

Teaching Statistics in the Age of Open Science

November 30, 2017



With a grant from the APS Teaching Fund, **Kevin P. McIntyre** has created a website that guides students through the reproduction of analyses in Psychological Science papers that include associated open data sets. (Photo credit: Trinity University)

The benefits of open science for promoting high-quality research are clear. Preregistration of hypotheses prevents p-hacking and other questionable research practices; open materials increase the fidelity of direct and conceptual replication studies; and open data allow for greater transparency in evaluating the strength of statistical evidence in support of a particular hypothesis. An unintended, but equally beneficial, outcome of the move toward open science is that those of us who teach statistics and research methods now have the ability to incorporate open data and materials into our courses.

I came to this realization about a year ago while in a moment of panic. I needed to create an activity for the undergraduate statistics class I was teaching later that day, and I had no good ideas. Desperate for inspiration, I was flipping through the August 2015 issue of Psychological Science and found an interesting article by Ella L. James and colleagues examining whether computer games could be used to reduce the frequency of intrusive memories following a traumatic event. As I was considering the ways that I might convert this paper into an activity for my students, I happened to notice that it had an Open Data badge, indicating that the authors had made their data publicly available using the [Open Science Framework](#). Rather than creating a fictitious data set that resembled the results reported in the original paper, as I had done in the past, I decided to give my students the actual data from the paper, along with an activity that would guide them through the reproduction of the analyses reported in the paper.

During the in-class activity, my students were enthusiastic and engaged. They seemed to connect with this activity in a way that they hadn't with my previous activities (which all used fictitious data). Working with real data helped my students see how our class was preparing them to conduct

psychological research, and they found that analyzing real data was more challenging than analyzing fictitious data. Unlike textbook data sets, which often have one independent variable and one dependent variable, actual data sets have many variables, and researchers need to make difficult decisions about the best way to analyze that data. Having to think about these issues is likely to help students develop skills they can use when analyzing the data sets for their own projects.

Shortly after realizing the benefits of using open data for my own teaching, I thought that other people may want to use these open materials too. So I applied for, and received, a grant from the APS Fund for Teaching and Public Understanding of Psychological Science to create a website that provides teachers and students with papers published in Psychological Science, their associated open data sets (in SPSS and .csv formats), and activities to guide students through the reproduction of the analyses in the paper. The resulting website is called Open Stats Lab (openstatslab.com), and it launched in early 2017. The site is free to use, and SAGE Publications even makes the articles freely available, so that anyone, even those without a subscription to the journal, can use the activities.