BioScience Advance Access published January 14, 2016

Would Ecology Fail the Repeatability Test?

STEFAN A. SCHNITZER AND WALTER P. CARSON

The ability to replicate scientific studies and confirm their results is a cornerstone of the scientific process. The low repeatability of empirical studies would be cause for alarm for nearly any scientific discipline. However, studies designed specifically to evaluate the degree that scientific studies are repeatable or reproducible are rare. In one of the first comprehensive studies of its kind, the Reproducibility Project (Nosek et al. 2015) selected 100 experimental and observational studies that were published in three psychological journals and repeated them over a 3-year period. Ideally, repeating these studies should have led to results and conclusions similar to the original ones. Moreover, the magnitude and direction of the findings (the effect sizes) should have varied only modestly above or below those of the published studies. Instead, the authors found that just over one-third of the repeated studies had significant results compared with 97 percent of the original studies, and only 39 percent of the repeated studies found the same results as the original ones did. Furthermore, fewer than 50 percent of the repeated studies had effect sizes matching those reported in the original studies. This lack of repeatability casts a shadow of doubt on the reliability of earlier studies and potentially challenges findings across much of the discipline.

Although scientific disciplines can differ in their level of repeatability, the findings of the Reproducibility Project provide a cautionary tale for all fields of science. The field of ecology may be particularly susceptible to a lack of repeatability for several reasons. First, species distributions, demography, and interactions vary widely over space and time, which may result in the findings of repeated studies' deviating from those of the original ones. Second, ecological field studies conducted over large spatial and temporal scales are the most definitive but the least likely to be repeated because of logistical and economic constraints. Stated simply, who is going to repeat enormously large studies, and who is going to fund them? Third, rapid environmental change may create different outcomes even if we can replicate the exact same field experiment; the ecological interactions, their magnitude, or even direction may have changed permanently.

The validity and repeatability of published studies in ecology are crucial assumptions for the continued development of the discipline. Although there is little direct empirical evidence for a lack of repeatability of key ecological findings, there are few examples in which the goal of researchers was to recreate and test the repeatability of a seminal study (e.g., Navarette and Menge 1996). Here, we discuss reasons for the potential lack of precision and repeatability of ecological studies. We also discuss some of the ways to move forward in ecology, even without a formal reproducibility project.

Impediments to reproducibility in ecology

The reproduction of ecological studies will be hampered by the tremendous variation inherent in natural ecosystems. Differences in species abundance, demography, and interactions, as well as changes in abiotic conditions over time and space, can cause enormous variation in repeated studies—even those conducted at the same study site. Although we would not expect fundamental ecological processes, such as the mechanisms that maintain species diversity and control species distributions, to change capriciously over time and space, the relative strengths of these mechanisms may differ considerably and change in their relative importance from site to site.

A huge potential issue is that ongoing environmental changes will complicate comparisons of recent studies with those that were conducted decades ago. Indeed, a rapidly changing climate can alter the relative abundance of species, which will change the outcomes of replicated studies. It would be a major challenge to tease apart whether different outcomes in original and repeated studies were caused by a lack of repeatability or by climate-induced differences in biotic and abiotic conditions.

Another significant barrier to repeating ecological studies is their large scale and temporal duration. Some of the most influential studies in ecology over the past 40 years have been extremely large in both space and time. Repeating many of these studies would require an enormous effort—entire careers spanning decades could be spent repeating influential ecological studies that were published years ago. Imagine the difficulty of repeating the enormous (but unreplicated) Hubbard Brook ecosystem study (Likens et al. 1977) or replicating 25 years of monthly collections of phenology and productivity data for woody plant species throughout a tropical forest in Panama (e.g., Wright et al. 2004). Similarly, establishing and maintaining new rodent exclosures for 25 years in the Arizona desert or a similar habitat is certainly daunting (Thibault et al. 2010).

Even if the substantial funding necessary to repeat these studies were available, what could we conclude if we found sharply contrasting results

between the original and repeated studies? When faced with two replicate studies that yield substantial differences in effect size, additional replicate studies would be necessary to determine which result, if either, is the most accurate. Therefore, repeating seminal studies may provide a rough estimate of the level of repeatability in ecological studies; however, to critically evaluate the theory being tested and to move beyond merely testing whether it is repeatable, we would need a large number of replicate studies over time and space.

Moving forward

Given the practical limitations of replicating large-scale ecological studies, we suggest that effort and funding would be better spent conducting new studies that provide novel and comprehensive tests of broad ecological theory. Large-scale, long-term, well-replicated studies that improve on previous methods will be a much better way to advance the field of ecology and give insight into whether previous findings are general and broadly applicable. For example, the Amazon Forest Inventory Network (Rainfor) and the Center for Tropical Forest Science have hundreds of forest plots around the world where all trees of all species are mapped and measured down to small size classes, which may be a good way to move forward with some level of repeatability. Meta-analyses of existing data are another way to converge on general principles in ecology. By combining appropriate data from multiple studies that address a single theory, one can provide a general test that includes variation in both time and space.

A more difficult problem to address is the publication bias that results from the difficulty of publishing negative results. Studies that fail to empirically support a popular theory are notoriously difficult to publish. The failure to publish negative results may lead ecologists to believe that favored ecological interactions are more pervasive and ubiquitous than they really are. Unfortunately, negative results may be relegated to dusty file drawers (or hard drives) after being repeatedly rejected for publication. Ecological societies, such as the Ecological Society of America or the British Ecological Society, could encourage the submission of negative results to their open-access online journals under a new "negative results" section if the research used an appropriate experimental design and levels of replication consistent with previous studies that found significant effects.

One final reason for a lack of repeatability of a key research finding is data fabrication or other unethical behavior, such as selectively removing data points. Although we would like to believe that this behavior is rare in ecology, this belief may be naive in light of the large number of retractions coming from other scientific fields (*www.retractionwatch.com*) and the real or perceived pressure to publish high-impact papers, which are often necessary to compete successfully for jobs and for increasingly scarce resources that support ecological research.

The increasing trend of publishing ecological data as a supplement to articles may help reduce fraud, although fabricated data still may be difficult to detect. Ultimately, however, the success of any scholarly field depends on the honesty and ethics of its citizens, which we hope are maintained in ecology.

The many scientists who participated in the Reproducibility Project should be applauded for their honest evaluation of their discipline. Their work is a reminder that all fields of science have some inherent weaknesses and that the lack of repeatability may be shared by many disciplines. The findings of the Reproducibility Project are particularly important to the field of ecology, which has many characteristics that make the reproduction of results challenging, and key findings and seminal studies in ecology typically remain unreplicated and unverified. Nonetheless, by combining large-scale, long-term replicated studies with meta-analyses, we can both test new theories and also verify the general findings of seminal studies that have a large influence on contemporary ecology.

Acknowledgments

We thank Jeremy Fox, Jonathan Pruitt, Kathryn Barry, Sergio Estrada, and Maria Rodriguez for their comments on the manuscript. Financial support was provided by NSF-DEB 0845071 and NSF-DEB 1019436 (to SAS).

References cited

Likens GE, Bormann FH, Pierce RS, Eaton JS, Johnson NM. 1977. Biogeochemistry of a Forested Ecosystem. Springer.

Navarette SA, Menge SA. 1996. Keystone predation and interaction strength: Interactive effects of predators on their main prey. Ecological Monographs 66: 409–429.

Nosek BA, et al. 2015. Estimating the reproducibility of psychological science. Science 349 (art. aac4716). doi:10.1126/science.aac4716

Thibault KM, Ernest SKM, White EP, Brown JH, Goheen JR. 2010. Long-term insights into the influence of precipitation on community dynamics in desert rodents. Journal of Mammalogy 91: 787–797.

Wright SJ, Calderon O, Hernandez A, Paton S. 2004. Are lianas increasing in importance in tropical forests? A 17-year record from Panama. Ecology 85: 484–489.

Stefan A. Schnitzer (s1@marquette.edu) is affiliated with the Department of Biological Sciences at Marquette University, in Milwaukee, Wisconsin. Walter P. Carson is affiliated with the Department of Biological Sciences at the University of Pittsburgh, in Pennsylvania.

doi:10.1093/biosci/biv176