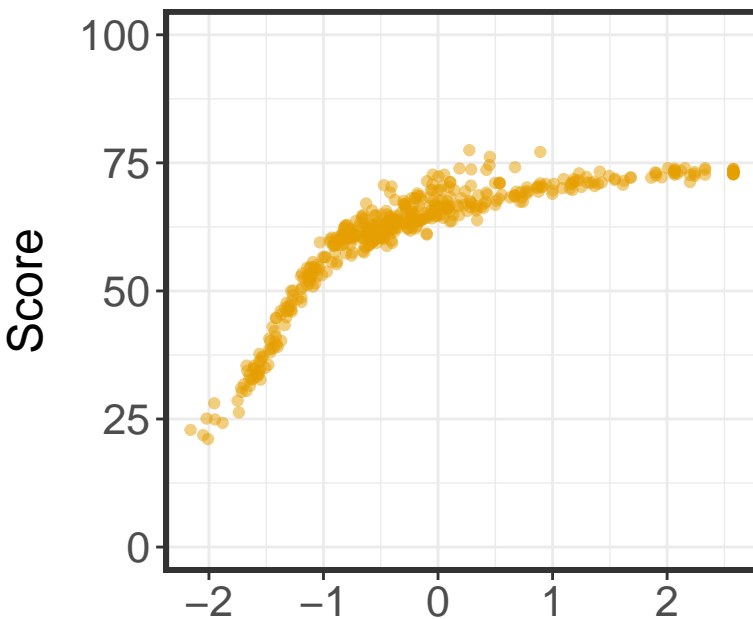
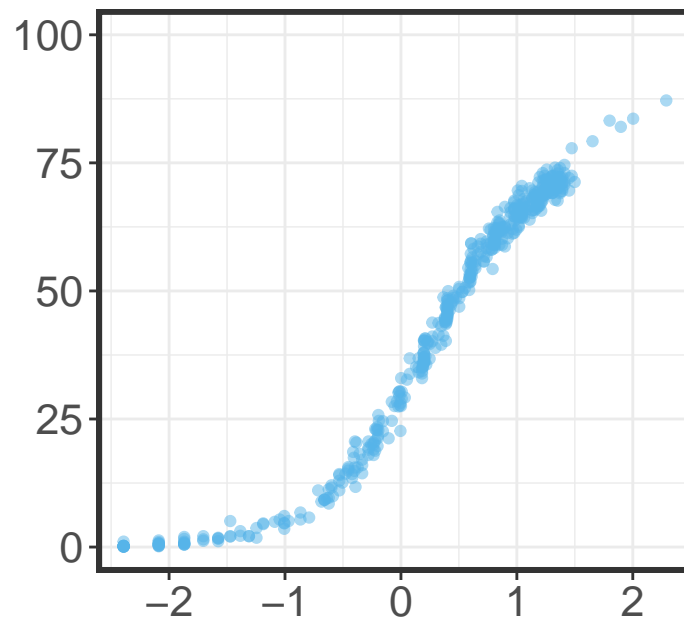


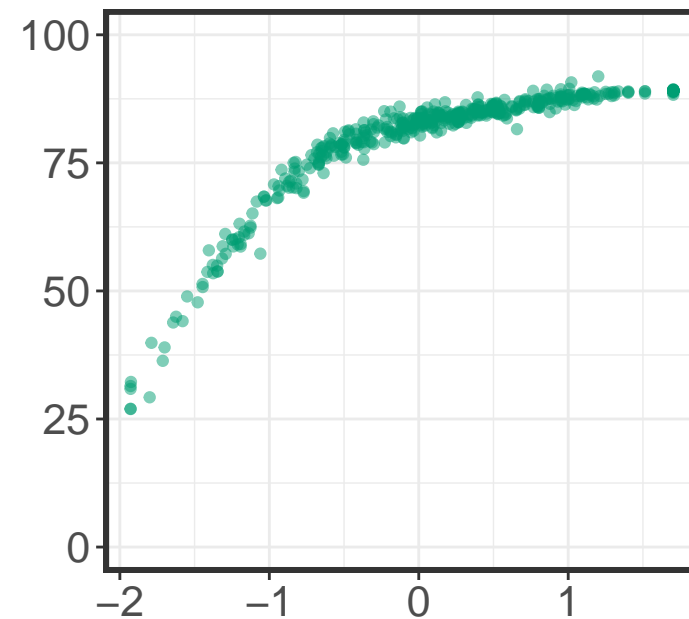
ARC* (d = 100)



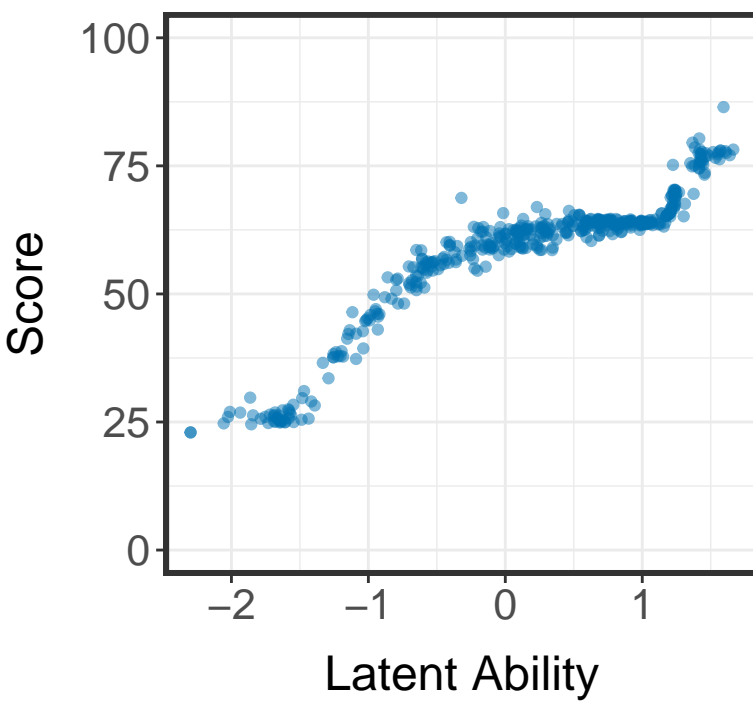
GSM8K* (d = 237)



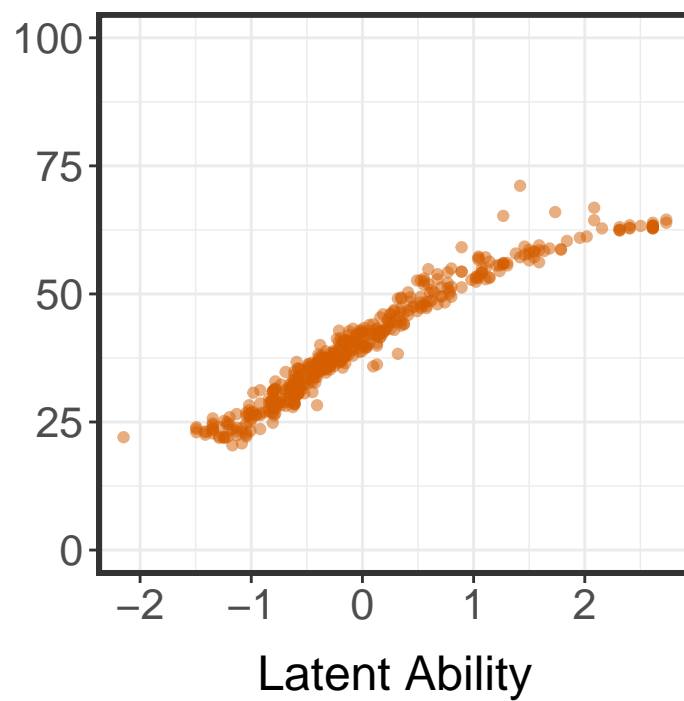
HellaSwag* (d = 58)



MMLU* (d = 102)



TruthfulQA* (d = 136)



WinoGrande* (d = 106)

