

Mean κ_w for all Evaluators across Language Tiers

