

Big Data Has Left the Station

December 27, 2012



Joseph E. Steinmetz, The Ohio State University

President's Note: *A big difference in the academy today compared with the academy we knew 25 years ago is the emergence of interdisciplinary approaches to research and scholarship. While there are a number of reasons for this trend, first and foremost there has been a growing recognition that the major problems facing the world are complex, requiring varied approaches and various levels of analysis for their solutions. Indeed, the days of the "lone wolf" researcher are on the decline, and this has led to a more team-oriented approach to research and scholarship. This has resulted in the emergence of fields at the intersection of our traditional disciplines. Most notably for psychology may be the emergence of neuroscience and cognitive science as well as a recent emphasis on translational research that connects basic and clinical research efforts. I asked Rich Ivry, who chairs the psychology department at the University of California, Berkeley, to comment on recent trends he has seen develop as we have moved into the world of interdisciplinary, team science in psychology. He provided the following insights, which highlight the "Big Data" nature of modern psychological science research.* **-Joseph E.**

Steinmetz

Will psychology departments exist on university campuses in 50 years? This may seem an absurd question, especially given that so many of us are thinking about how to control over-enrolled classes and impacted majors. Yet we have seen departments that have opted for a new name (MIT's Brain and Cognitive Sciences) or self-cleaved, with the cognitive and biological members splitting off to form new departments (UCSD, Rochester, Brown).



Rich Ivry, University of California, Berkeley

These moves reflect intellectual and methodological trends, with faculty self-organizing to make explicit the shifts in the knowledge domains that inform their work, as well as forming affiliations to maximize resource utilization.

I have had an interesting vantage point from which to consider the future of our field. I coordinate a professional development class in which different faculty members come by each week to provide our first-year graduate students with insights for surviving graduate school and beyond. A universal mantra is encouragement to be interdisciplinary, to look to other fields for the knowledge and tools that will support a sophisticated and comprehensive research program. This holds true for both basic and translational work, with the line between them becoming increasingly blurred, a trend that I suspect is not solely motivated in response to funding agency mandates.

Interestingly, an implicit aspect of the push for being interdisciplinary reflects an important trend in psychology, the move toward Big Data. Reflecting our roots in philosophy, psychology has been a theory-driven discipline. Even in the heyday of behaviorism, experimental work was dictated and motivated by theory. But I strongly suspect that a data-driven approach will continue to assert itself in psychology, much as it has in other disciplines, such as physics, biology, and computer science. As my colleague, Jack Gallant, says, “theory will always get blown away by data.”

We can see the emergence of Big Data in all areas of psychology: MRI studies not only generate massive amounts of data, but require increasingly multivariate analytic tools to make sense of the small hemodynamic signals that reflect mental activity. The cracking of the human genome provides opportunities for large-scale studies to look at complex gene-environment influences on behavior. Clinical sciences are turning to objective measures, both behavioral and genetic, that lend themselves to the Big Data movement — providing new ways to characterize mental disorders that are increasingly seen as dimensional rather than categorical. Social psychology has embraced the Internet and Mechanical

Turk, harvesting exponentially larger sample sizes than have been afforded by the standard campus-based subject pool with the potential to situate research in a rich context that can begin to consider the effects of social, cultural, and economic factors. And databases are readily available to explore just about any question, ripe for the latest mining techniques. Even our traditional statistics are being supplanted by methods arising from Big Data, with arbitrary significance thresholds replaced by estimation procedures to obtain exact probabilities. All of these trends require expertise from many fields. I don't know if *Psychological Science* in 2062 will feature articles with 3,000 co-authors, approaching the record-setting contributor list on the recent Higgs boson publications in particle physics. But we can anticipate large collaborative multifaceted research programs becoming the standard, replacing the traditional single-investigator-organized lab.

Will we recognize psychology in the Big Data world? Will it continue to have a unique signature? The problems that fascinate psychologists will remain, but much as we have seen philosophers, computer scientists, engineers, and biologists take up thorny psychological problems such as consciousness and the self, our students and our students' students are unlikely to be uniquely recognized as "psychologists." They will, and should, seek the training that can enable them to master the methods and technologies that allow them to generate and analyze Big Data. This does not mean our training won't include a grounding in psychological theory. As Nate Silver warns in *The Signal and the Noise*, it can't all be about the technology, "The numbers have no way of speaking for themselves. We speak for them." But we should also be prepared to find that the kinds of theories that emerge from such enterprises may not have the same form as those that have framed our field in its first 100 plus years.