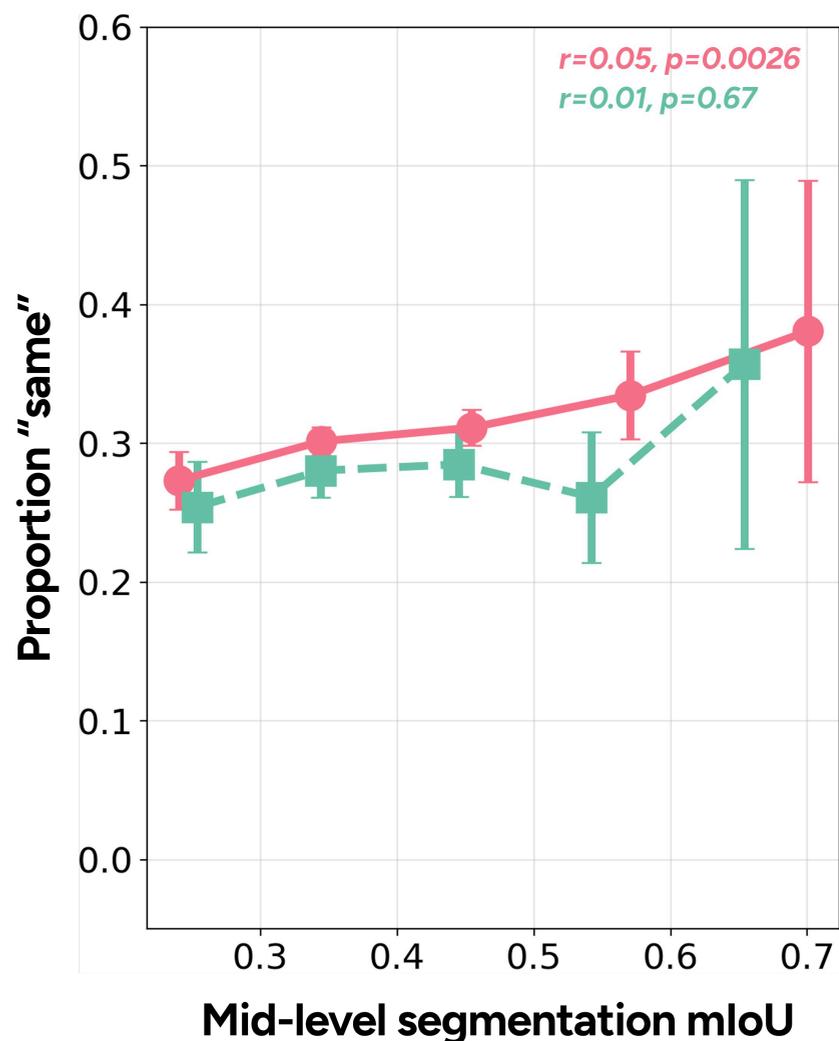
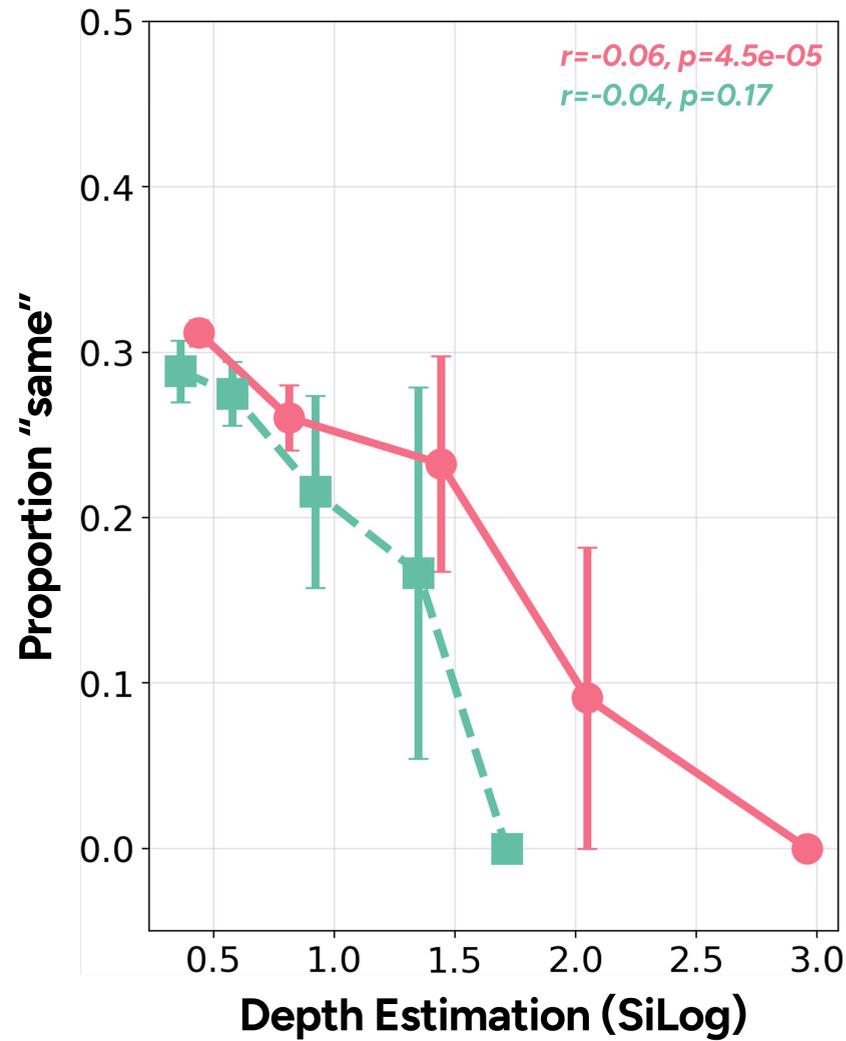


# mid-level depth & proto-object segmentation



# high-level visuo-linguistic alignment

