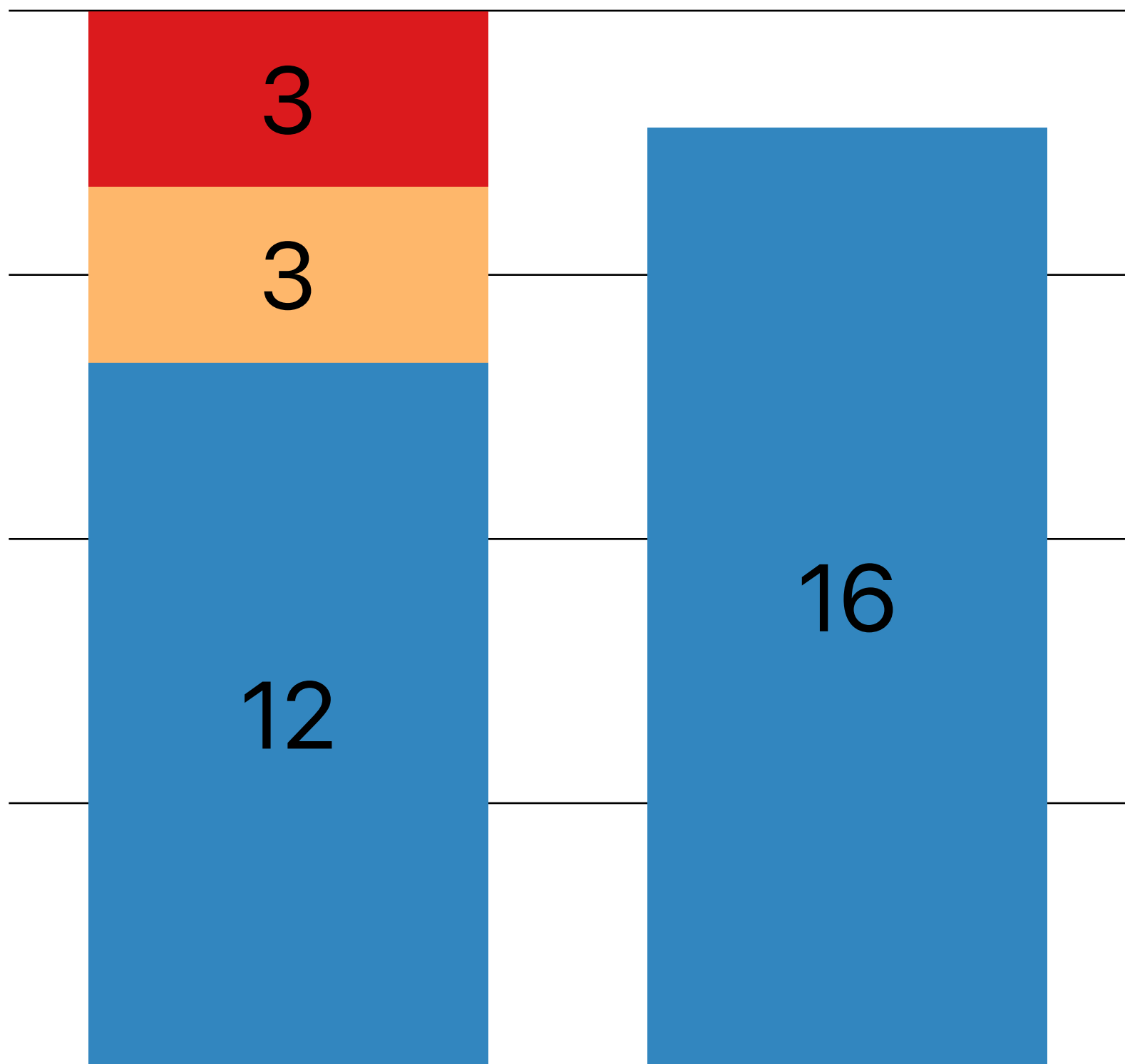




**Results: Effects on Over-/Under-Reliance**



**No-LP**

**LP**

Example of under-reliance:  
Not accepting correct AI  
suggestions but writing their  
own code

Example of over-reliance:  
Trusting AI-generated  
comments that incorrectly  
describe the code

"[LP] saved me the effort of writing multiple print statements."

# Validation Suites in Fixed-Prompt Tasks

**By lowering the cost of validation by execution, LP mitigates  
over-/under-reliance on AI-generated code**



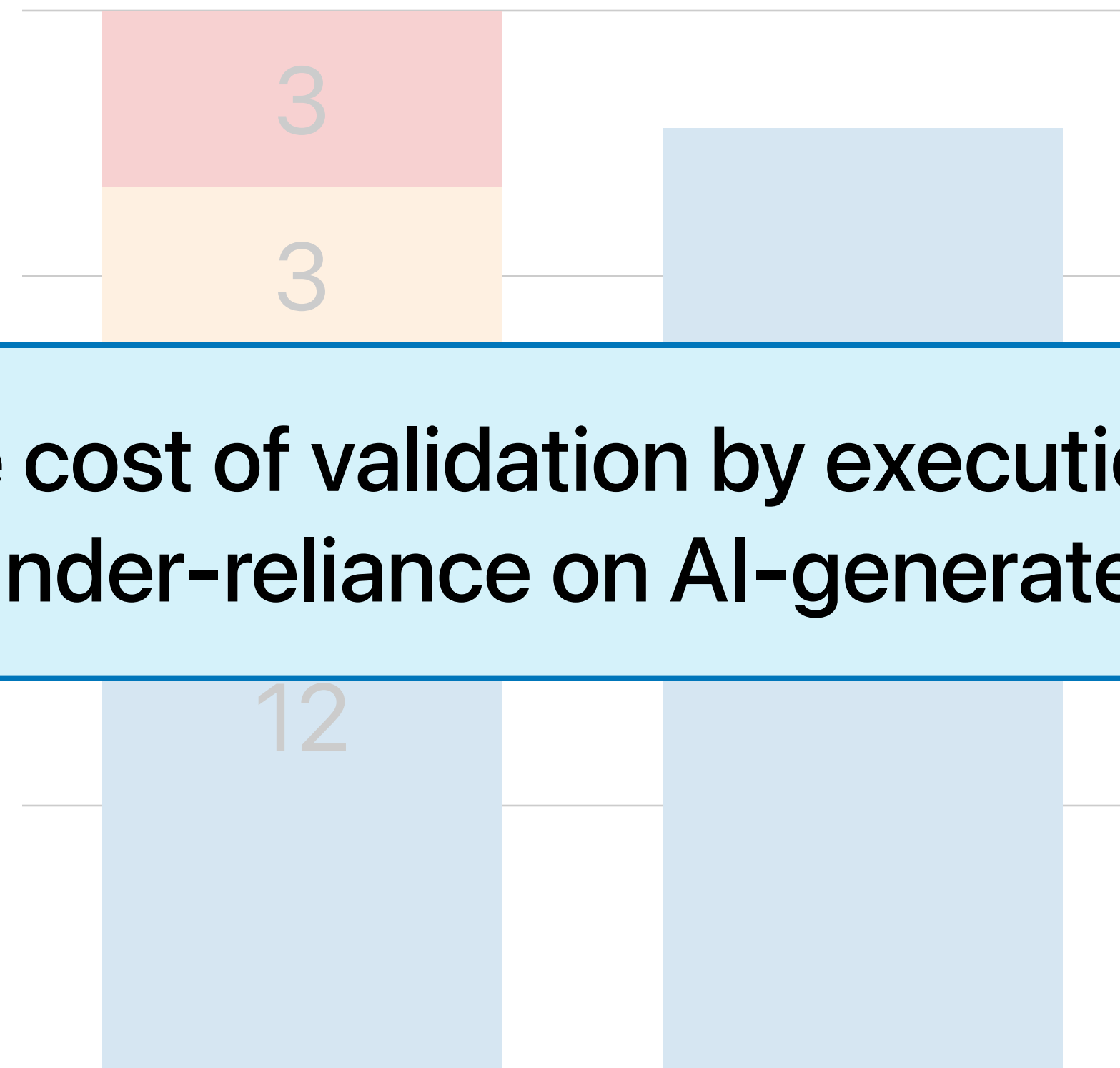


# Results: Effects on Over-/Under-Reliance

Example of over-reliance:  
Trusting AI-generated  
comments that incorrectly  
describe the code

Example of  
Not accepting  
suggestion  
own code

By lowering the cost of validation by execution, LP mitigates  
over-/under-reliance on AI-generated code



No-LP

LP

Validation Success in Fixed-Prompt Tasks

print statements."

the  
multiple

# Results: Effects on Validation Strategies

