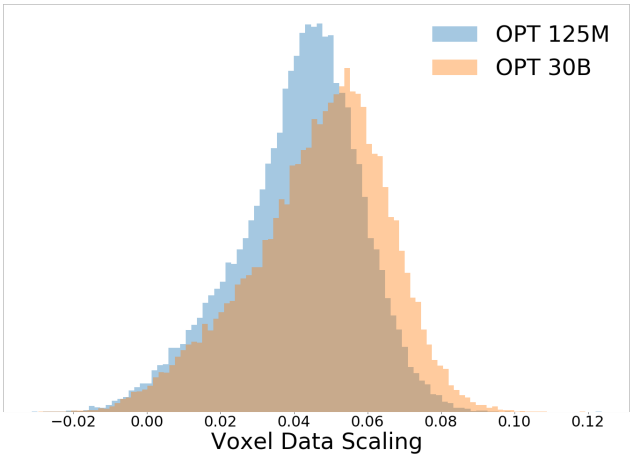
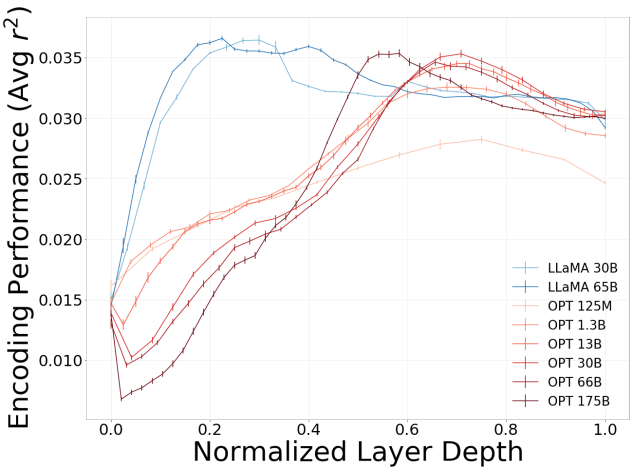
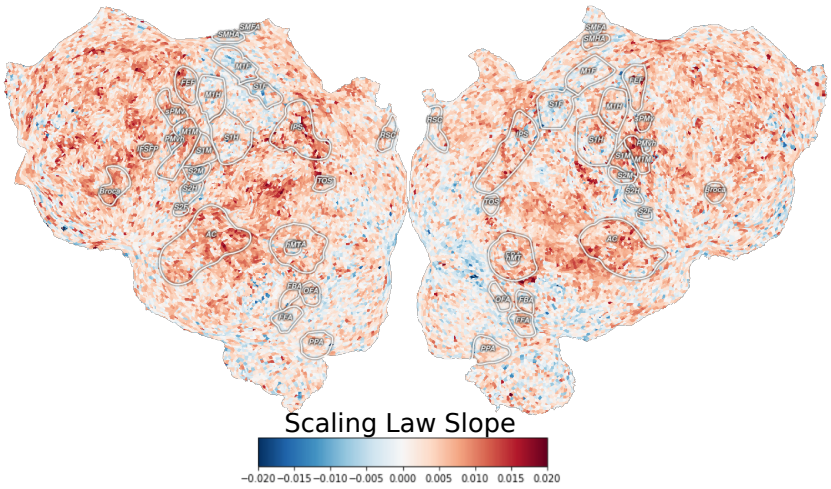


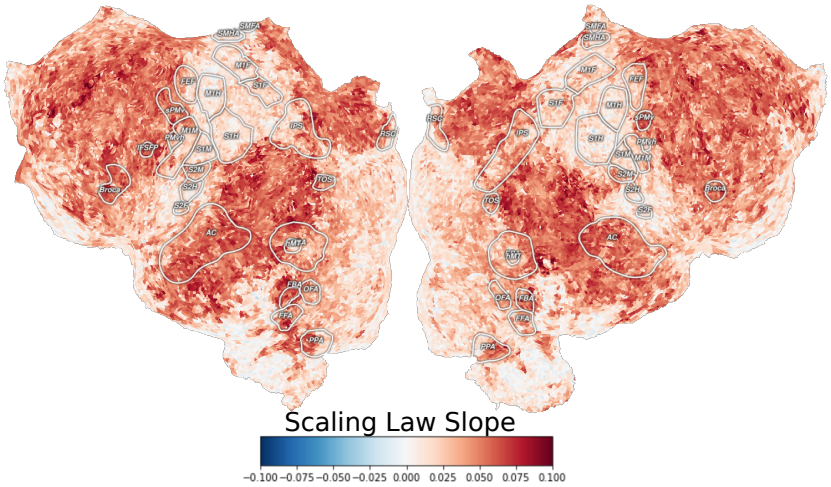
(1) *Joint Data-Parameter Scaling:* Histogram showing the slopes of voxelwise log-linear scaling laws for two model sizes. As model size increases, the marginal benefit of additional data increases.



(2) *Average r^2 Scaling:* Layerwise encoding performance across all models as measured by r^2. Error bars denote SNR-normalized standard error across subjects.



(3) *Voxelwise Model Size Scaling:* The voxelwise slope of the log-linear scaling law for model size. A value of 0.01 would mean that encoding performance (r) increases for that voxel by 0.01 when model size increases by a factor of e.



(4) *Voxelwise Dataset Size Scaling:* The voxelwise slope of the log-linear scaling law for dataset size. A value of 0.1 would mean that encoding performance (r) increases for that voxel by 0.1 when dataset size increases by a factor of e.