Stylized ImageNet (Cue-conflict)

Legend:
- humans vs. humans
- CNNs vs. CNNs
- ViTs vs. ViTs
- AlexNet vs. humans
- VGG−16 vs. humans
- GoogLeNet vs. humans
- ResNet−50 vs. humans
- ViT-B_32 vs. humans
- ViT-L_32 vs. humans
- ResNet-50 (ft.) vs. humans
- ViT-B_32 (ft.) vs. humans

Y-axis: Error consistency (κ)

X-axis: Error overlap expected by chance (c_{exp})