Research Practices for Scientific Rigor: A Resource for Discussion, Training, and Practice

Among human endeavors, science is unique because it yields progress: science advances our understanding of nature, yields new technologies, and improves human health. Progress depends on the consistent application of the highest standards in research methodologies, including experimental design, statistical analysis and reporting, and scientific communication, which we collectively refer to as "scientific rigor."

Recent scientific and public reports[e.g., 1–6] have raised concerns about lower-than-expected rates of replication, especially in preclinical research, and some have raised the serious question of whether the self-correcting nature of the scientific enterprise is being undercut. Funding agencies, scientific journals, professional organizations, and institutions have begun to examine the factors that underlie such concerns. Increasingly, the broad scientific community has become focused on the need to maintain and enhance rigor as part of our collective responsibility to the integrity of the scientific mission.

The Society for Neuroscience (SfN), like many professional societies, is evaluating scientific rigor within our own field. The Society is committed to helping ensure that neuroscientists are well trained in best research practices, that those practices are consistently adhered to, and that methods and design issues are reported in such a way so as to permit appropriate evaluation of results and facilitate attempts at replication. The document that follows is intended to support SfN members in our shared commitment to adhere strongly to principles of scientific rigor.

Resources for Neuroscientists
The SfN Scientific Rigor Working Group has developed the following set of research practices to serve as a foundation for ongoing field discussion. Additionally, the working group and other SfN leaders are encouraging awareness and discussion of scientific rigor in neuroscience through many SfN scientific venues and training forums, ranging from the SfN annual meeting and our journals to professional development and training programs.

The Research Practices section below is not a complete list of all research practices and is intended to serve as a foundation for trainees and experienced scientists alike to reference and use as the basis for conversation, training, and practice. Some of these practices are already established or straightforward, whereas other issues may be more complicated or unresolved. It is important to note that it may not always be possible to strictly adhere to every guideline, and that resulting research can still be rigorous. Given the complexity of these issues as they relate to individual research questions, SfN encourages full transparency, consideration, and communication regarding the implications of the choices made during research.

Policy Considerations
The Working Group stresses that the rigorous conduct of science can be influenced by

other factors beyond the actual conduct of science. Factors that warrant ongoing discussion and action by the field include:

- support for publication of negative results or results deemed inadequately "exciting," or "novel"[7–10];
- avoidance of "rushing" findings into publication without full investigation and proper self-replication[6,9];
- increasing incentives to retract incorrect or unreproducible findings[9];
- providing incentives and/or funding to perform replications[1,11–13];
- consideration of the proper balance between increasing numbers of animals for replication and the goals of "replacement, reduction, and refinement" in animal research[14];
- minimization of incentives that drive research conducted for reasons other than pursuit of truth (academic promotions, "publish or perish")[8]; and
- consideration of ways to counter the emerging trend in the peer review process in which additional experiments are requested on an abbreviated timeline, and pressure for results to be interpreted in ways that conform to previously-reached conclusions[15,16].

Research Practices
The Working Group recommends consistent attention and discussion regarding the following practices:

Experimental Design includes subject selection, use of controls, and other methodological concerns.

Topic	Possible Approaches
Unbiased sampling and data collection	• Systematic random sampling for all data collection, including selection of subjects, brain areas, cells, or cell parts (e.g., for behavior, stereology, neurophysiology, etc.) — any time the entire population isn't used[7,17]. • Use of methods to eliminate/minimize bias in experimental procedures[17]. For example, identity of samples and/or subjects should be blinded[7]; unbiased (usually random) assignment of subjects and/or stimuli to experimental groups; and timing of experiments balanced to account for sources of bias over time[18] (e.g., evolution of surgical skills, fatigue, change in personnel, test-order effects).
Experimental approach	• Use of positive and negative controls[18]. Use of replicate samples (including both technical and biologic replicates[7]) for experimental groups, when appropriate. • Use of validated and/or well-characterized reagents[1,18] (such as antibodies and pharmacological agents), procedures, and behavioral tasks[19,20]. • Consider limitations of models (behavioral, animal, cellular, etc.)[8,19], including possible contributions of genetic background[20]; selection of injury/disease models with reliability and validity[19,21]; recognition of pitfalls and caveats. • Work to established standards of the field, integrating industry or other perspectives when appropriate[22,23].

Thorough characterization of experimental effect	• Repetition of experiments within the laboratory to reduce likelihood of statistical flukes[18]. • Exploration of robustness or lack thereof across cell lines[18], animal models[24], species, and individual investigators, including use of multiple approaches or tools to interrogate specific mechanisms (molecular, cellular, circuits, etc.).

Data Analysis includes correct collection and analysis of data, and use of appropriate statistics and sample sizes.[7]

Topic	Possible Approaches
Where possible, analyses should be pre-planned[25]	• Pre-determination of sample size (power analyses) for each planned analysis to ensure experiments are appropriately powered[7,12–14,26,27]. • Pre-identification of stopping points, in order to avoid testing to a foregone conclusion (i.e., increasing "n" until a significant effect is seen or stopping the experiment once a significant difference is seen)[7,13,14,26].
Post-experiment data analyses	• Follow best practices in pooling of data across experiments (e.g., data collected at different times, collapsed across planned assessment points, or from different experimental groups)[28–32]. • Use of pre-defined procedures/criteria to deal with attrition or other missing data and data exclusion (of individual points or complete data sets)[7,13,33–35].
Statistical design[18]	• Selection of appropriate statistical tests (including testing of statistical assumptions, such as normality of data)[13]. • Control for multiple comparisons[7,13]. • Avoid "significance chasing" such as interpreting the data in different ways so that it passes the statistical test of significance or analyzing different measures until finding one on which groups differ[13]. • Distinguish between non-hypothesis driven, discovery experiments and those designed to directly test a proposed hypothesis[7].

Transparency includes reporting, publishing, or providing access to specific data, methods, or analyses.[7]

Topic	Possible Approaches
Data preservation	• Complete primary data set backed up and protected against alterations[36]. • Information security applied for sensitive material (videos of animals or for human subjects protection)[36]. • Data accessible to data owners and available to outside investigators if necessary[12,13,36].
Full transparency in data and methods reporting	• Report full details on methods and experimental design, such as timing of experiments, compilation of groups from experiments done over time, technical and biological replicates, animal surgeries and functional testing, methods

	for randomization and blinding, changes in staff carrying out experimental procedures, and self-replication efforts[5,7,13,26,37].
	• Fully report results of all analyses done as part of an experiment (including statistical controls for multiple comparisons and identification of pre- and post-hoc analyses) and any pooling of data from experiments done at different times[5,12,13,18,26].
	• Full transparency in collaborations between groups (to avoid false appearance of independence of findings and other issues related to conflicts of interest)[38].

Sources Cited

1. Wadman, M. NIH mulls rules for validating key results. *Nature* **500,** 14–16 (2013).
2. Trouble at the lab. *The Economist* (2013). at <http://www.economist.com/news/briefing/21588057-scientists-think-science-self-correcting-alarming-degree-it-not-trouble>
3. Hiltzik, M. Science has lost its way, at a big cost to humanity. *Los Angeles Times* (2013). at <http://articles.latimes.com/2013/oct/27/business/la-fi-hiltzik-20131027>
4. Couzin-Frankel, J. The Power of Negative Thinking. *Science* **342,** 68–69 (2013).
5. Collins, F. S. & Tabak, L. A. Policy: NIH plans to enhance reproducibility. *Nature* **505,** 612–613 (2014).
6. Zimmer, C. Rise in Scientific Journal Retractions Prompts Calls for Reform. *The New York Times* (2012). at <http://www.nytimes.com/2012/04/17/science/rise-in-scientific-journal-retractions-prompts-calls-for-reform.html>
7. Landis, S. C. *et al.* A call for transparent reporting to optimize the predictive value of preclinical research. *Nature* **490,** 187–191 (2012).
8. Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483,** 531–533 (2012).
9. Fanelli, D. Why Growing Retractions Are (Mostly) a Good Sign. *PLoS Med* **10,** e1001563 (2013).
10. Franco, A., Malhotra, N. & Simonovits, G. Publication bias in the social sciences: Unlocking the file drawer. *Science* 1255484 (2014). doi:10.1126/science.1255484
11. Steward, O., Popovich, P. G., Dietrich, W. D. & Kleitman, N. Replication and reproducibility in spinal cord injury research. *Exp. Neurol.* **233,** 597–605 (2012).
12. Button, K. S. *et al.* Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* **14,** 365–376 (2013).
13. Funder, D. C. *et al.* Improving the Dependability of Research in Personality and Social Psychology Recommendations for Research and Educational Practice. *Personal. Soc. Psychol. Rev.* **18,** 3–12 (2014).
14. Fitts, D. A. Ethics and animal numbers: informal analyses, uncertain sample sizes, inefficient replications, and type I errors. *J. Am. Assoc. Lab. Anim. Sci. JAALAS* **50,** 445–453 (2011).
15. Ploegh, H. End the wasteful tyranny of reviewer experiments. *Nat. News* **472,** 391–391 (2011).
16. Rockman, H. A. Waste not, want not. *J. Clin. Invest.* **124,** 463–463 (2014).
17. Ioannidis, J. P. A. Why Most Published Research Findings Are False. *PLoS Med* **2,** e124 (2005).

18. Begley, C. G. Reproducibility: Six red flags for suspect work. *Nature* **497,** 433–434 (2013).
19. Nestler, E. J. & Hyman, S. E. Animal models of neuropsychiatric disorders. *Nat. Neurosci.* **13,** 1161–1169 (2010).
20. Wahlsten, D. *et al.* Different data from different labs: lessons from studies of gene-environment interaction. *J. Neurobiol.* **54,** 283–311 (2003).
21. Willner, P. & Mitchell, P. J. The validity of animal models of predisposition to depression. *Behav. Pharmacol.* **13,** 169–188 (2002).
22. Schulz, K. F., Altman, D. G. & Moher, D. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMC Med.* **8,** 18 (2010).
23. Lapchak, P. A. Scientific Rigor Recommendations for Optimizing the Clinical Applicability of Translational Research. *J. Neurol. Neurophysiol.* **3,** (2012).
24. De Souza, N. Model organisms: Mouse models challenged. *Nat. Methods* **10,** 288–288 (2013).
25. Ruxton, G. D. & Beauchamp, G. Time for some a priori thinking about post hoc testing. *Behav. Ecol.* **19,** 690–693 (2008).
26. Simmons, J. P., Nelson, L. D. & Simonsohn, U. False-Positive Psychology Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychol. Sci.* **22,** 1359–1366 (2011).
27. Cortina, J. M. & Landis, R. S. The Earth Is Not Round (p = .00). *Organ. Res. Methods* **14,** 332–349 (2011).
28. Costafreda, S. G. Pooling fMRI Data: Meta-Analysis, Mega-Analysis and Multi-Center Studies. *Front. Neuroinformatics* **3,** (2009).
29. Demets, D. L. Methods for combining randomized clinical trials: Strengths and limitations. *Stat. Med.* **6,** 341–348 (1987).
30. Machlis, L., Dodd, P. W. D. & Fentress, J. C. The Pooling Fallacy: Problems Arising when Individuals Contribute More than One Observation to the Data Set. *Z. Für Tierpsychol.* **68,** 201–214 (1985).
31. Kendziorski, C., Irizarry, R. A., Chen, K.-S., Haag, J. D. & Gould, M. N. On the utility of pooling biological samples in microarray experiments. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 4252–4257 (2005).
32. Imbeaud, S. & Auffray, C. 'The 39 steps' in gene expression profiling: critical issues and proposed best practices for microarray experiments. *Drug Discov. Today* **10,** 1175–1182 (2005).
33. Schafer, J. L. & Graham, J. W. Missing data: Our view of the state of the art. *Psychol. Methods* **7,** 147–177 (2002).
34. Graham, J. W., Cumsille, P. E. & Elek-Fisk, E. in *Handbook of Psychology* (John Wiley & Sons, Inc., 2003). at <http://onlinelibrary.wiley.com/doi/10.1002/0471264385.wei0204/abstract>
35. Twisk, J. & de Vente, W. Attrition in longitudinal studies: How to deal with missing data. *J. Clin. Epidemiol.* **55,** 329–337 (2002).
36. Tenopir, C. *et al.* Data Sharing by Scientists: Practices and Perceptions. *PLoS ONE* **6,** e21101 (2011).
37. Simard, J. M. & Gerzanich, V. When replication teaches more than the original experiment--the saga of the unknown unknown. *Exp. Neurol.* **233,** 623–624 (2012).
38. Viswanathan, M. *et al.* A proposed approach may help systematic reviews retain needed expertise while minimizing bias from nonfinancial conflicts of interest. *J. Clin. Epidemiol.* (2014). doi:10.1016/j.jclinepi.2014.02.023