

Striving for science that is transparent, credible—and enjoyable: a comment on Ihle et al.

Walter D. Koenig^{a,b}

^aLab of Ornithology, Cornell University, 159 Sapsucker Woods Road, Ithaca, NY 14850, USA, and ^bHastings Reservation, University of California, 38601 E. Carmel Valley Road, Carmel Valley, CA 93924, USA

Ihle et al. (2017) outline a program designed to enhance the reliability, accuracy, and impact of behavioral ecology research. These are all worthy goals, and many of the suggestions they offer would enhance not only the transparency of the scientific process but also benefit researchers, particularly when it comes to reconstructing the sometimes convoluted process that went into analyses performed years previously. There is also a strong argument for embracing many, if not all, of the precepts of the open science movement of which Ihle et al.'s proposals are a part. The more accessible our data, analyses, and thinking are to future workers, the more valuable and useful our work will be to them.

Ihle et al. are explicitly interested in improving how we, as behavioral ecologists, conduct our science. Their proposals to pre-register and track questions and methods work well when it comes to experiments, which require a priori hypotheses, predictions, and explication of experimental design. However, I have significant reservations about both these processes when it comes to observational studies that seek to identify patterns and understand behaviors, particularly, but not necessarily limited to, when they involve data from long-term studies.

My own work, for example, has focused on projects that were initiated decades ago with the goals of addressing questions that ultimately became at best secondary. In the process of collecting data and analyzing results, new questions arose that led to unexpected questions, more analyses, and additional hypotheses to be tested, sometimes experimentally but more often by means of yet more exploratory analyses of existing data. As a result, data analysis has invariably been a learning process for me, regardless of how carefully I may have worked out the original questions and the statistical tests I thought I was going to perform beforehand. Part of this was due to the emergence of new analytical techniques, suggestions by reviewers, coauthors, and others, but more often it came about due to new and unexpected insights emerging from an initial set of analyses. More than once, the analyses in the ultimate version of a paper have borne little resemblance to those I originally envisioned. These changes have resulted not from bias, but from a combination of learning and self-correction.

Tracking and annotating all of the analyses and changes that went into the final version of my papers would have required a great deal of time. Would I have gained some benefit by investing all that effort? Absolutely. Am I sorry I haven't done so? Not in the least. I can't for the life of me think of how I would have had time to succeed if I'd had to document for the public record every thought and change going into the production of what were often relatively complex non-experimental papers, much less read and evaluate such documents produced by others.

Then there's pre-registration. What such a process does, as explained thoughtfully by Kosmala (2016), is shift a great deal of time and energy away from what to me are enjoyable and exciting parts of the scientific process—using data to discover the answer to a question—to a part I find relatively tedious, namely the analytical process itself. To me, having science be fun is worth the tradeoff of

a few extra false positives, especially since there are numerous alternatives to null hypothesis testing, where such errors are an issue (Johnson 1999, Anderson et al. 2000).

So, do we want science to be fun and creative, or tedious and a bit less flawed? Rigorous pre-registration and tracking will neither prevent cheating nor significantly enhance society's eroded confidence in science, which has much deeper roots than can be fixed by the solutions offered by Ihle et al. What would be tragic, however, would be to discourage aspiring behavioral ecologists of the future by forcing them to spend an inordinate amount of time planning their research rather than experiencing the joy of actually doing it. I, for one, am immensely grateful that I've had the opportunity to focus on the latter rather than the former.

FUNDING

My work has been supported by the National Science Foundation, most recently grants DEB-1256394 and IOS-1455900.

I thank J. Dickinson and J. Knops for their comments and suggestions.

Address correspondence to W.D. Koenig E-mail: wdk4@cornell.edu

Received 28 November 2016; editorial decision 23 December 2016; accepted 6 January 2017.

doi:10.1093/beheco/ax004

Editor-in-Chief: Leigh Simmons

REFERENCES

- Anderson DR, Burnham KP, Thompson WL. 2000. Null hypothesis testing: problems, prevalence, and an alternative. *J Wildl Manage.* 64:912–923.
- Ihle M, Winney I, Krystalli A, Croucher M. 2017. Striving for transparent and credible research: practical guidelines for behavioral ecologists. *Behav Ecol.* 28: 348–354.
- Johnson DH. 1999. The insignificance of statistical significance testing. *J Wildl Manage.* 63:763–772.
- Kosmala M. 2016. Thoughts on preregistering my research [cited 2016 November 18]. Available from: <http://ecologybits.com/index.php/2016/11/16/thoughts-on-preregistering-my-research/>.

Preregister now for an upgrade to Behavioral Ecology 2.0: a comment on Ihle et al.

Wolfgang Forstmeier^a

^aDepartment of Behavioural Ecology and Evolutionary Genetics, Max Planck Institute for Ornithology, 82319 Seewiesen, Germany

Let's face it: there is an undeclared conflict of interest in our daily work that was already spotted 40 years ago (Greenwald 1975) and that has been hampering the transparency that Ihle et al. (2017) are calling for. Researchers obtain direct benefits from finding significant effects in their data and from selectively reporting those effects (Smaldino and McElreath 2016). Significant results are easier to publish than null results; they receive more media coverage and