New stats summary:

* The main effect of big zone being different from the small zone in all of the experiments is still significant in the correct direction (smaller when image is smaller than object depicted, and larger when object is smaller than image)
* The AIC for models using the mean size from the norming is never higher than the models using prototype status
* Sometimes there is a significant effect of question order, but it is always small.
* There is never a significant effect of direction in the non-randomized conditions (whether the images are presented in ascending or descending order)
* For the experiments with tiny things (objects smaller than the image), there is never a significant effect of the mean, which is not super surprising.
* In the replication experiment, looking at the releveled models so that the prototypically big and neutral things are each compared to the prototypically small things, the prototypicality effects seem to be present, in addition to a significant interaction between adjective and prototypicality status, though I’m not sure what that means about what’s going on.
* In both the photo experiments (ordered and randomized) there mostly is no effect of prototype status, not even a difference between big and small things, though there is an effect of mean size and an interaction between adjective and mean in the model using mean size. (though not an effect of mean in randomized photos, just an interaction between adjective and mean).
* The cartoon experiment using bunnies and ducklings (instead of mice and chicks) so that it is more comparable to the photo experiment, behaves more like the cartoons (replication) than the photos, in that there is an interaction between prototype status and the adjective for both prototype statuses in the releveled models. However, I did not find a main effect of prototype status neither (compared to small), whereas I did find such a main effect with cartoons.