SANDIA REPORT

SAND2018-11186
Unlimited Release
Printed October 2018

Toward a Compatible Reproducibility Taxonomy for Computational and Computing Sciences

Michael A. Heroux, Lorena A. Barba, Manish Parashar, Victoria Stodden and Michela Taufer

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

SAND2018-11186
Printed October 2018
Unlimited Release

Toward a Compatible Reproducibility Taxonomy for Computational and Computing Sciences

Michael A. Heroux
Center for Computing Research
Sandia National Laboratories
P. O. Box 5800
Albuquerque, New Mexico 87185-1320

Lorena A. Barba
School of Applied Science and Engineering
George Washington University
800 22nd Street, NW
Washington, DC 20052 USA

Manish Parashar
Department of Computer Science
Rutgers University
110 Frelinghuysen Road
Piscataway, NJ 08854 USA

Victoria Stodden
School of Information Sciences
University of Illinois
501 E. Daniel St. MC-493
Champaign, IL 61820-6211 USA

Michela Taufer
Department of Electrical Engineering and Computer Science
University of Tennessee
Min H. Kao Building, Suite 401
1520 Middle Drive
Knoxville, TN 37996-2250 USA

Abstract

Reproducibility is an essential ingredient of the scientific enterprise. The ability to reproduce results builds trust that we can rely on the results as foundations for future scientific exploration. Presently, the fields of computational and computing sciences provide two opposing definitions of *reproducible* and *replicable*. In computational sciences, *reproducible research* means authors provide all necessary data and computer codes to run analyses again, so others can re-obtain the results (J. Claerbout et al., 1992). The concept was adopted and extended by several communities, where it was distinguished from *replication*: collecting new data to address the same question, and arriving at consistent findings (Peng et al. 2006). The Association of Computing Machinery (ACM), representing computer science and industry professionals, recently established a reproducibility initiative, adopting essentially opposite definitions. The purpose of this report is to raise awareness of the opposite definitions and propose a path to a compatible taxonomy.

ACKNOWLEDGMENTS

The authors thank Almadena Chtchelkanova and the U.S. National Science Foundation for sponsoring the July 25, 2017 workshop that provided the foundation for this report. The authors also thank other workshop attendees: Ronald Boisvert, National Institute of Standards and Technology; Dylan Chapp, University of Delaware; Bruce Childers, University of Pittsburgh; Juliana Freire, New York University; Carlos Maltzahn, University of California, Santa Cruz; Wilf Pinfold, Urban.Systems; Jeff Spies, Center for Open Science.

TABLE OF CONTENTS

TABLES

EXECUTIVE SUMMARY

The computational sciences (fields such as physics, engineering and biology, which use computing as a means for scientific discovery) and computing science (the science of computing itself) have evolved two distinct taxonomies for defining reproducibility concepts. Prevalent in the computational sciences is a taxonomy attributed to Claerbout, Donoho and Peng. The computing sciences have recently started using a newer taxonomy codified by ACM.

In order to advance reproducibility efforts in the computational sciences and computing science, this report describes the current state of reproducibility terminology in these fields, bringing particular attention to the fact that reproducibility and replicability have *opposite* definitions in the Claerbout/Donoho/Peng and ACM frameworks. We next propose a set of steps that we hope will lead to compatible, if not common, taxonomies in these fields.

The issues and strategy we propose are described in detail in subsequent sections, but we state the highlights here for quick reference:

1. ***Reproducibility taxonomies:*** *In order to give specific labels to distinct types of reproducibility efforts, scientists give particular definitions to common synonyms of reproducible. As a result, many communities have developed their own definitions of reproducible, replicable, repeatable, and related terms. These taxonomies differ from one community to another, depending on the history, needs and choices of the community.*

2. ***Opposite definitions for reproducibility and replicability:*** *The computing science community—scientists whose domain is computer science—have adopted a taxonomy that defines reproducibility as a different team producing the same computational results using a different experimental setup than the original authors' (different team, different experimental setup) and replicable as a different team producing the same computational result using the same experimental setup (different team, same setup). The computational sciences— scientists who use computing as a tool for scientific discovery in the physical sciences—have developed a different taxonomy for which reproducibility and replicability have essentially the opposite definition as in the computing sciences.*

3. ***Competing taxonomies for computational and computing sciences represent an impediment to progress:*** *Many scientific areas have different reproducibility taxonomies. In order to converse with scientists in a particular field, one must understand and use the taxonomy of that particular community. However, the computational and computing sciences have significant and increasing community overlap. The opposing definitions of reproducibility and replicability represent a serious impediment to progress and understanding within and across these communities.*

4. ***Raising awareness is the first step:*** *While our ultimate goal is to have compatible taxonomies, the first goal of this report is to raise awareness in the computational and computing sciences communities that reproducibility and replicability have opposite definitions depending on which body of literature and portion of these communities you are presently engaging. By raising awareness, we can at least start each conversation*

and publication with a statement of which taxonomy will be used.

5. ***Acknowledging historical precedent:*** *While each scientific community has a right to determine its own nomenclature, we believe that studying the history of reproducibility taxonomies in the computational and computing sciences illustrates that the computational sciences taxonomy significantly pre-dates the computing sciences taxonomy and is in wider use. Since the ACM taxonomy is just a few years old and the computational and computing sciences communities have significant overlap, we believe this historical precedence is noteworthy.*

6. ***Working toward a compatible taxonomy:*** *Both the computational and computing sciences reproducibility taxonomies have detailed definitions of a collection of terms. We do not think it is necessary that one community adopt the definitions of the other. However, we do believe that both communities will benefit from one community swapping its definition of reproducibility and replicability. We believe that the historical precedence of the computational science definitions of reproducibility and replicability suggests that the computing sciences should reverse its definitions.*

7. ***Establishing a taxonomy standard:*** *Achieving compatible reproducibility taxonomies may require or at least be assisted by establishing a standards effort for the computational and computing sciences. If required, we propose that IEEE or similar standardization body establish a committee toward this goal.*

1. INTRODUCTION

Reproducibility is an essential ingredient of the scientific enterprise. The ability to reproduce a result builds trust that we can rely on it as a foundation for future scientific exploration.

Experimental science has always strived to assure that the initial conditions and procedures of an experiment are adequately identified and controlled. Unknown or undetected input and procedure variation (Van Bevel, 2016) can lead to different results and conclusions. Thorough understanding, control and verification of inputs and procedures are essential for progress. The better we understand and control the experimental environment, the more we reduce risk, build trust in our results, and advance science.

Improved experimental environments are not all-or-nothing. For many years, scientific and analytical research communities have created taxonomies that use a collection of terms to describe the degree of reproducibility, the source of initial conditions, who is engaged in reproducing results and more. A common approach to building these taxonomies has been to use a group of words that are synonyms in normal discourse, and assign them subtle but important distinction. Assigning distinct yet related meaning to terms like reproducible, replicable and repeatable enable effective and efficient exchange of ideas within research communities. Similarly, some communities have established symbiotic relationships between words such as verification and validation.

A review of the literature shows that several communities have tackled reproducibility taxonomies. Internal medicine (Laine et al., 2007), applied econometrics (Koenker and Zeileis, 2009), computational biology (Sandve et al., 2013) and psychology (Stevens, 2017) are some of the fields that have adopted definitions. Each taxonomy has roughly the same set of terms, but creates a conceptual model with at least some nuanced difference in definitions.

This report is focused on the state of reproducibility taxonomies in two communities, which we label as:

- Computational Sciences: The computational sciences use computational modeling, simulation and data analysis as vehicles for scientific discovery. In this community, computing is primarily a tool.
- Computing Sciences: The computing sciences certainly include computing as a tool element, but also view computing itself as their primary research focus.

The conceptual models for computational and computing sciences are similar but not identical. Even so, there is a strong membership overlap between these communities. This is especially true in high-performance computing, where research in computing sciences can greatly enhance computational capabilities and where computational sciences can provide a rich set of challenge problems for computing research.

2. TAXONOMIES FOR COMPUTATIONAL & COMPUTING SCIENCES

The most common reproducibility taxonomy for computational sciences is what we will call the *Claerbout taxonomy* (Claerbout and Karrenbach, 1992; Buckheit and Donoho, 1995; Peng et al., 2006). A relatively new taxonomy for computing sciences is due to a formal effort sponsored by the Association for Computing Machinery (ACM), which we will call the *ACM taxonomy* (Stodden et al., 2013). As described in Barba (2018), both taxonomies have roots in a lineage that goes back to early discussions of scientific reproducibility. Claerbout refers back to experimental sciences, while ACM looks for its roots in the metrology literature.

As shown in Table 1, both the Claerbout and the ACM taxonomies use the terms reproduce or reproducible, and replicate or replicable. However, the meaning of the terms is swapped. In other words, the definition of reproducible in Claerbout is essentially equivalent to the definition of replicable in ACM, and the same for replicable in Claerbout and reproducible in ACM.

While different definitions of reproducibility terms are inevitable, and (we believe) difficult to universally reconcile, we find the opposite definitions in computational and computing sciences to be particularly confusing, and worth reconciling. Computational sciences and computing sciences have major community overlap: many people belong to both communities.

Table 1: Definitions of Reproducible and Replicable

Table 1: Claerbout/Donoho/Peng (Claerbout) and ACM definitions of Reproducible and Replicable. Claerbout definitions are prevalent in the computational science literature and have been used since the 1990s. The ACM definitions are used by ACM in its Artifact Review and Badging effort and first appeared in February 2013.

Term	Claerbout	ACM
Reproducible	Authors provide all the necessary data and the computer codes to run the analysis again, re-creating the results.	(Different team, different experimental setup.) The measurement can be obtained with stated precision by a different team, a different measuring system, in a different location on multiple trials. For computational experiments, this means that an independent group can obtain the same result using artifacts which they develop completely independently.
Replicable	A new study arrives at the same scientific findings as a previous study, collecting new data (with the same or different methods) and completes new analyses.	(Different team, same experimental setup.) The measurement can be obtained with stated precision by a different team using the same measurement procedure, the same measuring system, under the same operating conditions, in the same or a different location on multiple trials. For computational experiments, this means that an independent group can obtain the same result using the author's own artifacts.

2.1. The Challenge of Competing Taxonomies

Many scientific communities have developed a reproducibility taxonomy, each adopting their own qualitative sense of the degrees of reproducibility rigor, and the terms used to define these degrees. Econometrics, Political Science, Medicine and Biostatistics literature (see Barba, 2018) illustrate a few areas. In addition, there have been disagreements about the terminology. For example, in a 1991 economics publication, Nancy Cartwright argued for a definition of replicable and reproducible similar to the ACM taxonomy while Harry Collins (as described by Cartwright) uses the reverse definitions (Cartwright, 1991).

Differing taxonomies do not necessarily require reconciliation; it is reasonable to simply agree at the start of a discussion, conference, paper or other discourse which reproducibility taxonomy will be used. At the same time, for fields that have major overlap, such as the computational and computing sciences, it is desirable to seek compatibility, even if the same taxonomy is not possible for both communities. Compatibility is what we seek by writing this report.

2.2. The Historical Precedent of the Claerbout Taxonomy

While the ACM taxonomy is rooted in the metrology literature (JCGM, 2008), it is pre-dated by the Claerbout taxonomy, and the latter is widely used in a variety of scientific disciplines. Many scientists have come to understand the meanings of replicable and reproducible in a way that is consistent with Claerbout. The *Annals of Internal Medicine* espouses these meanings in a statement (Laine et al., 2007), so does the editorial in *Science* that introduced the special issue in Data Replication & Reproducibility (Jasny et al., 2011), and the reproducible-research initiative of the journal *Biostatistics* uses compatible definitions (Peng, 2009). The first comprehensive review of reproducibility in book form (Stodden et al., 2014) settles on the Claerbout/Peng terminology in its preface. Another recent book, consisting of 31 case studies in reproducible research, presents definitions aligned with the Claerbout taxonomy and states that they are broadly used throughout the book. The terminology is also compatible with the broad use of the term "replication studies" in empirical sciences (psychology, economics, political science), to mean new works attempting to independently confirm research findings.

The ACM taxonomy and definitions emerged only a few years ago (mentioned in Stodden, et al. 2013, approved June 2016) and appear to be unique to the computing sciences. Even so, the ACM badging system,[1] which is providing important incentives to advance the pursuit of reproducibility, has resulted in a growing awareness of reproducibility efforts overall and the ACM taxonomy in particular.

ACM badges are applied to papers that achieve some level of reproducibility, or at least proto-reproducibility. Several ACM-affiliated conferences such as the ACM/IEEE Supercomputing (SC) Conference series have integrated badges into the review and acceptance of publications. For example, an SC paper that contains an Artifact Description (AD) appendix is eligible for an *Artifacts Available* badge if the AD appendix is well written and provides links to openly available software and data used to produce the computational results in the paper.[2] Badges provide a tangible incentive for authors to invest time in reproducibility efforts. Fortunately,

[1] https://www.acm.org/publications/policies/artifact-review-badging
[2] https://sc18.supercomputing.org/submit/sc-reproducibility-initiative/

most badges do not use the terms reproducible or replicable. In fact, only the Results Replicated and Results Reproduced badges reveal the swapped ACM definitions.

A final note about the swapping of terms reproduce and replicate: a linguistics professor at University of Pennsylvania analyzed usage in the literature and traced the direct swap of terms to a single paper presented at a machine-learning workshop. This paper appears cited in most of the later publications that use the terminology now assimilated in the ACM taxonomy. The linguist's analysis also endorses the precedent of Claerbout and Peng (Liberman, 2015).

3. LESSONS FROM V & V DEFINITIONS

Verification and validation are two essential activities in science and engineering. Often combined ambiguously as V & V, these paired terms have more than one set of definitions. A history of V & V definitions is provided by Oberkampf and Roy (Oberkampf and Roy, 2010), where the authors describe the evolution and refinement of V & V definitions, exhibiting how the perspectives and sensibilities of particular communities can lead to seemingly subtle but important differences. From this book, we obtain several definitions.

From the Society for Computer Simulation (SCS), definitions of Model V & V (1979):

Model verification: substantiation that a computerized model represents a conceptual model within specified limits of accuracy.

Model validation: substantiation that a computerized model within its domain of applicability possesses a satisfactory range of accuracy consistent with the intended application of the model.

From the U.S. Department of Defense (DoD) (1994):

Verification: the process of determining that a model implementation accurately represents the developer's conceptual description of the model.

Validation: the process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model.

Some years later the American Institute of Aeronautics and Astronautics (AIAA) retained the DoD definition for validation, but changed verification to be (1998):

Verification: the process of determining that a model implementation accurately represents the developer's conceptual description of the model ***and the solution to the model****.*

While these definitions have important differences, it is worth noting that there is compatibility. In fact, a pithy definition of V & V that is consistent with the above definitions is:

Verification is doing things right and validation is doing the right things.

This pair of brief definitions omits important details provided by SCS, DoD and AIAA, but it is compatible with all three.

There are important lessons to be learned from the history of standardizing V & V definitions. Here we highlight two important points that are relevant to our discussion of reproducible and replicable:

1. Differences in communities and their purposes will lead to differences in the detailed definitions of V & V.
2. While there are differences, there is some degree of compatibility, as demonstrated by the pithy definition of V & V, which captures the essence of the SCS, DoD and AIAA definitions.

In fact, it is interesting to note that other pairs of terms can be classified as "doing things right" vs. "doing the right things." For example, management is doing things right and leadership is doing the right things. Other pairs are efficiency and effectiveness, and tactics and strategy

(Heroux, 2018). In all of these cases, doing things right is an internal consistency focus, while doing the right things is an external focus. The internal-external complementarity framework, which underpins the various definitions of V & V starkly illustrate the profound incompatibility between the Claerbout and ACM definitions of reproducible and replicable. For Claerbout, reproducible is doing things right and replicable is doing the right things. For ACM, the definitions are opposite.

4. ADDRESSING THE CHALLENGE

For scientists who belong to both the computational and computing sciences, the incompatibility of the Claerbout and ACM taxonomies is a serious concern. Each time we write a paper, give a presentation or have a conversation, we must make it clear which taxonomy we are using, or at least which basic definition for reproducible and replicable we are using. We may even face the need to use one taxonomy when publishing a paper for a particular journal or conference, and then the other for a different paper. This is certainly the situation at this time.

4.1. Raising Awareness

Because the Claerbout and ACM taxonomies are both in common use today, the first objective of this report is to raise community awareness of their incompatible definitions of reproducible and replicable. Scientists from the computational and computing sciences must be aware of the incompatibility in order to avoid serious misunderstandings in scientific discourse.

Therefore, it bears repeating that:

> **The definitions of reproducible and replicable developed and promoted by Jon Claerbout, David Donoho and Roger Peng in their publications are very similar and widely used in many scientific fields, including the computational sciences. As shown in Table 1, reproducible in this taxonomy indicates a weaker sense than replicable.**
>
> **In contrast, the definitions of reproducible and replicable developed and promoted by ACM are reversed and completely incompatible with the Claearbout/Donoho/Peng definitions. In the ACM taxonomy, replicable indicates a weaker level of confirmation than reproducible.**

If the computational and computing sciences communities are well aware of these incompatibilities, scientists can have fruitful exchanges as long as they are prefaced with a common understanding of the active taxonomy being used. As observed by LeVeque, Mitchell and Stodden (LeVeque, et. al. 2012), lack of awareness leads to confusion at meetings where authors use reproducible and replicable with opposite definitions, but provide the audience with no indication of which taxonomy is being used. The urgency of raising awareness and declaring which taxonomy is being used increases as other communities begin to focus on these topics. For example, LeVeque (LeVeque 2013), writing for the *SIAM News*, a publication of the Society for Industrial and Applied Mathematics (SIAM), implicitly adopts the ACM taxonomy by using replicate in a way that is consistent with ACM.

4.2. Revising the ACM Taxonomy

While raising awareness is sufficient for reproducibility efforts to successfully proceed in computational and computing sciences, we believe it is far better to arrive at compatible definitions of reproducible and replicable. Note that compatibility does not mean a common, detailed definition across these communities. In particular, we believe that swapping the ACM definitions of reproducible and replicable would be sufficient and would enable the computational and computing sciences communities to have effective discourse about reproducibility concerns.

We believe it would be in the best interests of the computational and computing sciences communities to swap the ACM definitions of reproducible and replicable for the following reasons:

1. Barba (2018) shows the historical precedence of the Claerbout taxonomy: Even though the ACM taxonomy is based on well-established metrology traditions, the extrapolation of metrology concepts to reproducibility in computing is new and has resulted in a taxonomy that is incompatible with existing use in computational science. To the best of our knowledge, extrapolating metrology concepts as the basis for the ACM taxonomy for reproducible and replicable has no precedent prior to the ACM effort discussed in this report.

2. If ACM were to swap its definitions of reproducible and replicable, the reversal could hopefully extend to its digital publications from previous years. It is already the case that ACM badges are applied to existing digital publications after the article has been already published in the digital library. We hope that a similar approach could be applied to achieve the reversal.

3. Swapping the ACM definitions of reproducible and replicable would minimally impact the ACM badging system since it would only require the Results Replicated and Results Reproduced badges to be swapped.

4. The benefits of quickly obtaining a compatible reproducibility taxonomy across the computational and computing sciences should far outweigh the cost and complications of implementing the swap at this early phase of ACM's approach. As other communities such as SIAM increase their focus on reproducibility concerns, the cost of change becomes higher and harder to implement. As mentioned previously, we have at least one example of SIAM (LeVeque, 2013) using the ACM taxonomy. SIAM serves both the computational and computing sciences communities. If they move in the direction of adopting the ACM taxonomy, we will experience further confusion. Achieving compatibility will only become more challenging as awareness of ACM's taxonomy grows with the success of its badging initiative.

5. CONCLUSIONS & NEXT STEPS

The computational and computing sciences have incompatible taxonomies for reproducibility, as shown in Table 1. Raising awareness of this incompatibility is essential for enabling effective discourse across these communities and avoiding needless confusion and repeated discovery of these incompatibilities by each individual scientist. We hope this report will aid in raising awareness across these communities.

While raising awareness is essential, we believe the computational and computing sciences communities have a timely opportunity to obtain compatible reproducibility taxonomies. We believe that delaying efforts to reach compatibility will only result in increased cost and effort in the future and challenging discourse across these communities in the meantime.

We believe that swapping the definitions of reproducible and replicable in the ACM taxonomy is the best approach to reach compatibility. This change will not be easy. If required, we recommend establishing a standards effort that would foster this change. There is surely more than one path to establishing a standard. ACM as an organization does not promote or establish standards. However, the IEEE does have a standards development process. Given the past collaborations between ACM and IEEE in promoting computing, we hope that an IEEE-sponsored effort to establish a standard reproducibility taxonomy could provide the means by which the computational and computing sciences could achieve compatibility.

REFERENCES

1. Barba, Lorena A. (2018). Terminologies for reproducible research, *arXiv preprint* arXiv:1802.03311.
2. Buckheit, J. B. and D. L. Donoho (1995). Wavelab and reproducible research. In Antoniadis A., O. G., ed., *Wavelets and Statistics*, Springer, New York, NY, vol. 103 of Lecture Notes in Statistics, pp. 55–81, doi:10.1007/978-1-4612-2544-7_5.
3. Cartwright, N. (1991). Replicability, Reproducibility, and Robustness: Comments on Harry Collins. *History of Political Economy*, vol. 23, no. 1, pp. 143–155., doi:10.1215/00182702-23-1-143.
4. Claerbout, J. F. and M. Karrenbach (1992). Electronic documents give reproducible research a new meaning, in *SEG Technical Program Expanded Abstracts 1992*, Society of Exploration Geophysicists, pp. 601–604, doi: 10.1190/1.1822162.
5. Heroux, M. A. (2018). Building Trusted Scientific Software. *Better Scientific Software.* Retrieved from https://bssw.io/blog_posts/building-trusted-scientific-software/.
6. Jasny, B. R., G. Chin, L. Chong, and S. Vignieri (2011). Again, and again, and again..., *Science*, 334(6060), 1225, doi:10.1126/science.334.6060.1225.
7. JCGM (2008): The international vocabulary of metrology—basic and general concepts and associated terms (VIM). Joint Committee for Guides in Metrology, 3rd ed. JCGM 200: 2012, https://www.bipm.org/utils/common/documents/jcgm/JCGM_200_2012.pdf.
8. Kitzes, J., D. Turek, and F. Deniz (2017). *The practice of reproducible research: Case studies and lessons from the data-intensive sciences.* University of California Press, ISBN: 9780520294752.
9. Koenker, R. and A. Zeileis (2009) On reproducible econometric research, *Journal of Applied Econometrics*, 24(5), 833–847, doi:10.1002/jae.1083.
10. Laine, C., S. N. Goodman, M. E. Griswold, and H. C. Sox (2007) Reproducible research: Moving toward research the public can really trust, *Annals of Internal Medicine*, 146(6), 450–453, doi:10.7326/0003-4819- 146-6-200703200-00154.
11. LeVeque, R. J., I. M. Mitchell and V. Stodden (2012), Reproducible Research for Scientific Computing: Tools and Strategies for Changing the Culture, *Computing in Science & Engineering*. Vol. 14, No. 4, p. 13-17.
12. LeVeque, R. J. (2013). Top Ten Reasons to Not Share Your Code (and why you should anyway), *SIAM News*, 46(3).
13. Liberman, M. (2015). Replicability vs. reproducibility — or is it the other way around? *The Language Log*, http://languagelog.ldc.upenn.edu/nll/?p=21956.
14. Oberkampf, W., & Roy, C. (2010). *Verification and validation in scientific computing*. New York: Cambridge University Press.
15. Peng, R. D. (2009) Reproducible research and biostatistics. *Biostatistics*, 10(3), 405–408, doi:10.1093/biostatistics/kxp014.

16. Peng, R. D. (2011). Reproducible research in computational science, *Science*, 334(6060), 1226–1227, doi:10.1126/science.1213847.

17. Peng, R. D., F. Dominici, and S. L. Zeger, (2006). Reproducible epidemiologic research, *American Journal of Epidemiology*, 163(9), 783–789, doi:10.1093/aje/kwj093.

18. Sandve, G. K., A. Nekrutenko, J. Taylor, and E. Hovig (2013). Ten simple rules for reproducible computational research, *PLoS Computational Biology*, 9(10), e1003285, doi:10.1371/journal.pcbi.1003285.

19. Schwab, M., M. Karrenbach, and J. Claerbout (2000). Making scientific computations reproducible, *Computing in Science & Engineering*, 2(6), 61–67, doi:10.1109/5992.881708.

20. Stevens, J. R. (2017). Replicability and reproducibility in comparative psychology, *Frontiers in Psychology*, 8, 862, doi:10.3389/fpsyg.2017.00862.

21. Stodden, V., F. Leisch, and R. D. Peng, eds., 2014: *Implementing Reproducible Research*. CRC Press, ISBN 9781466561595, preprint at osf.io/s9tya/.

22. Stodden, V., D.H. Bailey, J. Borwein, R.J. LeVeque, W. Rider, and W. Stein, *Setting the default to reproducible: Reproducibility in computational and experimental mathematics*, February 2, 2013; https://icerm.brown.edu/topical_workshops/tw12-5-rcem/icerm_report.pdf

23. Van Bavel, J. (2016). *Why do so many studies fail to replicate?* (Sunday Review Desk)(Gray Matter)(Column). *The New York Times, 10(l)*, 10.

DISTRIBUTION

4	Lawrence Livermore National Laboratory Attn: N. Dunipace (1) P.O. Box 808, MS L-795 Livermore, CA 94551-0808		
1	MS1324	S. Scott Collis	01400
1	MS0899	Technical Library	9536 (electronic copy)

Sandia National Laboratories