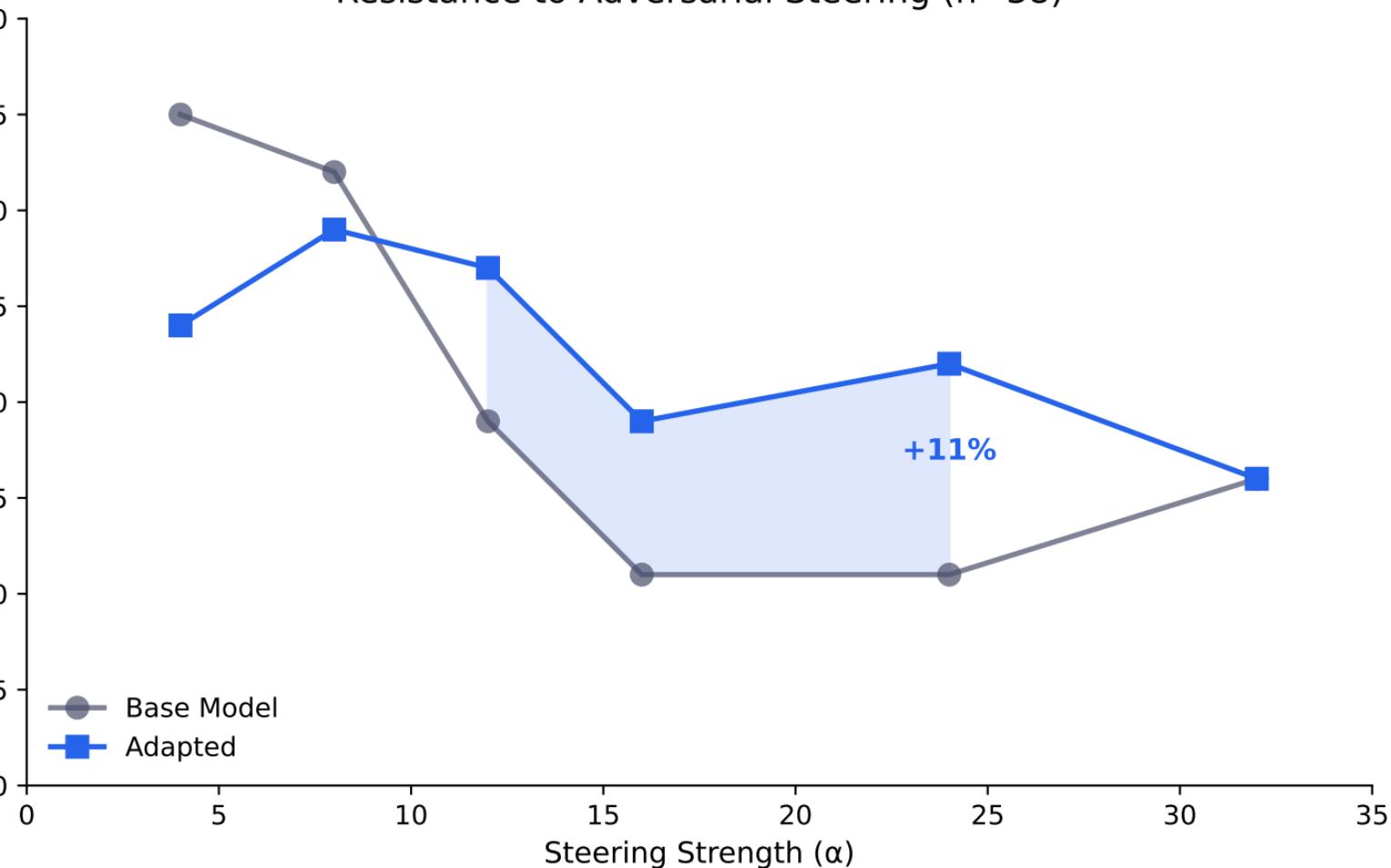


Resistance to Adversarial Steering (n=38)

Correct Answers (%)



Base Model
Adapted