

Formal "Aha!" ratio sweep (problem-level)  
MATH-500, Qwen2.5-1.5B

