INFECTION AND IMMUNITY, Dec. 2010, p. 4972–4975 Vol. 78, No. 12
0019-9567/10/$12.00 doi:10.1128/IAI.00908-10

EDITORIAL

Reproducible Science[▽]

The reproducibility of an experimental result is a fundamental assumption in science. Yet, results that are merely confirmatory of previous findings are given low priority and can be difficult to publish. Furthermore, the complex and chaotic nature of biological systems imposes limitations on the replicability of scientific experiments. This essay explores the importance and limits of reproducibility in scientific manuscripts.

"Non-reproducible single occurrences are of no significance to science."
—Karl Popper (18)

There may be no more important issue for authors and reviewers than the question of reproducibility, a bedrock principle in the conduct and validation of experimental science. Consequently, readers, reviewers, and editors of *Infection and Immunity* can rightfully expect to see information regarding the reproducibility of experiments in the pages of this journal. Articles may describe findings with a statement that an experiment was repeated a specific number of times, with similar results. Alternatively, depending upon the nature of the experiment, the results from multiple experimental replicates might be presented individually or in combined fashion, along with an indication of experiment-to-experiment variability. For most types of experiment, there is an unstated requirement that the work be reproducible, at least once, in an independent experiment, with a strong preference for reproducibility in at least three experiments. The assumption that experimental findings are reproducible is a key criterion for acceptance of a manuscript, and the Instructions to Authors insist that "the Materials and Methods section should include sufficient technical information to allow the experiments to be repeated."

In prior essays, we have explored the adjectives descriptive (6), mechanistic (7), and important (8) as they apply to biology, and experimental science, in particular. In this essay, we explore the problem of reproducibility in science, with emphasis on the type of science is that routinely reported in *Infection and Immunity*. In exploring the topic of reproducibility, it is useful to first consider terminology. "Reproducibility" is defined by the *Oxford English Dictionary* as "the extent to which consistent results are obtained when produced repeatedly." Although it is taken for granted that scientific experiments should be reproducible, it is worth remembering that irreproducible one-time events can still be a tremendously important source of scientific information. This is particularly true for observational sciences in which inferences are made from events and processes not under an observer's control. For example, the collision of comet Shoemaker-Levy with Jupiter in July 1994 provided a bonanza of information on Jovian atmospheric dynamics and *prima facie* evidence for the threat of meteorite and comet impacts. Consequently, the criterion of reproducibility is not an essential requirement for the value of scientific information, at least in some fields. Scientists studying the evolution of life on earth must contend with their inability to repeat that magnificent experiment. Gould famously observed that if one were to "rewind the tape of life," the results would undoubtedly be different, with the likely outcome that nothing resembling ourselves would exist (12). (Note for younger scientists: it used to be fashionable to record sounds and images on metal oxide-coated tape and play them back on devices called "tape players.") This is supported by the importance of stochastic and contingent events in experimental evolutionary systems (4).

Given the requirement for reproducibility in experimental science, we face two apparent contradictions. First, published science is expected to be reproducible, yet most scientists are not interested in replicating published experiments or reading about them. Many reputable journals, including *Infection and Immunity*, are unlikely to accept manuscripts that precisely replicate published findings, despite the explicit requirement that experimental protocols must be reported in sufficient detail to allow repetition. This leads to a second paradox that published science is assumed to be reproducible, yet only rarely is the reproducibility of such work tested or known. In fact, the emphasis on reproducing experimental results becomes important only when work becomes controversial or called into doubt. Replication can even be hazardous. The German scientist Georg Wilhelm Reichmann was fatally electrocuted during an attempt to reproduce Ben Franklin's famous experiment with lightning (1). The assumption that science must be reproducible is implicit yet seldom tested, and in many systems the true reproducibility of experimental data is unknown or has not been rigorously investigated in a systematic fashion. Hence, the solidity of this bedrock assumption of experimental science lies largely in the realm of belief and trust in the integrity of the authors.

Reproducibility versus replicability. Although many biological scientists intuitively believe that the reproducibility of an experiment means that it can be replicated, Drummond makes a distinction between these two terms (9). Drummond argues that reproducibility requires changes, whereas replicability avoids them (9). In other words, reproducibility refers to a phenomenon that can be predicted to recur even when experimental conditions may vary to some degree. On the other hand, replicability describes the ability to obtain an identical result when an experiment is performed under precisely identical conditions. For biological scientists, this would appear to be an important distinction with everyday implications. For example, consider a lab attempting to reproduce another lab's finding that a certain bacterial gene confers a certain phenotype. Such an experiment might involve making gene-deficient variants, observing the effects of gene deletion on the phenotype, and, if phenotypic changes are apparent, then going further to show that gene complementation restores the original phenotype. Given a high likelihood of microevolution in microbial strains and the possibility that independently synthesized gene disruption and replacement cassettes may have sub-

[▽] Published ahead of print on 27 September 2010.