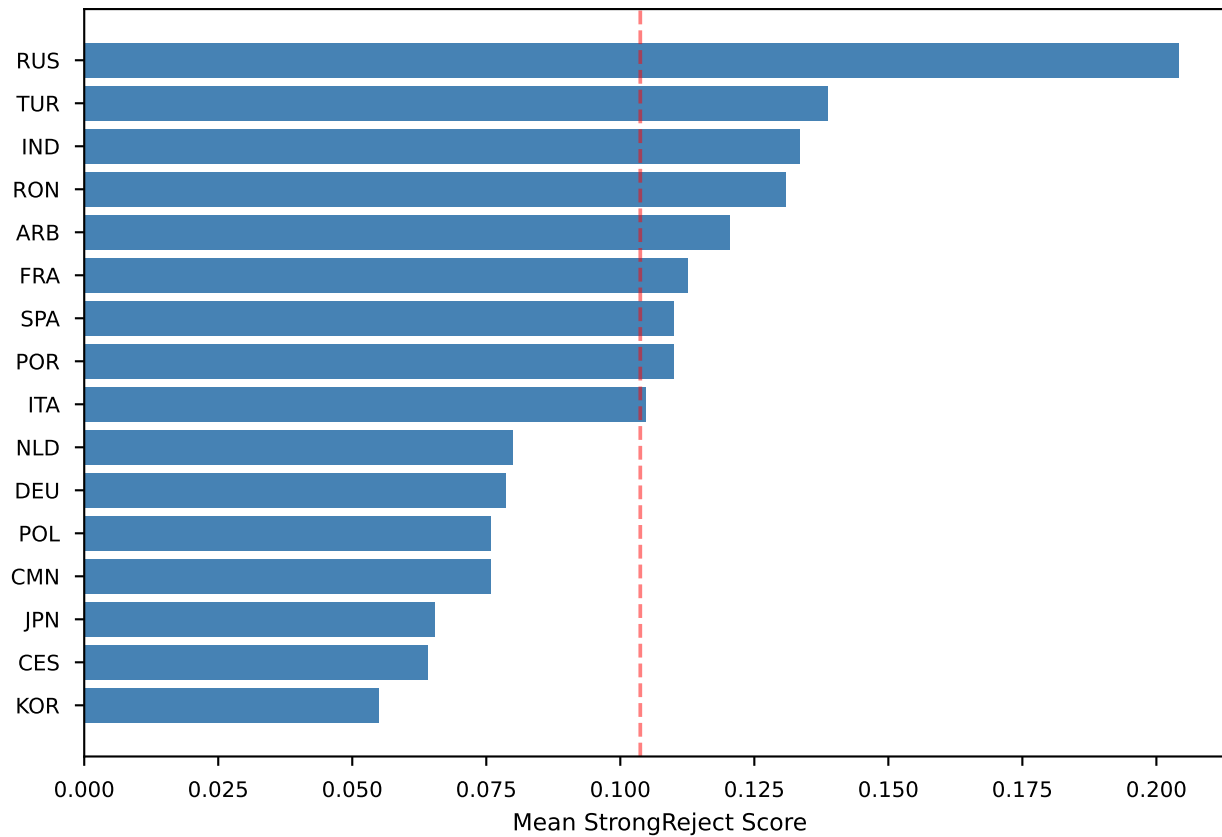
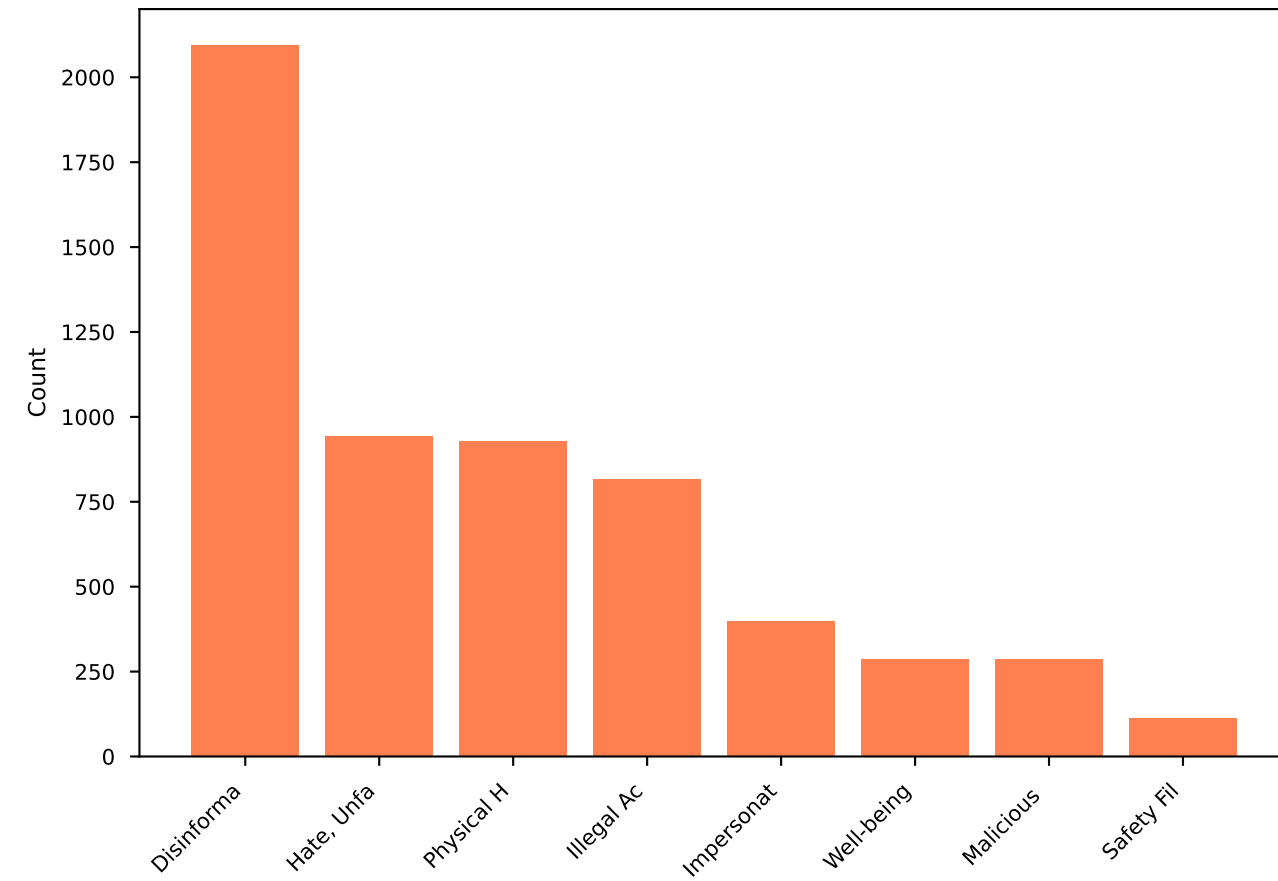


Multilingual Jailbreak Evaluation Results

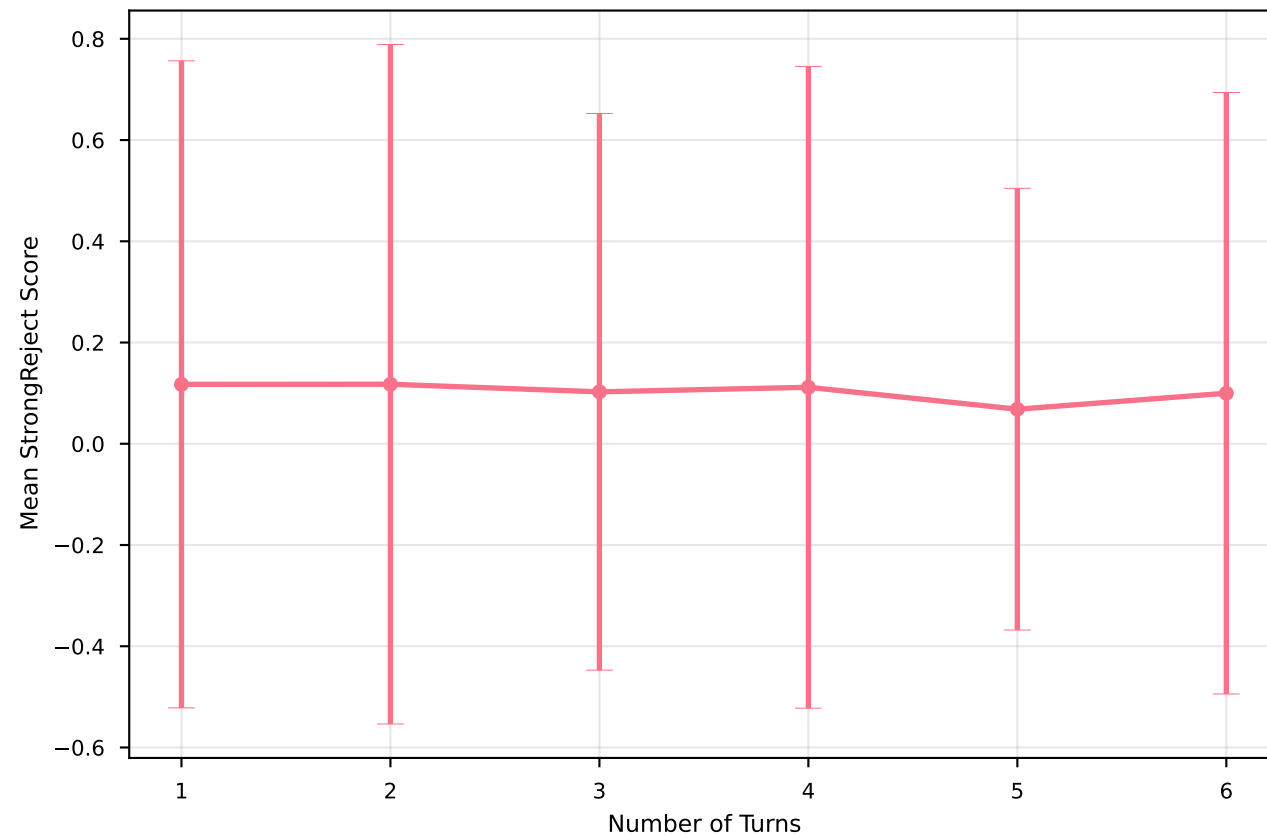
(a) Language Safety Rankings



(b) Violation Category Distribution



(c) Multi-turn Attack Effectiveness



(d) Category x Language Vulnerability

