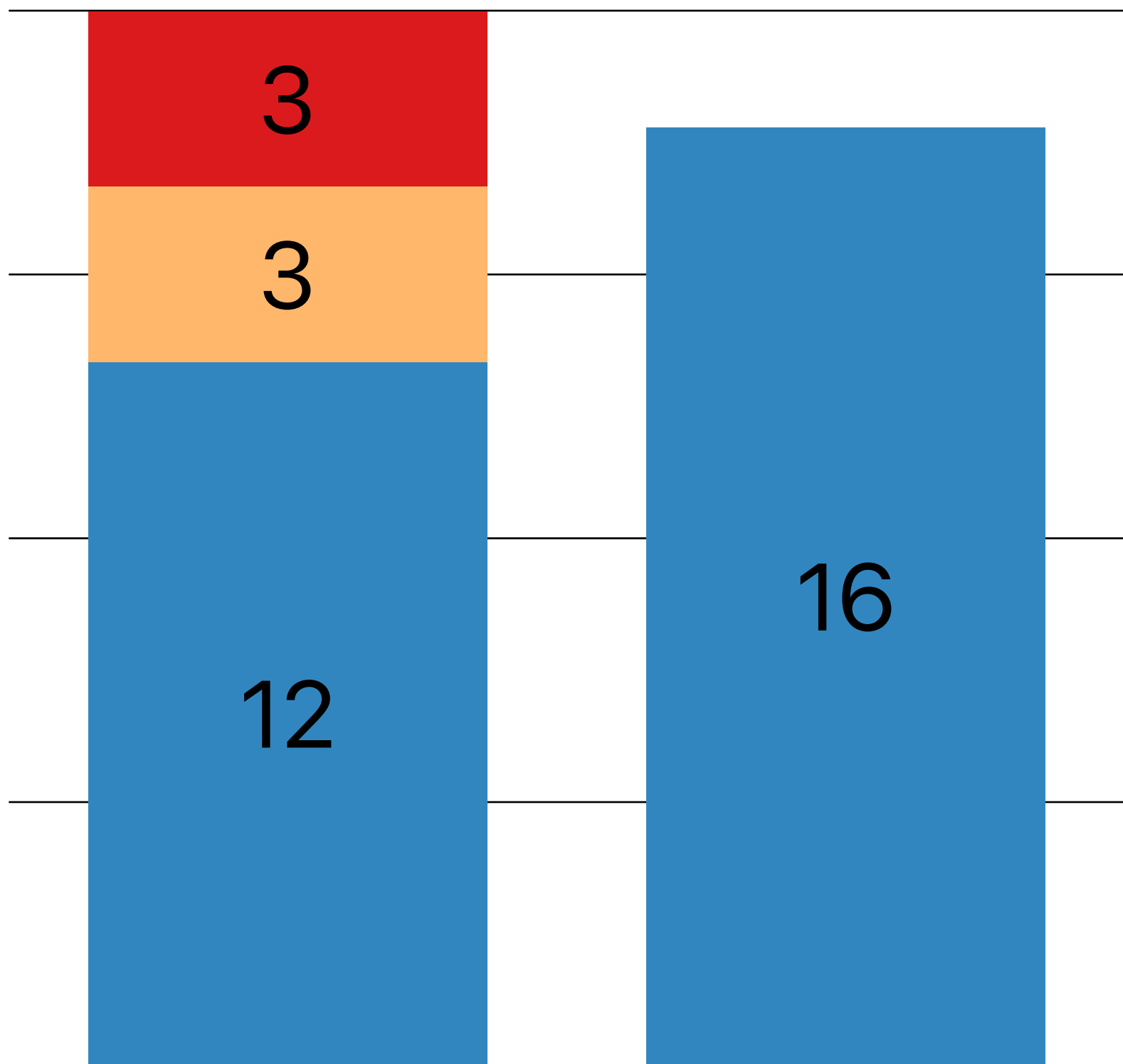


Results: Effects on Over-/Under-Reliance



No-LP

LP

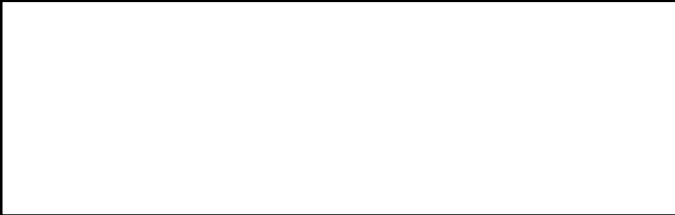
under-trusted the AI, leading
to validation timeout (under-
reliance)

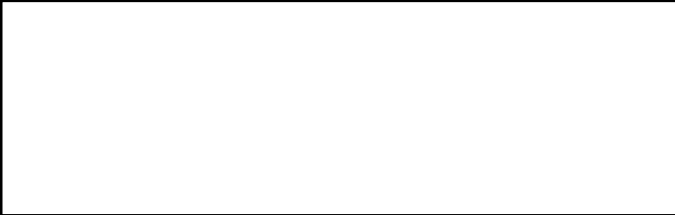
blindly trusted the AI
suggestion (over-reliance)

properreliance on AI

Validation Suites in Fixed-Prompt Tasks







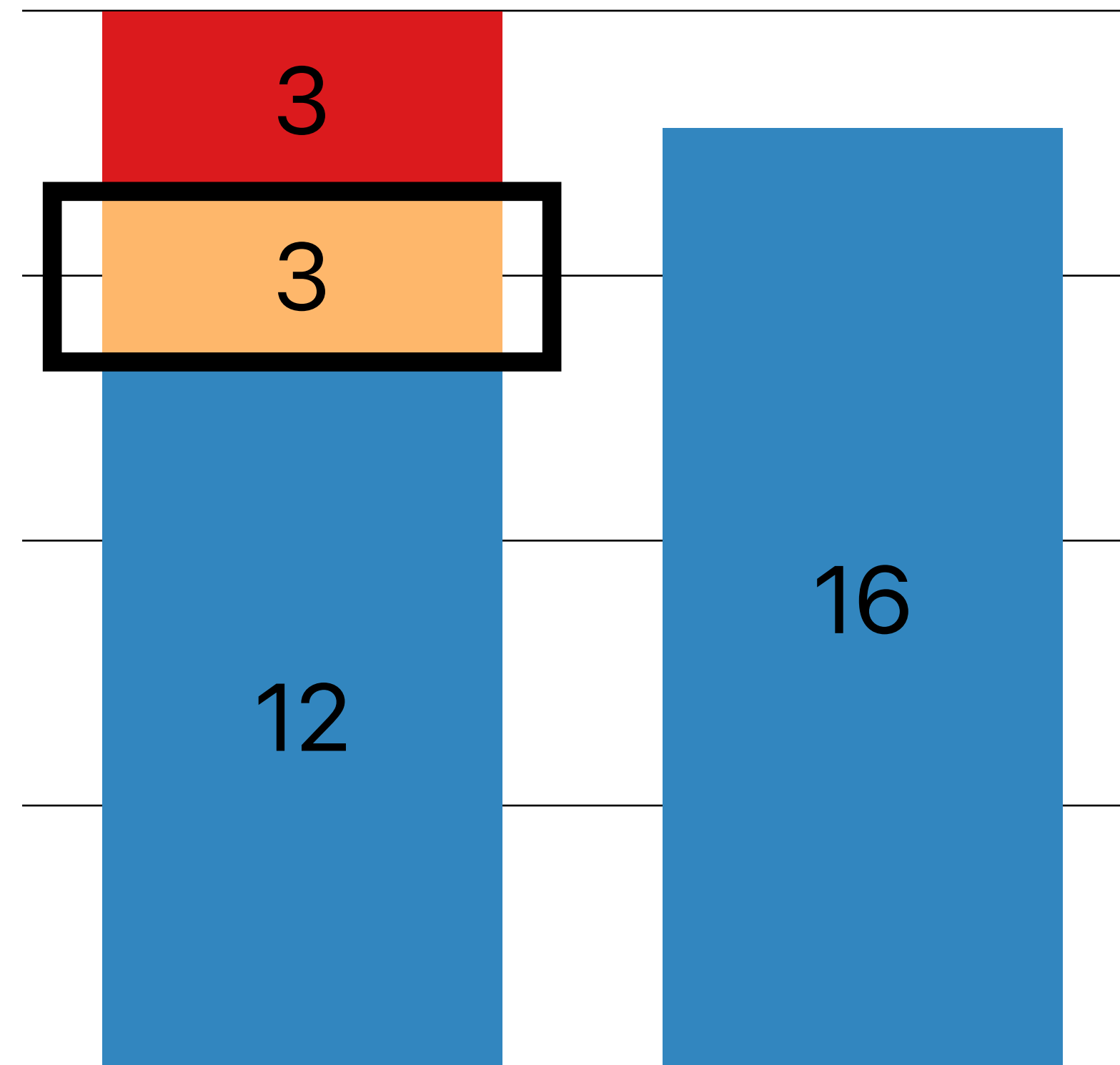
2

6

Results: Effects on Over-/Under-Reliance

blindly trusted the AI
suggestion (**over-reliance**)

under-trusted the AI, leading
to validation timeout (**under-
reliance**)



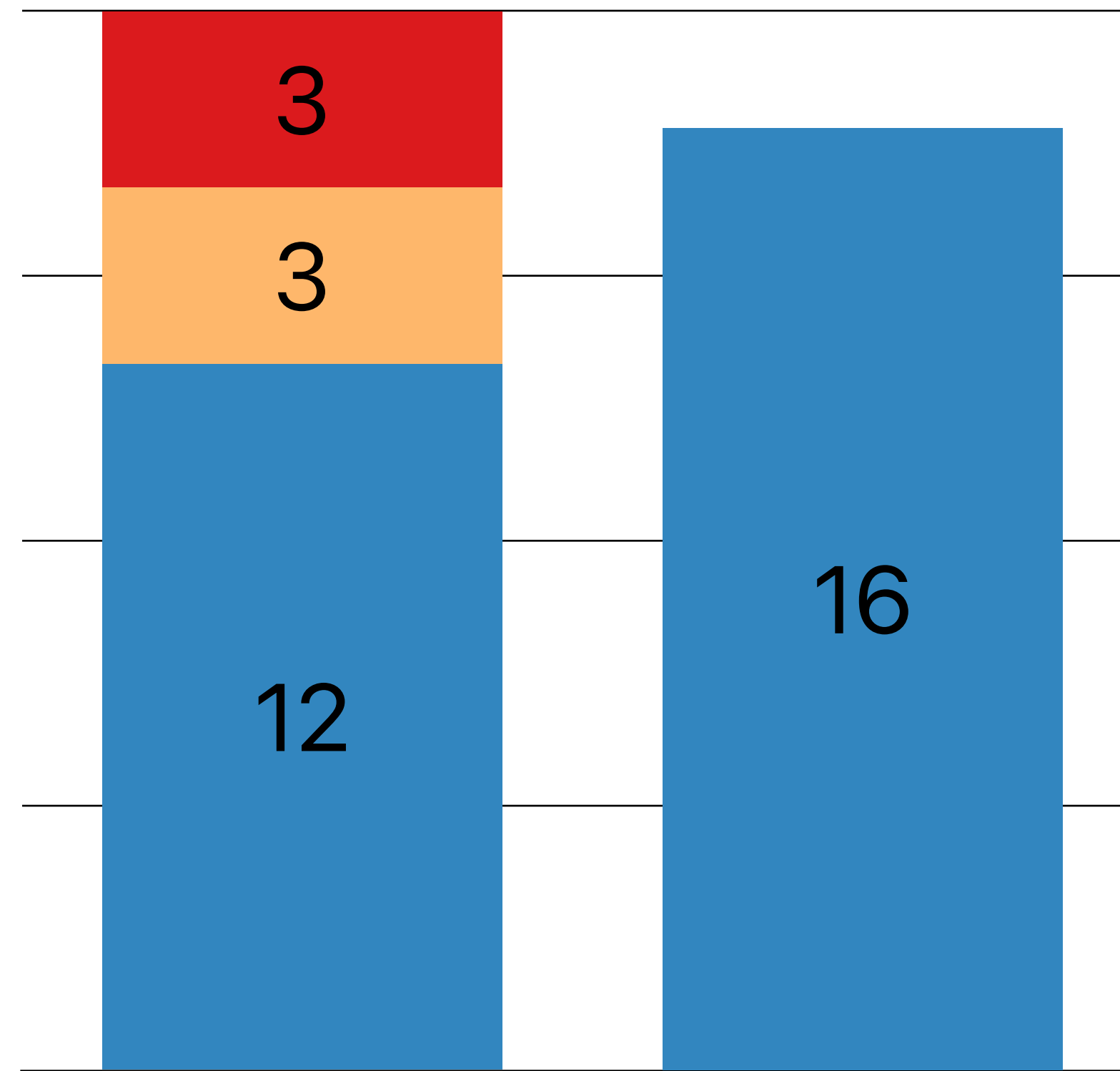
proper reliance on AI

No-LP

LP

Validation Success in Fixed-Prompt Tasks

Results: Effects on Over-/Under-Reliance



No-LP

LP

Validation Success in Fixed-Prompt Tasks