

11 July 2017.

OK, I have finally found the stomach to return here. I've been away, cavorting in Alaska for 3+ weeks, just away from "it all." Bike touring for a week from Skagway to Haines, recording birds, sitting among grizzlies . . . But, here goes. I'll comment on the Reviews of my last attempt to address the critics of my Forum article. My thoughts will be indented, prefaced by "DK:"

A few hours later, I return here. I've made a substantial number of changes, as can be seen in the "track changes" feature of Word. I have accepted many reviewer recommendations outright and deleted passages, and have modified other passages to try to comply with the intent of the comments.

If there are any attachments uploaded by the reviewers, you may view these by logging in as an author and checking the folder named "Submissions Needing Revision" and in the Action column select "View Reviewer Attachments".

DK: I hope there are none. I see no indication of such.

PLEASE NOTE: The journal would like to enrich online articles by visualising and providing geographical details described in Animal Behaviour articles. For this purpose, corresponding KML (GoogleMaps) files can be uploaded in our online submission system. Submitted KML files will be published with your online article on ScienceDirect. Elsevier will generate maps from the KML files and include them in the online article.

DK: NA

Animal Behaviour features the Interactive Map Viewer, <http://www.elsevier.com/googlemaps>, allowing you to visualize geospatial data with your article in a GoogleMap. Our author tool, <http://elsevier-apps.sciverse.com/GoogleMaps/verification>, enables creation of simple Interactive Map files by typing coordinates and a description, and choosing place markers. For more advanced maps, a GIS system or Google Earth can be used. For those, the author tool can be used for inspecting how the file will display online. In both cases, the created .kml/.kmz file should be uploaded along with the manuscript in the supplementary files section

Date Revision Due Jul 16, 2017

Dear Don,

DK: Not easy to feel very "dear" after 3 years of this stuff.

Here are my comments and comments from Rev. 2 (the Rev. 2 of the previous version) of your reply to Podos et al. I trust you will find them helpful in your final revision. Again, I have one

general point to make, and it is one on which Rev. 2 and I appear to be in complete agreement. And it is that you benefit by making these comments as impersonal as you can, i.e., criticism should be directed at the science, not the person. Your feelings about the people are well known, and need no additional emphasis. The target audience really is made up of uncommitted researchers, and researchers down the road, who will be interested in the scientific questions, not the politics. These are the folks who will contribute to progress in this field by careful research and dispassionate reviews of the literature. I hope you will find persuasive the points of consensus between the reviewer and myself - and there are many - given that the two of us wrote our comments entirely independently.

DK: I will try ☺

My comments:

Page 1. Delete the sentence: "What is especially disturbing about this literature is that those who author, promote, and defend it are among the perceived leaders of the Animal Behavior Society: three presidents, two recipients of the Exemplar Award, and one recipient of the Young Investigator Award". It amounts to an ad hominem attack on Podos, Nowicki and Searcy, as well as on the Animal Behavior Society and as well on the members of ABS who voted them in as President or nominated them for these awards. Note that apart from the presence of the Animal Behaviour Executive Editor on the ABS Executive Committee, there is no real connection between the Journal Editorial Office and the Society. Think about it: if there were, this commentary likely would not have been reviewed at all given the present leadership of ABS. We need to continue the approach of criticizing the science, or the writing, not the scientists/writers (except, of course, by implication).

DK: My goal here is nothing personal. My goal is an attempt to point out that publishing this kind of nonsense is a winning strategy for a career (evidence: awards by ABS, presidents of ABS), and that is disturbing. I see nothing ad hominem about my goal. And I don't see the "think about it" comment as really relevant. My statement stands whether ABS and the journal are linked or not; I do realize they are independent, as has been pointed out by Susan Foster. I'm not going to fight this. The thoughts behind this sentence probably belong in a separate commentary on how one succeeds in (non)science. . . . OK, I'VE DELETED THE SENTENCE (and adjusted the previous one, as per the referee). The point is covered sufficiently elsewhere

Page 2. Delete the clause "resorting to defense of the indefensible" so that the sentence becomes simply "Podos distracts from the big issues largely by addressing minor, secondary issues or nonissues altogether". In my opinion this becomes a stronger criticism with the distracting overkill removed.

DK: Ok, I'll try that.

Not sure that the litany of Vehrencamp's negative adjectives gets you anywhere. If you are going to say something about her commentary, you should make it parallel to your criticism of Podos; that is, you need to describe the weaknesses it has, rather than simply list the unpleasant things said about your manuscript. After all, even from the highly edited final version of your commentary, I am sure one could put together a similar long string of strong adjectives. Perhaps it would suffice if you just lumped Vehrencamp with Podos, since in my reading, they both fail to fully address your criticisms.

DK: I had considered deleting this paragraph altogether, but I like your suggestion of lumping Vehrencamp with Podos. But it does pain me somewhat not to point out what a silly opening paragraph Vehrencamp has offered.

Concerning the word 'performance'. I personally agree that it is generally instructive to read one of these papers with 'performance' replaced with a neutral word of one's choice. But you just say this, rather than demonstrating your point. And the paper you cite at this point, Podos et al 2016, is probably not a very good choice. For example, here is what I get when I substitute 'deviation' for 'performance' in the abstract of that paper (I also shorten the second sentence of the abstract):

"Our understanding of the evolution and function of animal displays has been advanced through studies of vocal deviation. Vocal deviation is limited by being applicable only to vocal trills, and also overlooks certain fine-scale aspects of song structure that might reflect vocal deviation. In light of these limitations we here introduce a new index of vocal deviation, 'frequency excursion'. Frequency excursion calculates, for any given song or song segment, the sum of frequency modulations both within and between notes on a per-time basis. We calculated and compared the two deviation metrics in three species: chipping sparrows, swamp sparrows, and song sparrows. The two metrics correlated as expected, yet frequency excursion accounted for subtle variations in deviation overlooked by vocal deviation. In swamp sparrows, frequency excursion values varied significantly by song type but not by individual. Moreover, song type deviation in swamp sparrows, according to both metrics, varied negatively with the extent to which song types were shared among neighbours. In song sparrows, frequency excursion values of trilled song segments exceeded those of nontrilled song segments, although not to a statistically significant degree. We suggest that application of frequency excursion in birds and other taxa will provide new insights into diverse open questions concerning vocal deviation, function and evolution".

It does lose something, but because of the topic, not as much as some other example might. So if you want to make this point again here, you should take a better example and provide the quote with the substitution.

DK: Well, I'm going to be a bit lazy here. I'm going to modify the first sentence in this section and then the first sentence of the second paragraph, all without providing a quote that substitutes neutral words.

You go on to say that following such a substitution we would get "a good description [which] will last forever and would contribute more to our understanding of the natural world than all of the performance experiments I have critiqued". I totally agree that we should always start with good description, and that this line of research would have benefited from such an approach, but what you are saying literally here is that this word substitution will turn (what you think are) very bad experiments into good description. You don't really mean that, right? so you need to clarify your point here.

DK: No, I certainly don't mean that bad experiments become good with word substitution.

You get more specific about Cardoso and Atwell 2016, which is good, but I found this section confusing. they report that 'shared' songs are of lower performance (lower vocal deviation I presume). I haven't read their paper but of course your summary should not assume the reader has read it, and so should give enough detail that the reader can follow. You rightly question use of the term 'shared' and suggest that simpler syllables with fast trill rates would be more likely to be classified as 'shared' because they have less detail that would enable the classifier to call them 'unshared' (different). You predict "a bias in their analyses with simpler syllable with fast trill rates would be more likely to be classified as shared". but this is confusing because fast trill rates are 'high performance' so how is that they find that 'shared' songs are more likely to be 'low performance'? In other words, you may be right about the (misleading) criterion they use for their 'shared' classification, but the result appears to be the opposite of what they find. I obviously don't get something here, so assuming I am the average reader, this means you need to provide more information here than you have.

DK: Yes, astute of you to notice that. I think that what it boils down to is this: The Cardoso method of eyeballing frequency range, together with the nothingness of vocal deviation in the first place, gives nothing predictable, unless one is so convinced from the outset of what the results should be that the results come to be what they "should be." I could go into that with more explanation, or just leave it as it is. I'm tempted to just leave it as it is and let astute readers figure it out for themselves. I think I've said enough.

Page 4. You conclude your discussion of Cardoso and Atwell (2016) with this statement: "By cutting through the performance verbiage ... I came to doubt that the classification of songs as shared or unshared has any relevance to the birds themselves...much less to sexual selection or any index of song performance". But changing a word can't in itself tell us anything how the birds perceive these stimuli. That is an empirical question that can only be answered by putting the question to the birds, i.e. by doing experiments in which the birds respond to songs classified as 'high performance' or 'low performance'. The substitution exercise can make you personally come to doubt their conclusion, but the real issue is that they imply, without sufficient basis, that these differences are perceived. It is a secondary issue that you, not being convinced by the word magic, continue to doubt it (and are upset that other folks seem to be swayed by the word magic).

But just putting your conviction ('I doubt it') against their implied conviction ('I don't doubt it, and furthermore my friends don't doubt it') is not productive. Again that makes it about what people are thinking at the moment, not about the eternal truth of what the science is designed to uncover. That is, do not focus on the fact that you are not persuaded, but instead on the fact that there is no direct evidence for - and in most cases not even a direct test of - the performance hypotheses (that birds try to develop the highest performance songs they can, and that listeners respond differentially to high vs low performance songs). As I have said all along, the point you want to make is not that you personally are convinced that the emperor has no clothes on, nor that you are outraged that so many people fail to see this, but instead that people should just use their senses (read the performance papers carefully) so that they will see that in fact the emperor has no clothes on.

DK: Maybe my wording changes in the last sentence of this section helps. Maybe. Maybe not.

Page 5. "Science is the search for truths about the natural world, but when the chosen methods cannot reveal truths, science suffers, as does the credibility of all scientists in all disciplines, including climate science". Gee, it doesn't seem fair to blame Podos et al for climate change skepticism. I doubt Trump and friends have read any of the song performance literature. ☹

DK: *I fundamentally disagree.* It is pseudo-scientists acting as scientists in all disciplines that degrade all scientific endeavors and the trust of the public. If climate deniers and science skeptics had a read on what passes for science in this supposedly scientific literature on animal behavior, they'd have a field day.

Other than that, I thought the last two paragraphs are fine, and serve to place your criticism of this specific literature into a larger scientific context. It thus can be viewed as a potential case study indicating the difficulties lurking in hypothesis-driven research, a topic of broad concern in the scientific community these days.

DK: Hmm. There are 4 paragraphs to go. You're going to let me get away with my nutritional stress swipe?

I look forward to your revision. If I possibly can, I will accept it as is (or close to it)!

DK: Maybe closer to "it"? How could you possibly accept anything from me as is?

all the best,

Mike

Michael Beecher

Guest Editor

DK: Thank you for your abundantly constructive thoughts.

Reviewer #2:

I like the "less is more" approach you have taken in writing a short, pithy conclusion to the performance exchange. Reading Mike Beecher's editorial comments about the prior version of the manuscript and your responses to them, I can see that 1) Mike feels that you, the journal, and our field are best served if you avoid statements that convey or seem tinged by personal animus, and 2) you feel frustrated by this restriction because you believe that the authors you targeted are exemplars of deep problems in the way science is conducted, and should be explicitly called out for these failings. I suppose that your frustration is intensified by the unprofessional and at times unethical actions (and inactions) taken by some of the principals in this exchange.

I think that you have been gracious in trying your best to accommodate the constraints of journal-appropriate discourse even though doing so cuts against your preference for bluntness. I also think that your restraint has thus far probably increased the chances that your critique will be received favorably by those who most need to hear it, namely the researchers who did not start out in one camp or the other, but might now decide not to go thoughtlessly down the performance rabbit hole. So, with that in mind, as I read your revised commentary, I subjected each paragraph to a "flamethrower test," looking for things that were inflammatory but could be changed or removed without undercutting what I take to be your main points. Here is my report:

DK: Well-said. Yes, I feel frustrated as stated above, but I am also not stupid. If Al Franken can listen to his advisers, perhaps I can too. I'm doing my best.

paragraph 1 - Here you say that the targeted authors defend their work on the basis of its popularity. Quotes support this assertion about the authors. No problems here.

DK: Great Start!!!

paragraph 2 - The natural follow-up to the set-up in paragraph 1 would be something to the effect that despite widespread acceptance of the vocal deviation hypothesis, your detailed critique showed that the available evidence is flawed and insufficient to support the hypothesis. And you do kind of make that point, but in a way that does not pass the flamethrower test:

-- "no credible scientific data" is a little harsh-sounding. It would be better to convey this idea with more measured language.

DK: "no available scientific data support these ideas"—maybe that's an improvement

-- "prolific, influential authors" does not seem relevant to your point, serves mainly to personalize the target (don't worry, readers will understand who you mean without this)

DK: OK, I'll delete the "subject" of the passive voice sentence and see how that goes. I think it becomes less personal and still pointed. I slip in the word "repeatedly" to help make the point.

-- "intuitively appealing but false story" implies that the perpetrators are intentionally lying, which might be an accurate representation of your belief, but is an inflammatory accusation.

DK: maybe "intuitively appealing but unsupported story"?

-- "uncritically accepted by others" seems kind of contemptuous of many readers, most of whom probably didn't previously think too deeply about vocal deviation because the work was in well-written articles published in respected journals (and was intuitively appealing, as you say). Anyway, the "others" form a group you want to persuade, so why alienate them?

DK: Well, all it took was another sentence to fail the flames. How about making the point more about what I have done than what others have failed to do: "all of which has until now escaped critical evaluation." OK, I let the uncritical citers off the hook. Maybe the revised sentence is now OK????

New version: "The considerable literature that has developed on this topic is instead a premiere illustration of how highly flawed methods have been repeatedly used to tell an intuitively appealing but unsupported story, all of which has until now escaped a critical evaluation."

-- final sentence goes after the "establishment" in the form of ABS, and names names in a thinly veiled way. Too personal to pass the flamethrower test.

Overall, it seems that the first part of the paragraph (up to the final sentence) could pretty easily be reframed in more neutral language (you're just reiterating a conclusion of your critique: the data are flawed and don't support the hypothesis). The last sentence is really on a different topic, the culpability of scientific institutions and the professional reward system in perpetuating bad science. I think that this point is actually well covered in the final two paragraphs of your essay, in an equally pointed but less snarky way.

DK: I've deleted the final sentence, as discussed under Beecher comments above. But I feel that something needs to be said about how this happens, other than what appears in the last paragraph. Maybe my "premiere illustration" sentence, as adjusted, is ok. All of this literature is really extraordinary in how it has happened, and how no one seems to have questioned it. Is it really "contemptuous" to say that readers have uncritically accepted this work? And the authors are prolific and influential (ABS officers, awardees, NSF panels, etc.). I struggle with letting go of some of this stuff, but I think I've done an

admirable job of toning it all down???? Maybe the revision will even pass the f-t test?
Maybe still a little candle power there?

paragraph 3 - Here you briefly comment on each of the three replies; no problem with doing that, of course. The comment on Podos would pass the flamethrower test if you removed "defense of the indefensible." Shortening this sentence to "Podos instead addresses only minor, secondary issues" would get the point across in a flame-free way. The comment on Vehrencamp et al. makes the point that their reply consists mainly of name-calling. I found the way that you made this point using quotes kind of humorous, but it's possible that other readers might not understand what you're doing. Still, I don't see anything terribly objectionable here, or in the comment about Cardoso's reply.

DK: I've come to think that there's an unnecessary put-down of Vehrencamp, though intended to be a little humorous (and to point out how bizarre it is), but Beecher suggested a good way out of it, so I'll just combine Podos and Vehrencamp et al. in a single sentence.

paragraphs 4 through 7 - The replies from the targeted authors seemed to either ridicule or not understand your criticism of use of the word "performance," so I thought that these four paragraphs were interesting and useful. This topic does end up being the only topic from your original critique that you really expand on at length in this essay, which might seem a little random to some readers, but I think it is an interesting choice. (An aside: the most recent issue of Behavioral Ecology includes a commentary and a bunch of responses on the topic of whether there is any value to the literature on animal personality. The main commentary includes a discussion analogous to yours; it's about the ill effects of using the loaded word "boldness" to describe what is actually a measurement of an animal's distance from an object or another animal. So you are not alone in worrying about the corrosive effects of "marketing" language in animal behavior research). Anyway, I don't see anything in this multi-paragraph passage that runs afoul of the flamethrower test.

DK: Good. Beecher worked some of this over, so I've changed some of the wording to accommodate his thinking. It is the hidden implications of the word "performance" that seem to me to be the problem. Replacing "performance" with neutral words, as I have suggested, perhaps doesn't work all that well.

paragraph 8 - This paragraph contains some blunt statements about the nutritional stress literature and about a particular paper that you have previously criticized. Only a few changes would be needed to cool this paragraph down slightly so that there are no questions with respect to flamethrower issues. In particular, you could 1) change "in which readers are led to believe" to "which concludes that" 2) change "promoted over a hundred times" to "cited over a hundred times," 3) change "exploited" to "used."

DK: all 3 changes made

Some readers might see this paragraph as tangential to the vocal deviation issue, but I think it fits as a warning that other parts of the birdsong literature are beset with similar problems. (It's hard for me to be objective, though, as I have a pre-existing low opinion of the nutritional stress and repertoire literatures).

DK: Credit me for leaving out the paragraph on the "sexy syllables" in canaries. The foundations of this literature are equally sketchy, and it has flourished without thoughtful evaluation.

paragraph 9 - This paragraph says that people who cite the vocal deviation literature lack credibility and that one gains credibility by not citing that literature. I don't think that there is any way to make this point and still pass the flamethrower test. It kind of insults anyone who has cited one of these papers (a group that probably includes most people who will read the exchange). Anyway, explicitly making this point is not really necessary because the overall point of your original critique is that the vocal deviation literature is bogus and one shouldn't believe a word of it. Anyone who accepts the critique will not be citing this literature, at least not in an approving way, and does not need an explicit warning. Also, the meta-point here, that we should stop citing flawed science if we want science to get less flawed, is implied in the next paragraph (in the line about shared responsibility for producing good science). If you wanted to make this point explicit there, you could easily add a phrase to that paragraph making it clear that not citing flawed research is one of the responsibilities of authors.

DK: Yes, I suppose, but let me nuance just a bit. The wording in my paragraph isn't just about "citing" flawed research; it's more about using flawed research to promote and validate one's own research. That's a little different, though maybe not enough different to turn the flames off. This point feels sufficiently important to me that it warrants a full paragraph. Using flawed research to bolster one's own ideas does destroy the credibility of the authors who use that approach. I'll rethink this paragraph one more time when I reread it . . . OK, I've adjusted a few words to stress that it is not just citing these flawed papers but relying on them to reinforce discussions and arguments. Maybe that passes muster? . . . I've tried inserting a sentence into the next paragraph, but just can't make it work satisfactorily, and this paragraph is a point that I think is important to stress. . . . I've come back to this paragraph repeatedly, trying to figure out what to do with it. I really want a full, explicit statement that says "cite responsibly," because, for me, that's where everyone's critical thinking must begin when writing a paper. It's irresponsible to just throw a bunch of references into a paragraph and feel that the job is done; "you are who your citations are" . . .

paragraph 10 - The excellent summation in this paragraph is why you don't need the "hot" personal call-outs that appeared in some earlier spots in the manuscript. You speak in general terms about science, and then somewhat more specifically about the performance literature as a whole. But because of the context, every reader will know exactly who you were thinking about

when you wrote those words. So you are making an implicit attack on your targets, but without inflammatory personalization.

paragraph 11 - Ditto what I said for the prior paragraph. Tying the quotes from Gitzen to the performance literature allows you to make a critique that is very strong, but not explicitly personal.

DK: OK. I'll leave the last two paragraphs as is.

Overall, then, it seems to me that you could conform to the requisite norms by rewriting paragraph 2 from scratch, ditching paragraph 9, and cleaning-up certain sentences in other paragraphs.

DK: I'm having trouble ditching paragraph 9 about who cites what to bolster one's arguments.

I'm not sure that Mike will agree, but I think that your final two paragraphs use acceptably broad and general language, but still leave no doubt about how strongly you feel about the performance literature and its authors.

DK: Mike didn't seem to object.

--Signed Review

NOTE: Upon submitting your revised manuscript, please upload the source files for your article. For additional details regarding acceptable file formats, please refer to the Guide for Authors at: <http://www.elsevier.com/journals/animal-behaviour/0003-3472/guide-for-authors>

When submitting your revised paper, we ask that you include the following items:

Response to Reviewers (mandatory)

This should be a separate file labeled "Response to Reviewers" that carefully addresses, point-by-point, the issues raised in the comments appended below. You should also include a suitable rebuttal to any specific request for change that you have not made. Mention the page, paragraph, and line number of any revisions that are made.

Manuscript and Figure Source Files (mandatory)

We cannot accommodate PDF manuscript files for production purposes. We also ask that when submitting your revision you follow the journal formatting guidelines. Figures may be embedded within the source file for the submission as long as they are of sufficient resolution for Production. For any figure that cannot be embedded within the source file (such as *.PSD Photoshop files), the original figure needs to be uploaded separately. Refer to the Guide for Authors for additional information.

<http://www.elsevier.com/journals/animal-behaviour/0003-3472/guide-for-authors>. However, tables should be formatted in Word, not embedded in Word from a different program.

Highlights (mandatory)

Highlights consist of a short collection of bullet points that convey the core findings of the article and should be submitted in a separate file in the online submission system. Please use 'Highlights' in the file name and include 3 to 5 bullet points (maximum 85 characters, including spaces, per bullet point). See the following website for more information

<http://www.elsevier.com/highlights>

Animal Behaviour features the Interactive Plot Viewer, see:

<http://www.elsevier.com/interactiveplots>. Interactive Plots provide easy access to the data behind plots. To include one with your article, please prepare a .csv file with your plot data and test it online at <http://authortools.elsevier.com/interactiveplots/verification> before submission as supplementary material.