in ecotoxicology and the environmental sciences cannot be ensured by impeccable policies or checklists. It is an attitude to be embraced, maintained, and enforced through the support, guidance and approval of one's peers through a community of practices.
2. Reliability, rigor, relevance and reproducibility of science are more important than novel advances.
3. Increased attention to a culture of quality management training and transparency could improve the confidence in published findings.
4. Studies that are not supported by primary data released through data repositories or detailed supporting information are not fully credible.
5. As a community, be aware of and disclose potential conflicting or competing interests that could contribute to bias; avoid and not tolerate extreme conflicts or bias.
6. Distinguish true uncertainties in science from economic, policy, or social implications of the science, and call out those who would conflate them.
7. Discourage judging science by its funder; rather, open-minded skepticism is applicable when the funder has a stake in the outcome of a study.
8. Scientists, like all people, have moral and ethical assumptions, based upon their values. These should not be intermixed with their interpretations and reporting of science. If scientists' values lead them to cross the lines from analysis to advocacy, they need to be particularly careful about distinguishing between science, values, assumptions, and opinion.

9. Professional societies such as SETAC have an important role in fostering respectful evidence-based dialog, in meetings and correspondence on published works.
10. Professional societies such as SETAC could support a standing training seminar on principles of scientific integrity, the transparent conduct of science and best practices for peer review in conjunction with its annual meetings.
11. Professional societies such as SETAC have a valuable role in facilitating balanced, expert reviews of controversial science topics, such as has been done with their Pellston Workshop series of meetings and publications.

Acknowledgements

This commentary and perspectives followed discussions by the Society of Environmental Toxicology and Chemistry (SETAC) ad hoc subcommittee on Scientific Integrity (Fairbrother 2016). No funding was provided for the writing of this manuscript. Author affiliations are given for identification, but do not necessarily imply endorsement. Further, the views are those of the authors and are not intended to represent those of SETAC. Author affiliations reflect a variety of academic, business, and governmental affiliations, and all authors have competing interests from our previous experiences and work environments, but hopefully these were somewhat balanced out. No financial conflicts of interest were declared by any author.

References

AAAS. 2000. The Role and Activities of Scientific Societies in Promoting Research Integrity. April 10-11, 2000. American Association for the Advancement of Science (AAAS) and the 24 pp. https://www.aaas.org/page/scientific-integrity.

Alberts, B. 2012. Editor's note [A bacterium that can grow using arsenic instead of phosphorus]. *Science*. 334(6034): 1149. http://dx.doi.org/10.1126/science.1208877

ARC. 2007. Australian Code for the Responsible Conduct of Research. Australian Research Council (ARC). http://www.arc.gov.au/research-integrity.

Arciszewski, T.J. and K.R. Munkittrick. 2015. Development of an adaptive monitoring framework for long-term programs: An example using indicators of fish health. *Integrated Environmental Assessment and Management*. 11(4): 701-718. http://dx.doi.org/10.1002/ieam.1636

Atkinson, R.L. and I. Macdonald. 2010. White hat bias: The need for authors to have the spin stop with them. *International Journal of Obesity*. 34(1): 83-83. http://dx.doi.org/10.1038/ijo.2009.269

Augspurger, T. 2014. SETAC Hosts Successful TSCA [Toxic Substances Control Act] Risk Assessment Science Seminar for Congressional Staff (6 November 2014). *SETAC Globe*. 15(11)

Aviv, R., 2014, A valuable reputation (February 10, 2014), *The New Yorker* at http://www.newyorker.com/magazine/2014/02/10/a-valuable-reputation.

Baba, A., D.M. Cook, T.O. McGarity, and L.A. Bero. 2005. Legislating "Sound Science": The Role of the Tobacco Industry. *American Journal of Public Health*. 95(S1): S20-S27. http://dx.doi.org/10.2105/AJPH.2004.050963

Baethge, C. 2013. The effect of a conflict of interest disclosure form using closed questions on the number of positive conflicts of interest declared – a controlled study. *PeerJ*. 1: e128. http://dx.doi.org/10.7717/peerj.128

Baker, M. 2016a. 1,500 scientists lift the lid on reproducibility. *Nature*. Nature. http://dx.doi.org/10.1038/533452a

Baker, M. 2016b. How quality control could save your science. *Nature*. 529: 456–458. http://dx.doi.org/10.1038/529456a

Bernhardt, E.S., E.J. Rosi, and M.O. Gessner. 2017. Synthetic chemicals as agents of global change. *Frontiers in Ecology and the Environment*. 15(2): 84-90. http://dx.doi.org/10.1002/fee.1450

Bes-Rastrollo, M., M.B. Schulze, M. Ruiz-Canela, and M.A. Martinez-Gonzalez. 2014. Financial Conflicts of Interest and Reporting Bias Regarding the Association between Sugar-Sweetened Beverages and Weight Gain: A Systematic Review of Systematic Reviews. *PLOS Medicine*. 10(12): e1001578. http://dx.doi.org/10.1371/journal.pmed.1001578

Blumenstyk, G. 2007. When Research Criticizes an Industry. *Chronicle of Higher Education*. 54(4): A21

Boden, L.I. and D. Ozonoff. 2008. Litigation-Generated Science: Why Should We Care? *Environmental Health Perspectives*. 116(1): 117-122

Bohannon, J. 2013. Who's Afraid of Peer Review? *Science*. 342(6154): 60-65. http://dx.doi.org/10.1126/science.342.6154.60

Boone, M.D., C.A. Bishop, L.A. Boswell, R.D. Brodman, J. Burger, C. Davidson, M. Gochfeld, J.T. Hoverman, L.A. Neuman-Lee, R.A. Relyea, J.R. Rohr, C. Salice, R.D. Semlitsch, D. Sparling, and S. Weir. 2014. Pesticide Regulation amid the Influence of Industry. *BioScience*. 64(10): 917-922. http://dx.doi.org/10.1093/biosci/biu138

Bornstein-Forst, S.M. 2017. Establishing Good Laboratory Practice at Small Colleges and Universities. *Journal of Microbiology & Biology Education*. 18(1): 18.1.10. http://dx.doi.org/10.1128/jmbe.v18i1.1222

Borrell, B. 2009. National Academy as *National Enquirer*? PNAS Publishes Theory That Caterpillars Originated from Interspecies Sex. *Scientific American*. August 24, 2009

Brain, R., J. Stavely, and L. Ortego. 2016. In Response: Resolving the perception of bias in a discipline founded on objectivity—A perspective from industry. *Environmental Toxicology and Chemistry*. 35(5): 1070-1072. http://dx.doi.org/10.1002/etc.3357

Brix, K.V., D.K. DeForest, and W.J. Adams. 2011. The sensitivity of aquatic insects to divalent metals: A comparative analysis of laboratory and field data. *Science of the Total Environment*. 409(20): 4187-4197. http://dx.doi.org/10.1016/j.scitotenv.2011.06.061

Brix, K.V., A.J. Esbaugh, K.M. Munley, and M. Grosell. 2012. Investigations into the mechanism of lead toxicity to the freshwater pulmonate snail, *Lymnaea stagnalis*. *Aquatic Toxicology*. 106–107(0): 147-156. http://dx.doi.org/10.1016/j.aquatox.2011.11.007

Bui, C., R. Parrish, M. Meredith, and J. Giddings, editors. 2004. *A History of SETAC Part 1: The Beginning*. SETAC Press, Pensacola, FL, 36 pp

Burton, G.A., R. Di Giulio, D. Costello, and J.R. Rohr. 2017. Slipping through the Cracks: Why is the U.S. Environmental Protection Agency Not Funding Extramural Research on Chemicals in Our Environment? *Environmental Science & Technology*. 51(2): 755-756. http://dx.doi.org/10.1021/acs.est.6b05877

Burton, G.A. and C.H. Ward. 2012. Editorial preface [letters to the editor on the Exxon Valdez oil spill and pink salmon]. *Environmental Toxicology and Chemistry*. 31(3): 469-469. http://dx.doi.org/10.1002/etc.1741

Buys, D.J., A.R. Stojak, W. Stiteler, and T.F. Baker. 2015. Ecological risk assessment for residual coal fly ash at Watts Bar Reservoir, Tennessee: Limited alteration of riverine-reservoir benthic invertebrate community following dredging of ash-contaminated sediment. *Integrated Environmental Assessment and Management*. 11(1): 43-55. http://dx.doi.org/10.1002/ieam.1577

Calfee, R.D., H.J. Puglis, E.E. Little, W.G. Brumbaugh, and C.A. Mebane. 2016. Quantifying Fish Swimming Behavior in Response to Acute Exposure of Aqueous Copper Using Computer Assisted Video and Digital Image Analysis. *Journal of Visualized Experiments*. (108): e53477. http://dx.doi.org/doi:10.3791/53477

Callaway, E. 2015. Faked peer reviews prompt 64 retractions (18 August 2015). *Nature*. http://dx.doi.org/10.1038/nature.2015.18202. http://dx.doi.org/10.1038/nature.2015.18202

Campbell, T. and J. Friesen. 2015. Why people "fly from facts": Research shows the appeal of untestable beliefs and how they lead to a polarized society (March 3, 2015). *Scientific American*.

Collins, S.L. and J.M. Verdier. 2017. The Coming Era of Open Data. *BioScience*. 67(3): 191-192. http://dx.doi.org/10.1093/biosci/bix023

Cope, M.B. and D.B. Allison. 2009. White hat bias: examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. *International Journal of Obesity*. 34(1): 84-88

Couzin-Frankel, J. 2017. Firing of veteran NIH scientist prompts protests over publication ban (February 22, 2017). *Science*. http://dx.doi.org/10.1126/science.aal0808

Curfman, G.D., S. Morrissey, and J.M. Drazen. 2005. Expression of Concern: Bombardier et al., "Comparison of Upper Gastrointestinal Toxicity of Rofecoxib and Naproxen in Patients with Rheumatoid Arthritis," N Engl J Med 2000;343:1520-8. *New England Journal of Medicine*. 353(26): 2813-2814. http://dx.doi.org/doi:10.1056/NEJMe058314

Davis, P., 2016, *Scientific Reports* On Track To Become Largest Journal In The World (August 23, 2016), *Scholarly Kitchen* at https://scholarlykitchen.sspnet.org/2016/08/23/scientific-reports-on-track-to-become-largest-journal-in-the-world/.

DeForest, D.K. and E.J. Van Genderen. 2012. Application of U.S. EPA guidelines in a bioavailability-based assessment of ambient water quality criteria for zinc in freshwater. *Environmental Toxicology and Chemistry*. 31(6): 1264–1272. http://dx.doi.org/10.1002/etc.1810

Descamps, H. 2008. Natural resource damage assessment (NRDA) under the European Directive on Environmental Liability: a comparative legal point of view. *Océanis*. 32(3/4): 439-462

Dixon, P.M. and J.H.K. Pechmann. 2005. A statistical test to show negligible trend. *Ecology*. 86(7): 1751–1756. http://dx.doi.org/10.1890/04-1343

Doremus, H. 2007. Scientific and Political Integrity in Environmental Policy. *Texas Law Journal*. 86: 1601-1653

Douglas, H. 2015. Politics and Science: Untangling Values, Ideologies, and Reasons. *The ANNALS of the American Academy of Political and Social Science*. 658: 296-306 http://dx.doi.org/10.1177/0002716214557237

Douglas, H.E. 2014. Scientific Integrity in a politicized world. Pages 254-268 *in* P. Schroeder-Heister, G. Heinzmann, W. Hodges, and P. E. Bour, editors. *Logic, Methodology and Philosophy of Science. Proceedings of the Fourteenth International Congress (Nancy, France)*.

Duke, C.S. and J.H. Porter. 2013. The ethics of data sharing and reuse in biology. *BioScience*. 63(6): 483-489. http://dx.doi.org/10.1525/bio.2013.63.6.10

Edwards, M.A. and S. Roy. 2016. Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*. 34(1): 51-61. http://dx.doi.org/10.1089/ees.2016.0223

Elliott, K.C. 2014. Financial conflicts of interest and criteria for research credibility. *Erkenntnis*. 79(5): 917-937. http://dx.doi.org/10.1007/s10670-013-9536-2

Elliott, K.C. 2016. Standardized Study Designs, Value Judgments, and Financial Conflicts of Interest in Research. *Perspectives on Science*. 24(5): 529-551. http://dx.doi.org/10.1162/POSC_a_00222

Enserink, M. 2014. Sabotaged Scientist Sues Yale and Her Lab Chief. *Science*. 343(6175): 1065-1066. http://dx.doi.org/10.1126/science.343.6175.1065

Enserink, M. 2017. A groundbreaking study on the dangers of 'microplastics' may be unraveling (March 21, 2017). *Science*. http://dx.doi.org/10.1126/science.aal0939

European Commission, 2016, Open Access to scientific information | Digital Single Market (October 28, 2016), at https://ec.europa.eu/digital-single-market/en/open-access-scientific-information.

Fairbrother, A. 2016. New Subcommittee on Scientific Integrity (18 February 2016). *SETAC Globe*. 17(2)

Fairbrother, A. and R.S. Bennett. 1999. Ecological Risk Assessment and the Precautionary Principle. *Human and Ecological Risk Assessment: An International Journal*. 5(5): 943-949. http://dx.doi.org/10.1080/10807039991289220

Fanelli, D. 2009. How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*. 4(5): e5738. http://dx.doi.org/10.1371/journal.pone.0005738

Fanelli, D. 2012. Negative results are disappearing from most disciplines and countries. *Scientometrics*. 90(3): 891-904. http://dx.doi.org/10.1007/s11192-011-0494-7

Farris, J.L. and J.H.V. Hassel, editors. 2006. *Freshwater Bivalve Ecotoxicology*. SETAC and CRC Press, 408 pp

Fenn, D.B. and N.M. Milton. 1997. Advocacy and the Scientist. *Fisheries*. 22(11): 4

Folta, K.M., 2015, Transparency Weaponized Against Scientists (August 16, 2015): Science 2.0, at http://www.science20.com/kevin_folta/transparency_weaponized_against_scientists-156873.

Geist, D.R., C.S. Abernethy, K.D. Hand, V.I. Cullinan, J.A. Chandler, and P.A. Groves. 2006. Survival, development, and growth of fall Chinook salmon embryos, alevin, and fry exposed to a variable thermal and dissolved oxygen regime. *Transactions of the American Fisheries Society*. 135(6): 1462–1477. http://dx.doi.org/10.1577/T05-294.1

Ghorayshi, A. 2016. "He thinks he's untouchable"-Sexual Harassment Case Exposes Renowned Ebola Scientist (June 29, 2016). BuzzFeedNews. https://www.buzzfeed.com/azeenghorayshi/michael-katze-investigation.

Gibbons, A. 2014. Sexual harassment is common in scientific fieldwork (July 16, 2014). *Science*.

Glanz, J. and A. Armendariz, 2017, Years of Ethics Charges, but Star Cancer Researcher Gets a Pass (March 8, 2017): New York Times, at https://www.nytimes.com/2017/03/08/science/cancer-carlo-croce.html.

Gleick, P.H. and 252 co-authors. 2010. Climate Change and the Integrity of Science. *Science*. 328(5979): 689-690. http://dx.doi.org/10.1126/science.328.5979.689

Goodstein, D. 1995. Conduct and Misconduct in Science. *Annals of the New York Academy of Sciences*. 775(1): 31-38. http://dx.doi.org/10.1111/j.1749-6632.1996.tb23124.x

Halpern, M. and M. Mann. 2015. Transparency versus harassment. *Science*. 348(6234): 479-479. http://dx.doi.org/10.1126/science.aac4245

Harris, C.A., A.P. Scott, A.C. Johnson, G.H. Panter, D. Sheahan, M. Roberts, and J.P. Sumpter. 2014. Principles of Sound Ecotoxicology. *Environmental Science & Technology*. 48(6): 3100-3111. http://dx.doi.org/10.1021/es4047507

Harris, C.A. and J.P. Sumpter. 2015. Could the Quality of Published Ecotoxicological Research Be Better? *Environmental Science & Technology*. 49(16): 9495-9496. http://dx.doi.org/10.1021/acs.est.5b01465

Harris, D. and M. Patrick, 2011, Is 'Big Food's' Big Money Influencing the Science of Nutrition? (June 21,2011): ABC News, at http://abcnews.go.com/US/big-food-money-accused-influencing-science/story?id=13845186.

Hayes, T.B. 2004. There Is No Denying This: Defusing the Confusion about Atrazine. *BioScience*. 54(12): 1138-1149. http://dx.doi.org/10.1641/0006-3568(2004)054[1138:TINDTD]2.0.CO;2

Henderson, D. and K. Thomson. 2017. Why Should Scientific Results Be Reproducible? (NOVA, WGBH, 15 minute video, released 19 January 2017). WGBH, NOVA. http://www.pbs.org/wgbh/nova/next/body/reproducibility-explainer/.

Hey, E. and I. Chalmers. 2010. Mis-investigating alleged research misconduct can cause widespread, unpredictable damage. *Journal of the Royal Society of Medicine*. 103(4): 133-138. http://dx.doi.org/10.1258/jrsm.2010.09k045

Hobbs, D.A., M.S.J. Warne, and S.J. Markich. 2005. Evaluation of criteria used to assess the quality of aquatic toxicity data. *Integrated Environmental Assessment and Management*. 1(3): 174–180. http://dx.doi.org/10.1897/2004-003R.1

Holdren, J.P. 2013. Increasing Access to the Results of Federally Funded Scientific Research. Memorandum for the heads of executive departments and agencies (February 22, 2013). Executive Office of the President, Office of Science and Technology Policy. , Washington, D.C. 6 pp.

Hopkin, M. 2006. Ecology: Caught between shores. *Nature*. 440(7081): 144-145. http://dx.doi.org/10.1038/440144a

Huff, D. and I. Geis. 1954. How to lie with statistics. W.W. Norton, New York and London. (1993 reissue). 142 pp

Hurlbert, S.H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*. 54(2): 187-211. http://dx.doi.org/10.2307/1942661

Hutchings, J.A. 1997. Is scientific inquiry incompatible with government information control? *Canadian Journal of Fisheries and Aquatic Sciences*. 54(5): 1198-1210. http://dx.doi.org/10.1139/f97-051

ICMJE. 2016. ICMJE Form for Disclosure of Potential Conflicts of Interest. International Committee of Medical Journal Editors. 3 pp. http://icmje.org/conflicts-of-interest/.

Ioannidis, J.P.A. 2005. Why Most Published Research Findings Are False. *PLoS Med*. 2(8): e124. http://dx.doi.org/10.1371/journal.pmed.0020124

Johnson, A.C., R.L. Donnachie, J.P. Sumpter, M.D. Jürgens, C. Moeckel, and M.G. Pereira. 2017. An alternative approach to risk rank chemicals on the threat they pose to the aquatic environment. *Science of the Total Environment*. 599-600: 1372-1381. http://dx.doi.org/10.1016/j.scitotenv.2017.05.039

Johnson, A.C. and J.P. Sumpter. 2016. Are we going about chemical risk assessment for the aquatic environment the wrong way? *Environmental Toxicology and Chemistry*. 35(7): 1609-1616. http://dx.doi.org/10.1002/etc.3441

Johnson, B.L. 2007. Intimidation of Scientists: A HERA Experience. *Human and Ecological Risk Assessment: An International Journal*. 13(3): 475-477. http://dx.doi.org/10.1080/10807030701340870

Johnson, D.R. and E.H. Ecklund. 2016. Ethical ambiguity in science. *Science and Engineering Ethics*. 22(4): 989-1005. http://dx.doi.org/10.1007/s11948-015-9682-9

Kapustka, L. 2016. Words matter. *Integrated Environmental Assessment and Management*. 12(3): 592-593. http://dx.doi.org/10.1002/ieam.1767

Keith, R., 2015, 13th retraction issued for Jesús Ángel Lemus (October 9, 2015), *Retraction Watch* at http://retractionwatch.com/2015/10/09/13th-retraction-issued-for-jesus-angel-lemus/.

Kidwell, M.C., L.B. Lazarević, E. Baranski, T.E. Hardwicke, S. Piechowski, L.-S. Falkenberg, C. Kennett, A. Slowik, C. Sonnleitner, C. Hess-Holden, T.M. Errington, S. Fiedler, and B.A. Nosek. 2016. Badges to Acknowledge Open Practices: A Simple, Low-Cost, Effective Method for Increasing Transparency. *PLOS Biology*. 14(5): e1002456. http://dx.doi.org/10.1371/journal.pbio.1002456

Kloor, K. 2015. Agricultural researchers rattled by demands for documents from group opposed to GM foods. *Science Insider*. . http://dx.doi.org/10.1126/science.aaa7846

Kollipara, P. 2015. Open records laws becoming vehicle for harassing academic researchers, report warns (Feb. 13, 2015). *Science*. http://dx.doi.org/10.1126/science.aaa7856

Kolok, A.S. 2016. The dose makes the poison. Pages 1-9 *in Modern Poisons: A Brief Introduction to Contemporary Toxicology*. Springer International Publishing.

Krimsky, S. 2005. The Funding Effect in Science and Its Implications for the Judiciary. *Journal of Law and Policy*. 13(1): 43-68

Krumholz, H.M., J.S. Ross, A.H. Presler, and D.S. Egilman. 2007. What have we learnt from Vioxx? *BMJ : British Medical Journal*. 334(7585): 120-123. http://dx.doi.org/10.1136/bmj.39024.487720.68

Lackey, R.T. 2001. Values, Policy, and Ecosystem Health. *BioScience*. 51(6): 437-443. http://dx.doi.org/10.1641/0006-3568(2001)051[0437:VPAEH]2.0.CO;2

Lackey, R.T. 2007. Science, scientists, and policy advocacy. *Conservation Biology*. 21(1): 12-17. http://dx.doi.org/10.1111/j.1523-1739.2006.00639.x

Laine, H. 2017. Afraid of Scooping; Case Study on Researcher Strategies against Fear of Scooping in the Context of Open Science. *Data Science Journal*. 16: 29. http://dx.doi.org/10.5334/dsj-2017-029

Levy, K.E. and D.M. Johns. 2016. When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society*. 3(1). http://dx.doi.org/10.1177/2053951715621568

Lewandowsky, S. and D. Bishop. 2016. Research integrity: Don't let transparency damage science. *Nature*. 529: 459-461. http://dx.doi.org/10.1038/529459a

Lexchin, J., L.A. Bero, B. Djulbegovic, and O. Clark. 2003. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 326(7400): 1167-1170. http://dx.doi.org/10.1136/bmj.326.7400.1167

Lindenmayer, D.B. and G.E. Likens. 2010. The science and application of ecological monitoring. *Biological Conservation*. 143(6): 1317-1328. http://dx.doi.org/10.1016/j.biocon.2010.02.013

Macleod, M. 2014. Some salt with your statin, Professor? *PLOS Biology*. 12(1): e1001768. http://dx.doi.org/10.1371/journal.pbio.1001768

Marcus, A. and I. Oransky. 2016. CRISPR controversy reveals how badly journals handle conflicts of interest (January 21, 2016). *STAT*. https://www.statnews.com/2016/01/21/crispr-conflicts-of-interest/

Martinson, B.C., M.S. Anderson, and R. de Vries. 2005. Scientists behaving badly. *Nature*. 435(7043): 737-738. http://dx.doi.org/10.1038/435737a

McClellan, F. 2008. The Vioxx Litigation: A Critical Look at Trial Tactics, the Tort System, and the Roles of Lawyers in Mass Tort Litigation. *DePaul Law Review*. 57: 509–538

McGarity, T.O. 2003. Our Science Is Sound Science and Their Science Is Junk Science: Science-Based Strategies for Avoiding Accountability and Responsibility for Risk-Producing Products and Activities. *Kansas Law Review*. 897-937

McGarity, T.O. and W.E. Wagner. 2012. Bending Science: How Special Interests Corrupt Public Health Research. Harvard University Press, Cambridge, MA

McNutt, M. 2016. #IAmAResearchParasite. *Science*. 351(6277): 1005-1005. http://dx.doi.org/10.1126/science.aaf4701

McNutt, M., K. Lehnert, B. Hanson, B.A. Nosek, A.M. Ellison, and J.L. King. 2016. Liberating field science samples and data. *Science*. 351(6277): 1024-1026. http://dx.doi.org/10.1126/science.aad7048

Mebane, C.A., R.J. Eakins, B.G. Fraser, and W.J. Adams. 2015. Recovery of a mining-damaged stream ecosystem. *Elementa: Science of the Anthropocene*. 3(1): 000042. http://dx.doi.org/10.12952/journal.elementa.000042

Mebane, C.A. and J.S. Meyer. 2016. Environmental Toxicology without Chemistry and Publications without Discourse: Linked Impediments to Better Science. *Environmental Toxicology and Chemistry*. 35(6): 1335–1336. http://dx.doi.org/10.1002/etc.3418

Melvin, S.D., K.R. Munkittrick, T. Bosker, and D.L. MacLatchy. 2009. Detectable effect size and bioassay power of mummichog (*Fundulus heteroclitus*) and fathead minnow (*Pimephales promelas*) adult reproductive tests. *Environmental Toxicology and Chemistry*. 28(11): 2416–2425 http://dx.doi.org/10.1897/08-601.1

Meyer, J.N. and A.B. Francisco. 2013. A call for fuller reporting of toxicity test data. *Integrated Environmental Assessment and Management*. 9(2): 347-348. http://dx.doi.org/10.1002/ieam.1406

Mozur, M.C. 2012. A global professional society. *Integrated Environmental Assessment and Management*. 8(1): 1-1. http://dx.doi.org/10.1002/ieam.1270

Mudge, J.F., T.J. Barrett, K.R. Munkittrick, and J.E. Houlahan. 2012. Negative Consequences of Using $\alpha = 0.05$ for Environmental Monitoring Decisions: A Case Study from a Decade of Canada's Environmental Effects Monitoring Program. *Environmental Science & Technology*. 46(17): 9249-9255. http://dx.doi.org/10.1021/es301320n

Munkittrick, K.R., C.J. Arens, R.B. Lowell, and G.P. Kaminski. 2009. A review of potential methods for determining critical effect size for designing environmental monitoring programs. *Environmental Toxicology and Chemistry*. 28(7): 1361-1371. http://dx.doi.org/10.1897/08-376.1

Murray, K.R., S.J. McCardell, and J.R. Schofield. 1999. Natural Resource Damage Trustees: Whose Side Are They Really On? *Environmental Lawyer*. 5: 407-467

NAS. 1992. Responsible Science: Ensuring the Integrity of the Research Process. National Academy of Sciences, National Academy Press, Washington, D.C. https://doi.org/10.17226/1864.

NAS. 2017. Fostering Integrity in Research. National Academy of Sciences, National Academy Press, Washington, D.C. 284 pp. https://doi.org/10.17226/21896.

Nature Editors. 2014. Why high-profile journals have more retractions (17 September 2014). *Nature*. http://dx.doi.org/10.1038/nature.2014.15951

Nature Editors. 2018. Nature journals tighten rules on non-financial conflicts. *Nature*. 554(6). http://dx.doi.org/10.1038/d41586-018-01420-8

Nature Medicine Editors. 2017. Take science off the stand. *Nature Medicine*. 23(3): 265-265. http://dx.doi.org/10.1038/nm.4303

Nelson, B. 2009. Data sharing: Empty archives. *Nature*. 461: 160-163. http://dx.doi.org/10.1038/461160a

Norton, S.B., S.M. Cormier, and G.W. Suter, II. 2002. The easiest person to fool. *Environmental Toxicology and Chemistry*. 21(6): 1099-1100. http://dx.doi.org/10.1002/etc.5620210601

Nosek, B.A. and 39 co-authors. 2015. Promoting an open research culture. *Science*. 348(6242): 1422-1425. http://dx.doi.org/10.1126/science.aab2374

Nosek, B.A. and T.M. Errington. 2017. Making sense of replications. *eLife*. 6: e23383. http://dx.doi.org/10.7554/eLife.23383

NRC-CNRC. 2013. NRC's Research Integrity Policy. National Research Council Canada. 28 pp. http://www.nrc-cnrc.gc.ca/eng/about/policies/research_integrity/.

NRC. 2002. Integrity in Scientific Research: Creating an Environment That Promotes Responsible Conduct. National Research Council (US), National Academy Press, Washington, D.C.

Nuzzo, R. 2015. How scientists fool themselves – and how they can stop. *Nature*. 526: 182–185 (08 October 2015). http://dx.doi.org/10.1038/526182a

Obama, B. 2009. Scientific Integrity (Presidential documents, Memorandum of March 9, 2009). *Federal Register*. 74(46): 10671-10672

Ogden, L.E. 2016. Nine years of censorship. *Nature*. 533 26–28. http://dx.doi.org/10.1038/533026a

Oransky, I. 2015. How Publish or Perish Promotes Inaccuracy in Science—and Journalism. *AMA Journal of Ethics*. 17(12): 1172-1175. http://dx.doi.org/10.1001/journalofethics.2015.17.12.sect1-1512

Oreskes, N., D. Carlat, M.E. Mann, P.D. Thacker, and F.S. vom Saal. 2015. Viewpoint: Why Disclosure Matters. *Environmental Science & Technology*. 49(13): 7527-7528. http://dx.doi.org/10.1021/acs.est.5b02726

Oreskes, N. and E.M. Conway. 2010. Defeating the merchants of doubt. *Nature*. 465(7299): 686-687. http://dx.doi.org/10.1038/465686a

Owen, S.F., D.B. Huggett, T.H. Hutchinson, M.J. Hetheridge, P. McCormack, L.B. Kinter, J.F. Ericson, L.A. Constantine, and J.P. Sumpter. 2010. The value of repeating studies and multiple controls: replicated 28-day growth studies of rainbow trout exposed to clofibric acid. *Environmental Toxicology and Chemistry*. 29(12): 2831-2839. http://dx.doi.org/10.1002/etc.351

Parker, K.R. and J.A. Wiens. 2005. Assessing recovery following environmental accidents: environmental variation, ecological assumptions, and strategies. *Ecological Applications*. 15(6): 2037-2051. http://dx.doi.org/10.1890/04-1723

Parker, T.H., W. Forstmeier, J. Koricheva, F. Fidler, J.D. Hadfield, Y.E. Chee, C.D. Kelly, J. Gurevitch, and S. Nakagawa. 2016. Transparency in Ecology and Evolution: Real Problems, Real Solutions. *Trends in Ecology & Evolution*. 31(9): 711-719. http://dx.doi.org/10.1016/j.tree.2016.07.002

Peterman, R.M. and M. M'Gonigle. 1992. Statistical power analysis and the precautionary principle. *Marine Pollution Bulletin*. 24(5): 231-234. http://dx.doi.org/10.1016/0025-326X(92)90559-O

PLOS, 2014, Data Availability, accessed March 2, 2017 at http://journals.plos.org/plosone/s/data-availability.

PLOS Medicine Editors. 2008. Making Sense of Non-Financial Competing Interests. *PLoS Med*. 5(9): e199. http://dx.doi.org/10.1371/journal.pmed.0050199

Power, M., G. Power, and D.G. Dixon. 1995. Detection and decision-making in environmental effects monitoring. *Environmental Management*. 19(5): 629-639. http://dx.doi.org/10.1007/BF02471945

Rekdal, O.B. 2014. Academic urban legends. *Social Studies of Science*. 44(4): 638-654. http://dx.doi.org/10.1177/0306312714535679

Renner, R. 2005. Proposed selenium standard under attack. *Environmental Science & Technology*. 39(6): 125A-126A. http://dx.doi.org/10.1021/es053215n

Resnik, D.B., T. Neal, A. Raymond, and G.E. Kissling. 2015. Research Misconduct Definitions Adopted by U.S. Research Institutions. *Accountability in Research*. 22(1): 14-21. http://dx.doi.org/10.1080/08989621.2014.891943

Resnik, D.B. and A.E. Shamoo. 2011. The Singapore Statement on Research Integrity. *Accountability in Research*. 18(2): 71-75. http://dx.doi.org/10.1080/08989621.2011.557296

Roberts, P.D., G.B. Stewart, and A.S. Pullin. 2006. Are review articles a reliable source of evidence to support conservation and environmental management? A comparison with medicine. *Biological Conservation*. 132(4): 409-423. http://dx.doi.org/10.1016/j.biocon.2006.04.034

Ross, T. 2017. Crime in the cancer lab (January 27, 2017). New York Times. https://www.nytimes.com/2017/01/28/opinion/sunday/a-crime-in-the-cancer-lab.html?_r=1.

Ruff, K. 2015. Scientific journals and conflict of interest disclosure: what progress has been made? *Environmental Health*. 14(1): 1-8. http://dx.doi.org/10.1186/s12940-015-0035-6

Schäfer, R.B., M. Bundschuh, A. Focks, and P.C. von der Ohe. 2013. To the Editor [authors should make all their raw data accessible]. *Environmental Toxicology and Chemistry*. 32(4): 734-735. http://dx.doi.org/10.1002/etc.2140

Schindler, D.W. 1998. Replication versus realism: The need for ecosystem-scale experiments. *Ecosystems*. 1(4): 323-334. http://dx.doi.org/10.1007/s100219900026

Scott, J.M. and J.L. Rachlow. 2011. Refocusing the Debate about Advocacy. *Conservation Biology*. 25(1): 1-3. http://dx.doi.org/10.1111/j.1523-1739.2010.01629.x

Scott, J.M., J.L. Rachlow, R.T. Lackey, A.B. Pidgorna, J.L. Aycrigg, G.R. Feldman, L.K. Svancara, D.A. Rupp, D.I. Stanish, and R.K. Steinhorst. 2007. Policy Advocacy in Science: Prevalence, Perspectives, and Implications for Conservation Biologists. *Conservation Biology*. 21(1): 29-35. http://dx.doi.org/10.1111/j.1523-1739.2006.00641.x

Sedlak, D. 2016. Crossing The Imaginary Line. *Environmental Science & Technology*. 50(18): 9803-9804. http://dx.doi.org/10.1021/acs.est.6b04432

Shaw, D. and P. Satalkar. 2018. Researchers' interpretations of research integrity: A qualitative study. *Accountability in Research*. 1-15. http://dx.doi.org/10.1080/08989621.2017.1413940

Shrader-Frechette, K. 2012. Research Integrity and Conflicts of Interest: The Case of Unethical Research-Misconduct Charges Filed by Edward Calabrese. *Accountability in Research*. 19(4): 220-242. http://dx.doi.org/10.1080/08989621.2012.700882

Skorupa, J.P., T.S. Presser, S.J. Hamilton, A.D. Lemly, and B.E. Sample. 2004. EPA's Draft Tissue-Based Selenium Criterion: A Technical Review. U.S. Fish and Wildlife Service, U.S. Geological Survey, U.S. Forest Service and CH2M Hill. 35 pp. https://wwwrcamnl.wr.usgs.gov/Selenium/library.htm

Smaldino, P.E. and R. McElreath. 2016. The natural selection of bad science. *Royal Society Open Science*. 3(9). http://dx.doi.org/10.1098/rsos.160384

Smith, R. 2006. Conflicts of interest: how money clouds objectivity. *Journal of the Royal Society of Medicine*. 99(6): 292-297

Smith, R. and I. Roberts. 2016. Time for sharing data to become routine: the seven excuses for not doing so are all invalid [version 1; referees: 2 approved, 1 approved with reservations]. *F1000Research*. 5(781). http://dx.doi.org/10.12688/f1000research.8422.1

Snyder, B.F. and L.M. Hooper-Bui. 2018. A Teachable Moment: The Relevance of Ethics and the Limits of Science. *BioScience*. bix165-bix165. http://dx.doi.org/10.1093/biosci/bix165

Solomon, K.R., J.A. Carr, L.H. Du Preez, J.P. Giesy, R.J. Kendall, E.E. Smith, and G.J. Van Der Kraak. 2008. Effects of Atrazine on Fish, Amphibians, and Aquatic Reptiles: A Critical Review. *Critical Reviews in Toxicology*. 38(9): 721-772. http://dx.doi.org/10.1080/10408440802116496

Stedeford, T. 2007. Prior restraint and censorship: acknowledged occupational hazards for government scientists. *William and Mary Environmental Law and Policy Review*. 31(3): 725–745.

Stein, R. and J. Eilperin. 2010. Obama administration issues guidelines designed to ensure 'scientific integrity' (December 17, 2010). *Washington Post*.

Steneck, N.H. 2006. Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*. 12(1): 53-74. http://dx.doi.org/10.1007/PL00022268

Stern, V. 2017. Updated: Why would a university pay a scientist found guilty of misconduct to leave? (2 October 2017). *Science*. http://dx.doi.org/10.1126/science.aan7182

Stokstad, E. 2012. Fracking Report Criticized for Apparent Conflict of Interest (July 24, 2012). Science. http://www.sciencemag.org/news/2012/07/fracking-report-criticized-apparent-conflict-interest.

Sumner, P., S. Vivian-Griffiths, J. Boivin, A. Williams, C.A. Venetis, A. Davies, J. Ogden, L. Whelan, B. Hughes, B. Dalton, F. Boy, and C.D. Chambers. 2014. The association between exaggeration in health related science news and academic press releases: retrospective observational study. *BMJ : British Medical Journal*. 349. http://dx.doi.org/10.1136/bmj.g7015

Suter, G.W., II and S.M. Cormier. 2015a. Bias in the development of health and ecological assessments and potential solutions. *Human and Ecological Risk Assessment*. 22(1): 99-115. http://dx.doi.org/10.1080/10807039.2015.1056062

Suter, G.W., II and S.M. Cormier. 2015b. The Problem of Biased Data and Potential Solutions for Health and Environmental Assessments. *Human and Ecological Risk Assessment*. 21(7): 1736-1752. http://dx.doi.org/10.1080/10807039.2014.974499

Suter, G.W., II., S.B. Norton, and S.M. Cormier. 2002. A method for inferring the causes of observed impairments in aquatic ecosystems. *Environmental Toxicology and Chemistry*. 21(6): 1101-1111. http://dx.doi.org/10.1002/etc.5620210602

Tollefson, J. 2015. Earth science wrestles with conflict-of-interest policies. *Nature*. 522: 403-404. http://dx.doi.org/10.1038/522403a

Topf, A., 2016, Imperial Metals sues engineering companies over Mount Polley disaster (July 10, 2016): Mining.com, accessed February 7, 2017 at http://www.mining.com/imperial-metals-sues-engineering-companies-mount-polley-disaster/.

U.S. Department of Interior, 2014, Code of Scientific and Scholarly Conduct (poster), at https://www.doi.gov/scientificintegrity/upload/DOI-Code-of-Scientific-and-Scholarly-Conduct-Poster-December-2014.pdf.

USEPA. 2013. Aquatic Life Ambient Water Quality Criteria For Ammonia – Freshwater 2013. EPA 822-R-13-001. 255 pp. http://water.epa.gov/scitech/swguidance/standards/criteria/aqlife/ammonia/index.cfm.

van Iersel, S., E.M. Swart, Y. Nakadera, N.M. van Straalen, and J.M. Koene. 2014. Effect of Male Accessory Gland Products on Egg Laying in Gastropod Molluscs. *Journal of Visualized Experiments*. 88: e51698. http://dx.doi.org/10.3791/51698

Van Kirk, R.W. and S.L. Hill. 2007. Demographic model predicts trout population response to selenium based on individual-level toxicity. *Ecological Modelling*. 206(3-4): 407-420. http://dx.doi.org/10.1016/j.ecolmodel.2007.04.003

van Kolfschooten, F. 2002. Conflicts of interest: Can you believe what you read? *Nature*. 416(6879): 360-363

Van Noorden, R. 2014. Publishers withdraw more than 120 gibberish papers (24 February 2014). *Nature*. http://dx.doi.org/10.1038/nature.2014.14763

Vines, T.H., A.Y.K. Albert, R.L. Andrew, F. Débarre, D.G. Bock, M.T. Franklin, K.J. Gilbert, J.-S. Moore, S. Renaut, and D.J. Rennison. 2014. The availability of research data declines rapidly with article age. *Current Biology*. 24(1): 94-97. http://dx.doi.org/http://dx.doi.org/10.1016/j.cub.2013.11.014

Wadman, M. 1997. $100m payout after drug data withheld. *Nature*. 388(6644): 703-703

Wagner, W. 2005. The perils of relying on interested parties to evaluate scientific quality. *American Journal of Public Health*. 95(S1): S99-S106. http://dx.doi.org/10.2105/AJPH.2004.044792

Wagner, W.E. and D. Michaels. 2004. Equal Treatment for Regulatory Science: Extending the Controls Governing the Quality of Public Research to Private Research. *American Journal of Law & Medicine*. 30(2-3): 119-154. http://dx.doi.org/10.1177/009885880403000202

Wiens, J.A. and K.R. Parker. 1995. Analyzing the effects of accidental environmental impacts: approaches and assumptions. *Ecological Applications*. 5(4): 1069-1083. http://dx.doi.org/10.2307/2269355

Wise, J. 1997. Research suppressed for seven years by drug company. *BMJ : British Medical Journal*. 314(7088): 1145-1145. http://dx.doi.org/10.1136/bmj.314.7088.1145

Womack, R.P. 2015. Research Data in Core Journals in Biology, Chemistry, Mathematics, and Physics. *PLoS ONE*. 10(12): e0143460. http://dx.doi.org/10.1371/journal.pone.0143460

Young, N.S., J.P.A. Ioannidis, and O. Al-Ubaydli. 2008. Why Current Publication Practices May Distort Science. *PLOS Medicine*. 5(10): e201. http://dx.doi.org/10.1371/journal.pmed.0050201