Research Reproducibility 2020
Educating for Reproducibility: Pathways to Research Integrity

THE IMPORTANCE OF QUANTITATIVE THEORY FOR REPRODUCIBLE SCIENCE

Peter D. Kvam[(1)]
[(1)]*University of Florida, pkvam@ufl.edu, @peterkvam*

ABSTRACT

Introduction: A pillar of research reproducibility is the ability to get the same results by running the same experiments under the same conditions. Results are defined in terms of the analyses we carry out: whether we compute a mean and standard deviation, or estimate a more sophisticated statistical or generative model, will determine what results we report and the accompanying conclusions we draw. If the model of the data is not appropriate, or if it makes assumptions that are not borne out, then the conclusions we draw will be systematically biased or simply wrong.

Methods: To illustrate this issue, I look at five common tasks – spanning social, clinical, and cognitive psychology – that are used to assess psychological constructs like interference, associative strength, and impulsivity.

Results: When using simple summary statistics like mean response times or accuracy to quantify behavior, these tasks appear unreliable: we do not get the same results from the same people when tested at different points in time. However, when we use a hierarchical Bayesian modeling approach that quantifies behavior in terms of the parameters of a generative model – accounting for within- and between-person variability in behavior – estimates of reliability increase by as much as .8 (on a -1 to 1 scale). Furthermore, the parameters of the generative model are better predictors of health outcomes (such as substance use), detect effects of experimental manipulations in the data where summary statistics fail, and allow us to connect data across measures (e.g., neural and behavioral data) to form a more complete picture of cognition.

Discussion: If we care about reproducibility, therefore, we must strive to incorporate modeling into scientific practices, particularly in psychology where it is often omitted. I discuss two solutions to this problem: (1) graduate and undergraduate training in computational modeling methods, and (2) the use of artificial intelligence and machine learning approaches to automate model fitting, making it more accessible to researchers without computational training. Put together, these remedies offer the opportunity to improve the reliability and validity of psychological measurement, promoting reproducibility across a variety of research domains.