Commentary: it is a good idea to have bad ideas in science.

Abstract

There are few truly bad ideas in authentic science. We need to embrace science as a process-driven human endeavour to better understand the world around us. Products are important, but through better transparency, we can leverage ideas, good and bad, ours and others, to do better science. In a brief commentary here inspired by a recent discussion of the topic and previous introspections by other ecologists, it is proposed that whilst it is a good idea to track ideas and all the processes that generate outcomes such as publications, there is inherent merit in all scientific ideas. That said, organizing and framing our ideas into the networks that we already use to examine hypotheses and questions in science is a window into our workflows including ideation, implementation, data analyses, and how we can better map ideas into open science outcomes. Formalizing and describing the linkages between ideas, data, and projects we produce as scientists will enhance and diversify the value of the work we do individually and collectively.

Preamble

A recent editorial suggests that science is all about sorting the wheat from the chaff (Kirwan 2017). To be clear and fair, the author conceptualizes bad ideas using suggestions from eminent scientists on the value and necessity of bad ideas to the advancement of good science. The bad-idea science paradigm proposed is a simple dichotomy defining good ideas as those that generate publications and bad as those that do not. This is a functional taxonomy for the purposes of a self-assessment of work done (i.e. effort allocated) and project ideas by the author, and to the defence of this analysis, every effort was made to qualitatively include ideas that were 'built on' for subsequent positive outcomes, i.e. publications. This working definition is an absolutely necessary and convenient short-cut, but it is not the end of the story. The accountability of ideas to the progress of a discipline such as ecology is not a new idea (Aarssen 1997) nor without debate (Weiner 1999). Common ground suggests that we should individually evaluate the merit of our ideas if for no other reason then to better prepare our work for the review process of others. We can even conceptualize some of the less productive ideas as stepping stones to more useful ones or as a counterpoint to frame and anchor the relative merit of those that succeed in whatever capacity we elect to define the positive outcomes. Creation, divine or otherwise, is not a perfect system. Evolution needs mutations. That ideas are bad if they do not either directly or indirectly generate publications is perhaps too limiting. New ideas are novel, and new *and* useful ideas are creative (Runco and Jaeger 2012). The former can become the subtrate for the latter. Useless today can become indispensible tomorrow. Introspection of our workflows and the relationship between ideation and implementation can streamline the mapping of ideas and hypotheses to effective testing through evidence. However, we also need to ensure that ideas do not get lost through self-criticism, lack of data, availability of a mechanism to test today, or the extensive volume of information we process. In the spirit of replication science (Brandt et al. 2014; Mulkay and Gilbert 1986), I examined my relative idea management and outcomes to explore whether this more generous interpretation of merit and my workflow kept most of the wheat and some of the chaff.

Replication

A simplified, direct replication of the concept of idea merit self-assessement from the most recent editorial that inspired this commentary was done. A check of idea efficacy scrapes the files and structures used to support a scientific workflow (hosted locally, but see below on how this landscape is changing). This in and of itself is a superb idea. The need and pressure to publish can encourage one to be less cognizant of the processes that support the final paper particularly if there are many steps or if time to final acceptance is significant. Many of the simple rules for data and experimental provenance (Kazic 2015) also apply to what can be similarly term 'scientific idea provenance' including version control, integration, object labelling, and reviewing the semantics. The author in the former idea provenance self-study examined all projects completed in career, recorded initiation date, scored effort for each, and developed an outcome classification

scale that included direct and indirect publications from each project (Kirwan 2017). A project was defined as a 'project folder' that likely represented parent directories organized around each independent research thread. The author concluded that 25% of projects generated a publishable result. The assumption of this workflow is that every scientific task is assigned to/nested within a project folder. The folder dimension of my scientific workflows is similar but not a perfect match (and it is fundamentally evolving). I use project folders to organize protocols and ideas, dataset folders for data and some other forms of scientific evidence such as camera trap pictures, an 'idea archive' parent folder to store all ideas, and a 'papers-in-progress' parent folder with subfolders to store ideas that have some development in written form. However, this workflow and structure for organzing the processes that support scientific inquiry is dramatically changing thanks to GitHub and RStudio where I couple code, annotation, and written interepretations with the associated datasets. Scoring effort and time allocated, pre-Github and without effective, time-stamped version control was not viable in terms of reproducibility with my former workflow and its files. Consequently, I restricted analyses primarily to counts. However, I propose that the number of objects within a project, i.e. files, in addition to parent/child folders that support a research thread is an important building block in many instances in the process of turning ideas into more tangible outcomes. This conceptualization of ideas and project management more fully embraces provenance and can include both folder (parent and child) and file frequencies. Hence, I use a more topological approach to model the merit of ideas herein but expanded the supporting structures associated with ideas and workflows.

Literature cited

Aarssen, L.W. 1997. "On the Progress of Ecology." Journal Article. *Oikos* 80 (1): 177–78.

Brandt, Mark J., Hans Ijzerman, Ap Dijksterhuis, Frank J. Farach, Jason Geller, Roger Giner-Sorolla, James A. Grange, Marco Perugini, Jeffrey R. Spies, and Anna van 't Veer. 2014. "The Replication Recipe: What Makes for a Convincing Replication?" Journal Article. *Journal of Experimental Social Psychology* 50: 217–24. doi:http://dx.doi.org/10.1016/j.jesp.2013.10.005.

Kazic, Toni. 2015. "Ten Simple Rules for Experiments' Provenance." Journal Article. *PLoS Comput Biol* 11 (10): e1004384. doi:10.1371/journal.pcbi.1004384.

Kirwan, J. 2017. "It's Good to Have Lots of Bad Ideas." Journal Article. *Nature* 548: 491. doi:10.1038/nj7668-491a.

Mulkay, Michael, and G. Nigel Gilbert. 1986. "Replication and Mere Replication." Journal Article. *Philosophy of the Social Sciences* 16 (1): 21–37. doi:10.1177/004839318601600102.

Runco, M.A., and G.J. Jaeger. 2012. "The Standard Definition of Creativity." Journal Article. *Creativity Research Journal* 24: 92–96.

Weiner, J. 1999. "On Self-Criticism in Ecology." Journal Article. *Oikos* 85 (2): 373–74.