Distinguishing Reproducibility Issues from the Heterogeneity of Biological Responses

Stephen Keith Chapes, Kansas State University

Summary:There are concerns about the reproducibility of scientific experiments in several scientific fields including psychology, biology, biochemistry and biomedicine. This problem may be one of the factors in a diminished acceptance of science as a societal impact factor in driving decisions. Less informed decisions will lead to poorer outcomes and will not bode well for the future of society. The problem of reproducibility is also compounded by the underlying fact that there is natural biological variability. Therefore, one of the problems that we face in science today is to separate issues of reproducibility with biological heterogeneity. This essay will address these issues.

In recent years, the issue of data reproducibility has become a significant concern. In the field of behavioral psychology there is a crisis surrounding the ability to reproduce key findings that have become central to behavioral dogma (https://doi.org/10.1111/psyp.12787). There are issues with the reproducibility of preclinical cancer drug and treatment therapies (https://cancerres.aacrjournals.org/content/79/13_Supplement/90.article-info) and there are issues with the reproducibility of human tissues (https://doi.org/10.1016/j.yamp.2018.07.003). Not only is this is a problem when trying to translate basic research into usable medical practice, the inability of scientist to show consistent results draws criticism and doubt from an increasingly scientifically skeptical public. Therefore, the scientific community must address this issue.

Data reproducibility not only is an issue of scientific precision, it also must be viewed in the context of biological heterogeneity. Most biological responses are not exact and exhibit a normal distribution of activities (https://doi.org/10.1111/anae.13666). This heterogeneity can be quite narrow or can be highly variable depending on the response. Unfortunately, we don't always have good measures of this heterogeneity or the variability is superimposed on the imprecision of the experimental design (methodology and instrumentation). Therefore, data reproducibility is not an easy question to address and the heterogeneous response gradient makes it a more complex problem.

Why is it critical to address this problem now? We live in an era where leaders question the scientific community (https://www.nytimes.com/2019/05/27/us/politics/trump-climate-science.html). To maintain credibility, science and scientist must not appear to be slipshod or bumble-heads! The work must be accurate and it must be informative.

There are a number of barriers to improved data reproducibility. One significant issue is the long history of principal investigator (PI)-centered labs where the PI mentors students and supervises technical staff and controls the research focus. Data from those labs is kept in-house with the publications giving insight into what was done. Often the publications don't have complete methodology and maybe some of the metadata are missing from the primary publication. Post publication, especially after some of the students graduate or technicians leave the lab, the metadata or methods are hard to document.

Ethics of animal use and scientific standards of replication is also an issue. In general, Institutional Animal Care and Use Committees (IACUC) frown on the excessive use of animals. Power analysis is used to define the needed animal numbers and the statistics rule the final number of animals to be used. A weakness of power analysis is that biological variation may not be accurate and there may be more biological variation than anticipated. It is very hard to convince IACUCs to increase animal numbers even when the Guide for the Use of Animals in Research says it is scientifically justified (https://grants.nih.gov/grants/olaw/guide-for-the-care-and-use-of-laboratory-animals.pdf).

In wild or “natural” populations, it is difficult to know the history of the wild animals. Therefore, variation caused by their different life histories will not be known. This problem is compounded by genetic diversity in those populations; necessitating the understanding of the genetics. Indeed, this is true even for those studying animals under controlled laboratory situations. Mechanisms to improve this understanding must be facilitated.

There are both easy and hard tasks at hand in overcoming the barriers to good scientific replication. Regardless of whether PIs are going to maintain a single PI lab or they are going to be part of multi-disciplinary teams, there needs to be strict requirements for the archiving of meta data, primary data and detailed methodology. These requirements can be mandated both at the funding level AND at the publication level of the scientific process. A critical issue here is going to be establishing a standard format for the data archiving. The National Institutes of Health (NIH), the National Science Foundation (NSF) or the National Academy of Sciences (NAS) will need to take the lead on this. The requirement must be universal for all biological and possibly all scientific disciplines.

Changes in the thinking about experiment reproduction also are necessary. An acceptable avenue, and funding, for repetitive work is needed. IACUC standards need to accommodate large samplings to validate what is normal variation. Improved methodology is part of improved experimental reproduction. This not only includes greater precision of the instrumentation used to collect the data, but also a rethinking of how experiments are done. Different techniques might be used to address the question from different perspectives. This is where cross biological disciplines become a factor and multiple investigators plan an experiment to examine multiple systems. Instead of a scientist manipulating an animal to understand what goes on in one organ, without examining the rest of the animal, other scientists look at the bone, muscle, brain, etc. Again, this requires a change in thinking about the way science is done. More planning, fewer experiments but more data will be collected and a more comprehensive picture may emerge from the manipulation. Some imagination will be necessary to assemble these teams but this is where adaptation and new approaches can be encouraged by the large funding agencies (NIH, NSF, USDA, etc.). Programs to promote better understanding of the factors that cause variation will also be important (e.g. funding, conferences and collaborations).

How do we apply his approach to ecology and field work? Many ecologists already are part of multidisciplinary teams that collect comprehensive data sets. Long-Term Ecological Research (LTER) projects funded by the NSF have a long history of comprehensive data collection. Therefore, there is a precedent for this to extend to the "outdoor" people. A better unified system of data collection and archiving to cross disciplines will be a challenge but is necessary. This is where the NAS or NSF can take leadership. Planning to coordinate experiments across many labs and possibly species can be pre-planned with data collection and meta data analysis archived in appropriate, shareable fashion.

A big challenge for biologists of all persuasions is experimentally validating how much population heterogeneity contributes to experimental variation. Even in inbred animals, where the mice are supposed to be "clones" the biological measurements have standard deviations. We have statistical rules to help us determine how many animals we need to measure that variation but it will be important to make sure we can give the scientist working on genetically heterogeneous populations the resources to establish that variation.

Another advantage modern biologists have today is the development of super computing and machine learning. This is another opportunity to reintegrate biology with other disciplines to use computational approaches and collaborations with bioinformaticians to find data signatures that are meaningful. Good data archiving (as described above) and the encouragement of data mining instead of doing "novel wet-lab" experiments is needed. In the past, scientists thought that every experiment had to be done in their lab with their own animals or that they needed to be out in the field to make their own observations. The stigma of using others' data has diminished in the last few years but we need our scientific leaders to truly make this the first methodological approach before proceeding to hands-on experimentation. For those looking for refinement and reduction on the IACUC, this should be appealing. This approach should also help in improving data reproducibility. Of course, the data handling pipelines and methodologies are also sources of variation (https://onlinelibrary.wiley.com/doi/pdf/10.1096/fba.1017). Therefore, we also must encourage standardization and strong documentation of these methodologies.

Many of the necessary tools needed for this reintegration of biology are in place. What is necessary is a revolution in the standard way of doing science. This may require Universities to reevaluate their tenure assessments and what constitutes "productivity". Fewer publications with higher quality have to be judged for what they are and not on past standards. Leadership at the organizational level (NIH, NSF, etc.) will be necessary to change expectations.

The improvement in data reproduction also serves to reintegrate biology by encouraging many biologists to use a single data set. Freely accessible, reliable data sets, meta data and methodologies are attractive tools to biologists from other subfields. Scientists don't want to go down dead-end roads which often occurs when the data are questionable. They want the assurance that data they might access and collaborate on are reliable and reproducible.

It is my hope that program managers, scientists and funding leaders at the NIH and NSF heed this call for data reproducibility. Help (i.e. funding) to improve the precision in experimental measurements must be part of this effort. An understanding of biological variability will be among the outcomes and improved scientific collaboration will be driven by this effort. The reintegration of biology will be one of the beneficial consequences.