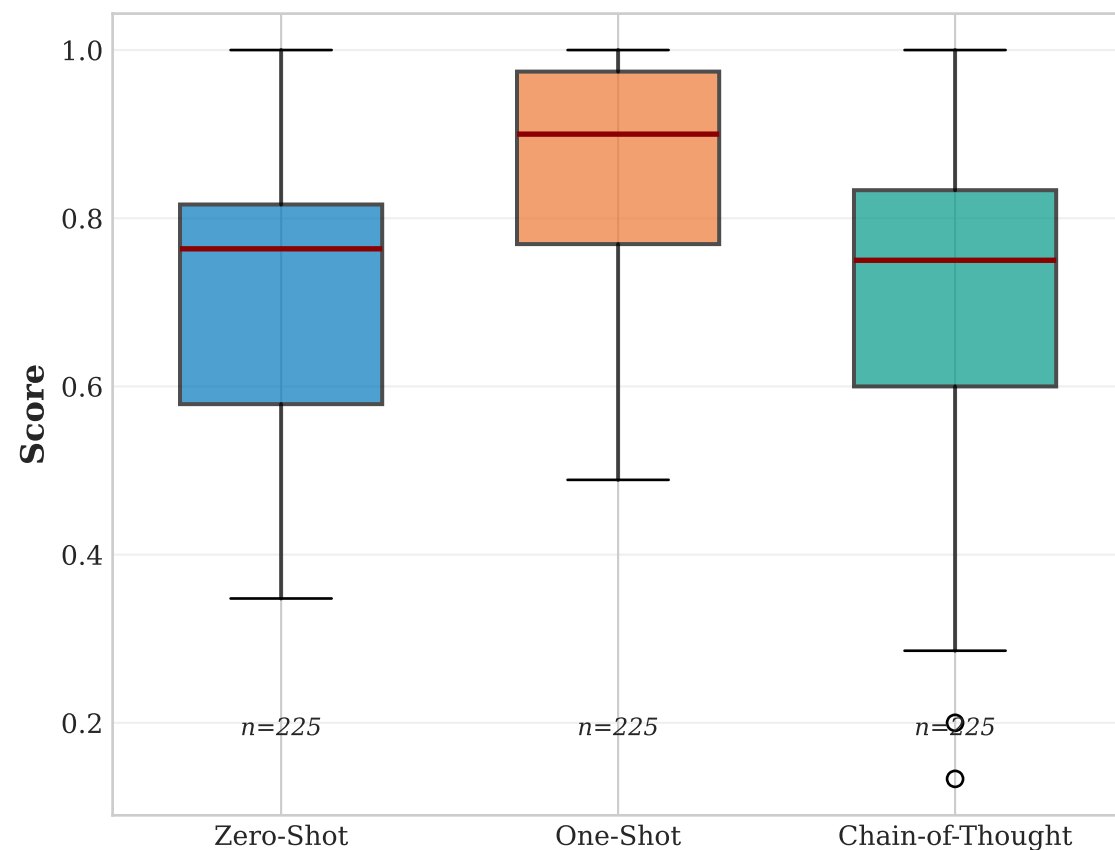
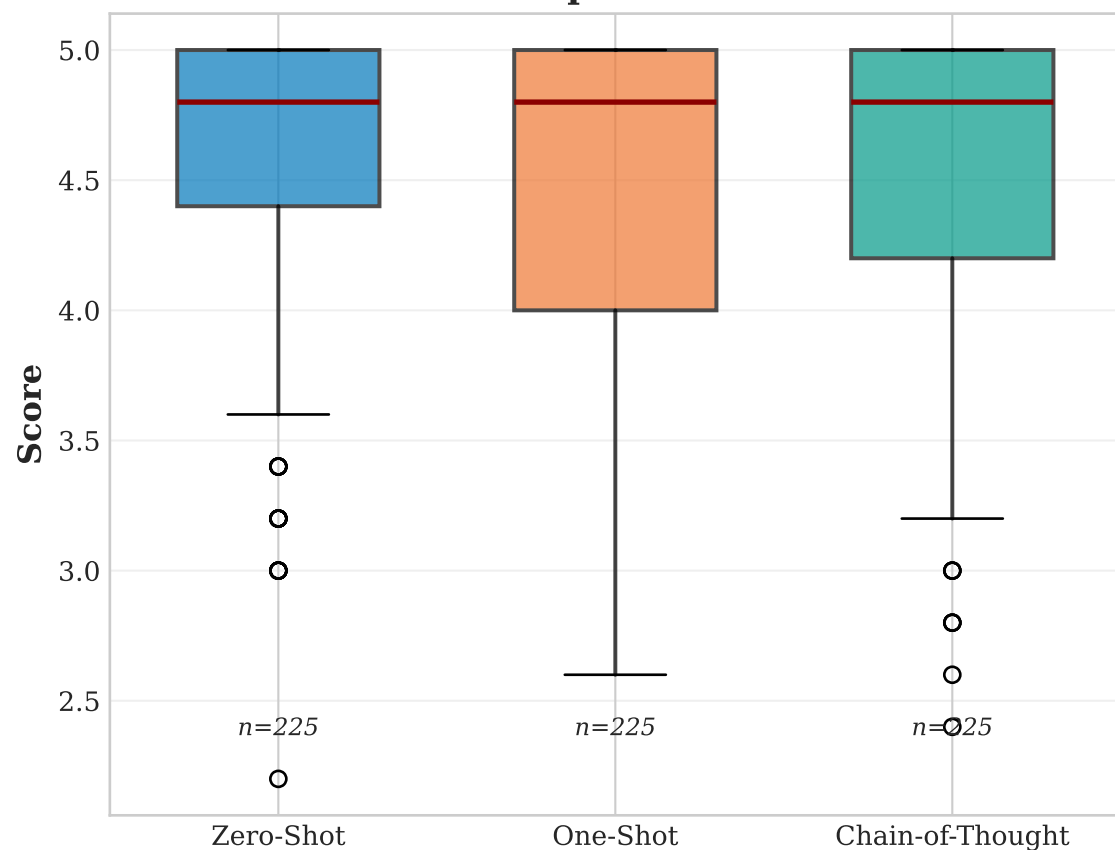


Score Distributions by Prompting Strategy (H2)

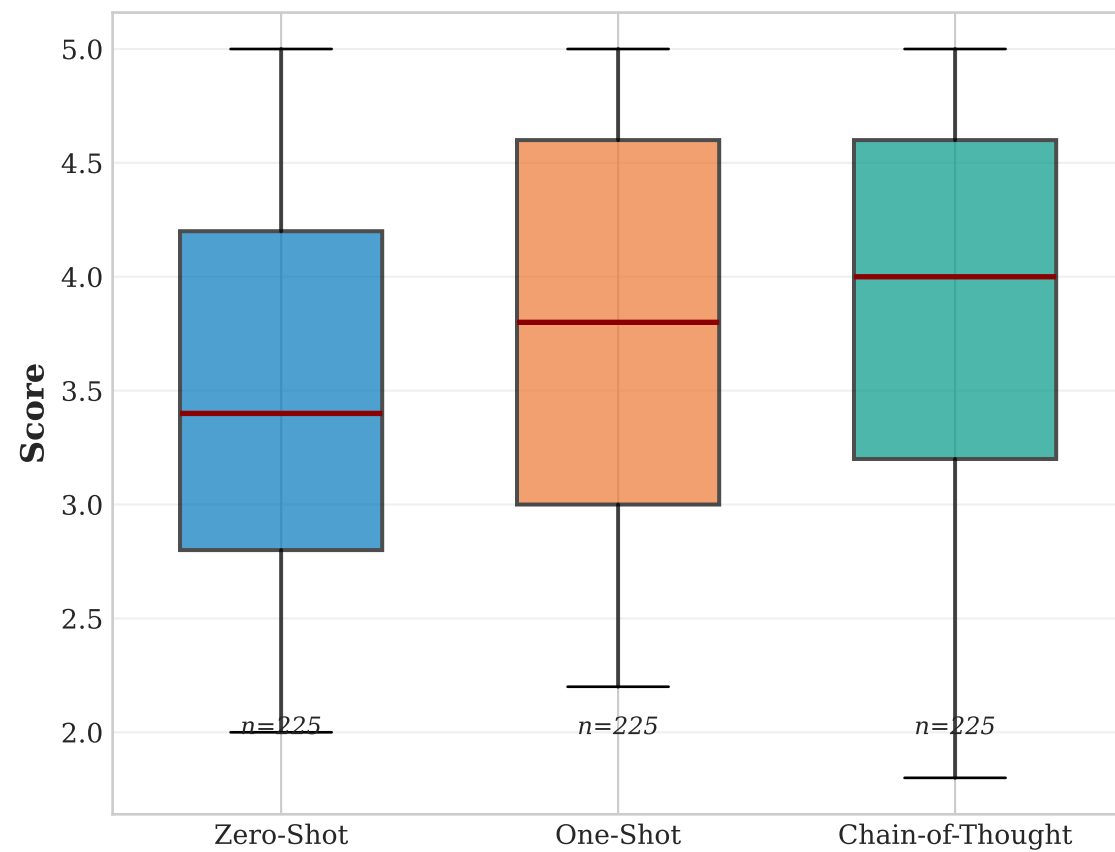
F1 Score



Completeness



Correctness



Hallucination

