

Rigor and Transparency in the Conduct of Research at USC

April 7, 2016

Background:

The topic of reproducibility in research has gained progressive attention in both the academic community and media in recent years. This has resulted in a focus on the need for improved rigor and transparency in research while raising concerns regarding investment of significant resources in research that cannot be reproduced. As an example, one recent analysis estimated that \$28 billion is spent by the United States each year on basic biological research that is irreproducible.¹ More importantly, however, is the pursuit of faulty leads that result from irreproducible research, where billions of dollars are lost on following up on research that is ultimately not accurate and had led to misguided scientific inquiries. Although some propose that research that is not reproducible can still be accurate, the concerns regarding irreproducible research are difficult to dispel, and the importance of improving rigor and transparency in research is inescapable. Research is defined by a systematic investigation, including development, testing and evaluation, and is usually designed to advance generalizable knowledge. Generating research that is both robust and transparent may therefore ensure that the research output can be reproducible and thus generalizable.

The reproducibility debate has emphasized that irreproducible results are rarely caused by scientific misconduct. Instead, a complex host of factors are responsible, including: 1.) Researchers are often poorly trained in experimental design and modern statistical methods, 2.) An emphasis by journals on making provocative statements from research studies has resulted in a bias towards publication of positive results and a paucity of venues to publish negative data or papers that point out scientific flaws in previously published work, 3.) Publications not reporting basic elements of experimental design, including study design, laboratory protocols, biological reagents and reference materials, and data analysis and reporting has resulted in these elements frequently being ignored. 4.) Journals, funding agencies and academic institutions have failed to establish or enforce policies that insist on access to research data. 5.) Authentication of materials and resources (such as biological and chemical agents) is rarely performed. 6.) No systems are in place to encourage or require the replication of research results.

Although many of these factors are not directly related to the policies of academic centers, a number of concerns exist at the institutional level. Most importantly, academic centers place a high value on journal publications (such as for achieving promotion and tenure, enhancing academic reputation, and facilitating grant funding) and may overvalue the publication of

research findings relative to assuring research accuracy. Due to the inclination of many journals, especially high-impact journals, for reporting positive over negative results, the resulting bias is pervasive in research performed at academic institutions. In addition, in the current budgetary environment, rapid submission of research results is also encouraged to the detriment of careful efforts at replicating findings.

Funding agencies such as the National Institutes of Health (NIH) and the National Science Foundation (NSF) have brought attention to the matter of reproducibility and changed their policies to address it. In addition, scientific journals (including high-profile journals such as *Science* and *Nature*) have also taken steps to revise their review and acceptance policies to address this issue. The public focus remains on this topic, however, and lack of reproducibility in research continues to be highlighted by high-profile studies on reproducibility that generate significant media interest (such as a recent study on the very low reproducibility rates of psychology research articles).

Academic centers such as USC have an important role to play in fostering an environment of rigorous and transparent research. The University Research Committee (URC) believes that USC can be an international leader among academic institutions in assuring rigor and transparency in research. Through the recommendations that follow, we aim to create an innovative institutional framework through which we raise the quality and impact of research produced at USC. By so doing, we assist the broader research community in maintaining the public trust in research investments and demonstrate that academic research produces results that are novel, impactful, and accurate.

Recommendations:

USC should endorse a rigorous and transparent approach to research. Towards that end, USC should commit to encouraging *good practices* in research, thereby maximizing the reproducibility of research conducted at USC. In order to fulfill this commitment, USC should provide the necessary infrastructure and resources towards supporting these efforts by its research community, including:

1.) **Promoting Transparency:** Detailed descriptions of collection methods, raw data and research analyses better accommodate future efforts to reproduce research results. Sharing of this type of data and information can therefore be crucial for efforts by others to reproduce and validate research. In order to disseminate this information, an appropriate medium for reporting such data is required. USC should continue to support the infrastructure for this form of data sharing as follows:

a.) Data storage and sharing sites should be made available for researchers to have ready access to resources for publicly sharing data that make this type of transparency possible (addressed in the URC document on Academic Computing). The infrastructure and software for data sharing are often far more complex than mere storage of data, and a structure for facilitating data sharing for the highly variable research that occurs in the USC research community is required. Such data sharing networks are freely available and should be considered as sites for data sharing at USC.

b.) When feasible for hypothesis-driven research, an electronic disclosure of the research hypotheses should be logged prior to the onset of data collection. This is analogous to current expectations of federally and industry-funded research through clinicaltrials.org, where disclosure of the hypothesis is mandated prior to the clinical study. A distinction can often be made between discovery research and confirmation research. Although generation of a clear hypothesis is not possible for many forms of discovery research, project entry (whether private or public) into a database still allows the opportunity to pre-register a project prior to data collection. The USC research community should be encouraged to pre-register all research projects with an expectation that the information included (including hypothesis) is dependent on the type of research being performed.

c.) Beyond individual use by researchers at the university, USC should consider a broader collaboration with a data sharing partner whereby such data sharing is encouraged and consolidated. In addition to the additional project management support and training possible with this arrangement, a collaborative partnership can encourage USC's researchers to adhere to these principles while demonstrating USC's true commitment to transparency in research.

2.) Encouraging Good Institutional Practices: USC should support the following practices which encourage research integrity and foster objectivity:

a.) Training programs and courses in rigorous experimental design, research standards, statistics, meta-analyses, and objective evaluation of data should be offered by the university. An exclusive reliance on significance testing (P-values) in data analysis has been a focus of the reproducibility discussion, and each school should ensure that its courses, including statistical classes, remain current and up to date. Similarly, appropriate methodological training should be made available to bolster rigorous research that is not experimental or quantitative in nature. This includes ensuring that courses are made available to qualitative social scientists and researchers in the humanities that teach the most appropriate methods to address their questions of interest. In addition, education in the "Responsible Conduct of Research" should be encouraged throughout the university.

b.) Compliance with research standards, including data-sharing, should be supported by the university. A reporting system through the Office of Research should be charged with providing a forum for informal and confidential discussions regarding research practices at the university. A mechanism for formal reporting of concerns should be established with an expectation for impartial review of these reports.

c.) A method for systematic data collection should be sought at the institutional level. This could potentially be accomplished by use of "electronic laboratory notebooks" which can be used to facilitate collaboration, supervision and record keeping and can link records to the original data. Although these notebooks do not need to be open or publicly shared, they provided a medium for recording research progress.

3.) Consideration in Merit Review and Promotion: While traditional metrics of academic achievement, including journal publications and grant funding, remain important measures in the evaluation of merit and promotion, strategies that further encourage robustness of research

design, data and code sharing and high-quality mentoring should also be considered by the University. Infusing these incentives into the research culture of USC is the ultimate objective, and this can be facilitated if coupled to the processes of merit evaluation and promotion. The URC therefore encourages inclusion of the principals of rigor and transparency in research in the merit review and promotion process. This includes a strong consideration of the impact of research work, specifically as it relates to the investigator's work later being validated or reproduced. In addition, a process whereby faculty are encouraged and incentivized to document their personal efforts and successes in achieving transparency and rigor as it applies to their research work should be considered. A system for including metrics of research work not published in the traditional research format, such as data collection and hosted data sets, could also be established, especially when the data is more formally published and used by other investigators. During this process, researchers should continue to recognize and consider the tradeoffs of such public disclosure. Lastly, research projects specifically designed to systematically validate or reproduce original research should be understood to have value. These reproducibility efforts should be evaluated with a consideration that although this type of work may result in less academic success (presentations, publications, follow-on funding) than "ground-breaking" research, it nonetheless can provide an important scholarly service and be important to the advancement of knowledge.

4.) Participating in Reproducibility Work: Efforts to reproduce prior research have not traditionally been encouraged at academic institutions as a research or teaching priority. This type of research has not been able to garner the type of publication success expected from ground-breaking original research, nor is it a typical recipient of external funding. As a result, and due to the emphasis and "requirements" for authorship in high-profile journals for academic promotion, reproducing previous research is rarely performed. The recent statement "Innovation points out paths that are possible; replication points out paths that are likely; progress relies on both" nonetheless emphasizes the need to validate research directions prior to misdirected assumptions of truth.² USC should therefore encourage its researchers to both participate in reproducibility efforts and to report and share efforts to reproduce research as follows:

- a.) Incentives to encourage participation in reproducibility work should be established at the university, where consideration of the importance of this work is given, including in the merit evaluations and consideration for promotion as stated above.
- b.) In many fields, efforts at replicating prior data are already often performed at the initiation of research projects in order to ensure that a laboratory has the capacity to undertake work in a given area. These authentication efforts should be included as technical reports and be publicly shared and available for evaluation outside of the institution.
- c.) As an institution, USC should make every effort to encourage external funding sources to support the validation of others' research, either as stand-alone efforts or as a preliminary phase of new research. This includes encouraging external funding sources to allocate funds for this specific issue, to accommodate this work as a deliverable in broader research, and/or to support these types of investigations via internal resources.

5.) Increasing Visibility of the Topic of Reproducibility: Although there have been some concerns by the wider academic community that a strong focus on reproducibility in research

may result in a vigilante-like policing and censoring of research, USC is committed to directly addressing this topic of reproducibility in research (including rigor and transparency) with an academic fervor and using resources available only at an academic institution such as USC. Given this commitment, the URC makes the following suggestions:

- a.) The topic of reproducibility in research should be incorporated in the broader curriculum throughout the schools and departments of USC. These could include courses specifically addressing a historical perspective on misguided research directions as well as ideas around enhancing research practices.
- b.) A focus on transparency and rigor in research practices should be both discussed and encouraged amongst all members of the research community, including students, technicians, assistants, and other staff. These topics should therefore be addressed during group meetings of individual research teams.
- c.) Speaker series and other lecture series should be encouraged to include the topic of reproducibility.
- d.) The institution should host debates on the topic of reproducibility, allowing well thought out and comprehensive discussions of the pros and cons of methods for ensuring reproducibility in the conduct of research.
- e.) A collaborative approach to the topic should be encouraged (e.g., a focus on reproducibility as a funded project via the USC Collaboration Fund).

6.) Authenticating Key Resources: Reproducibility depends on controlling the accuracy and consistency of all research inputs and research methods. There has been particular attention on the finding that cell lines reported to contain certain biological information are often not accurate and result in erroneous research conclusions. The NIH has recognized the importance of this issue, and starting in January of 2016, all NIH research proposals require an explanation of the steps that will be taken to authenticate cell lines prior to initiating research. Although this type of authentication may appear time-consuming and costly, the NIH mandate now makes at least the consideration of cell line authentication necessary. A similar authentication effort may therefore be expected for other key biological and chemical resources (including chemicals, antibodies and other biologics).³ Furthermore, a broader view on this issue involves the validation of all key inputs used in research (including software code, physical materials and devices) as well as the importance of precisely calibrating research tools.⁴ An understanding of the complexity and variability of systems is mandatory, and the responsibility of considering the validity of all input materials and models should remain with the research investigator. The university, however, should encourage and support these efforts to authenticate key research resources as follows:

- a.) USC should commit resources to building an infrastructure (including external sources when appropriate) for authenticating cell lines and other biological and chemical resources upon arrival at the university. Many of these types of tests are already readily available, but should be vetted on a departmental / school level with an expectation that these authentication practices are standardized by the university.

b.) Periodic calibration and testing of research tools should be made systematic and traceable by researchers and the university.

c.) In most cases, funding for authentication of biological and chemical resources as well as system and instrument calibration and quality checks will be available through sponsored agreements. If funding is not available to the PI for these authentication services, the university should commit to paying for this type of authentication.

d.) The consideration and efforts to validate all research inputs (including biological or chemical, hardware or software) should be clearly documented by the investigator. The closed electronic digital notebook described above can be used as a medium for recording these practices, but consideration should be given to more open and public reporting of authentication efforts.

References:

- 1.) Freedman LP, Cockburn IM, Simcoe TS. The Economics of Reproducibility in Preclinical Research. PLoS Biol. 2015 Jun 9;13(6):e1002165.
- 2.) Open Science Collaboration. PSYCHOLOGY. Estimating the reproducibility of psychological science. Science. 2015 Aug 28;349(6251):aac4716.
- 3.) Enhancing Research Reproducibility: Recommendations from the Federation of American Societies for Experimental Biology. 2016 Jan 14.
<https://www.faseb.org/Portals/2/PDFs/opa/2016/FASEB_Enhancing%20Research%20Reproducibility.pdf>
- 4.) Baker M. How quality control could save your science. Nature. 2016 Jan 28;529(7587):456-8.