XQuAD on Gemini - Mean LLM Judge Score per Language 6 4.61 ± 1.14 4.67 ± 1.02 4.73 ± 0.95 4.74 ± 0.90 4.68 ± 0.92 4.78 ± 0.83  $4.94 \pm 0.28$ 5 FLM Judge Score zh th el ar tr en ru