

Jo Etzel, PhD

Staff Scientist

- [An analysis] was coming out too good; ...some of the effects were just implausibly large, and not affected by the conditions like they should have been. ...We discovered it was a communication error resulting in incorrect condition labels. I hadn't preprocessed the dataset myself; a colleague ran the experiment at a different university and we had several meetings about its files. ...I coded up the conditions as described in my notes from the meetings ... but my notes were wrong.
- This only resulted in lost time and aggravation. But it was really tricky, since some of the results came out better, not worse, and the error was in my initial task/file description (so wasn't found by me comparing the code with my notes - those matched)



Brian Carpenter, PhD

Prof of Psychology

- “I've made too many data management mistakes to name, but among the most time consuming have been data transformations, such as cases-to-variables and variables-to-cases, that I didn't document properly and then had no way of knowing what data I was working with after the transformations. Now I'm obsessive about documenting every single tiny step and change I make so I can go back and remember and recreate if needed. We work with a lot of non-independent data in our studies (e.g., multiple family members in the same data set), so those transformations get a little confusing even when my thinking is clear!”

