Auxiliary Task Ablation: Visual Generalization Benchmark Results

Legend:
- No perturbation Baseline IQM
- Perturbed Condition IQM
- Error-bar = 95% CI (on Perturbed IQM)