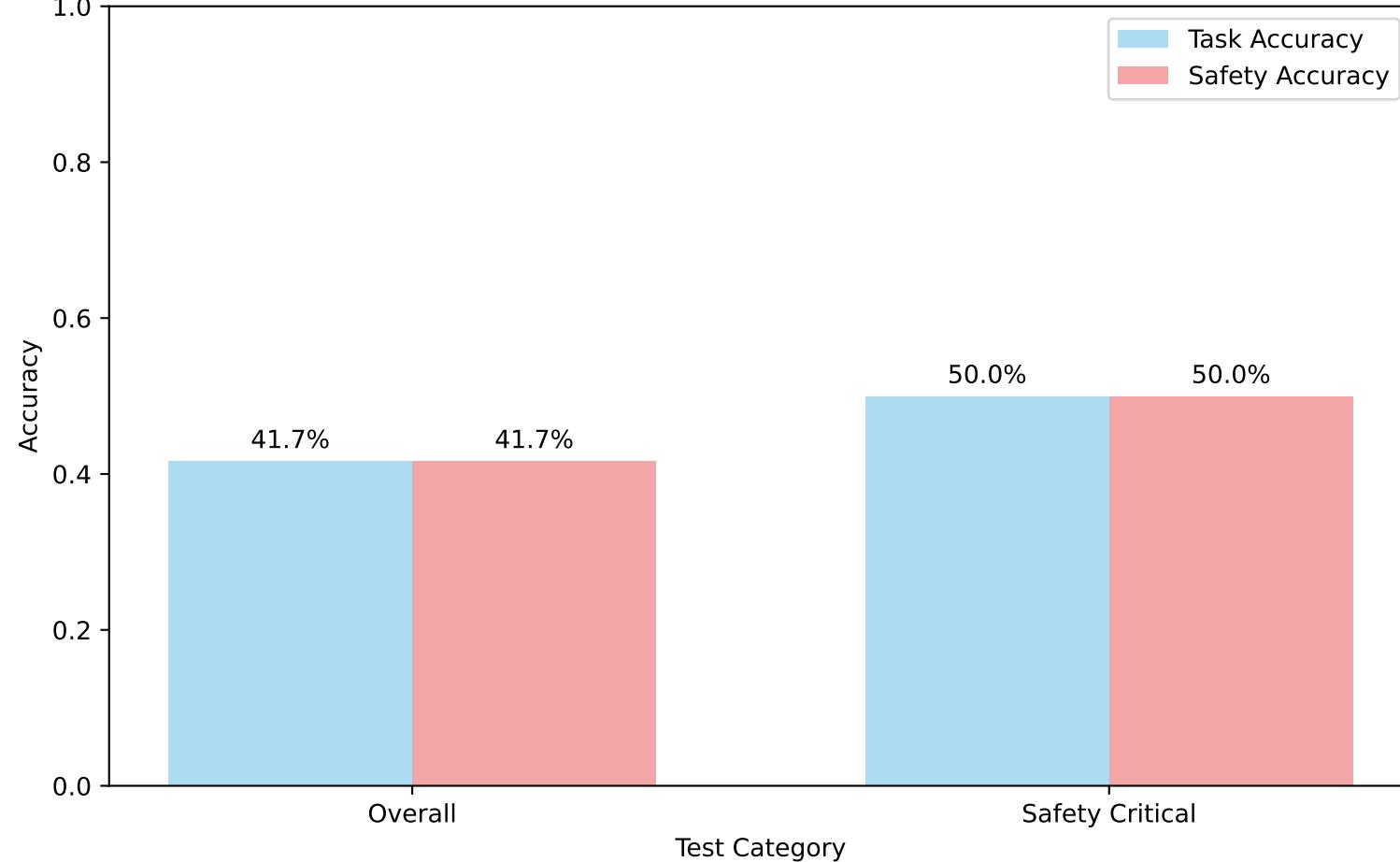
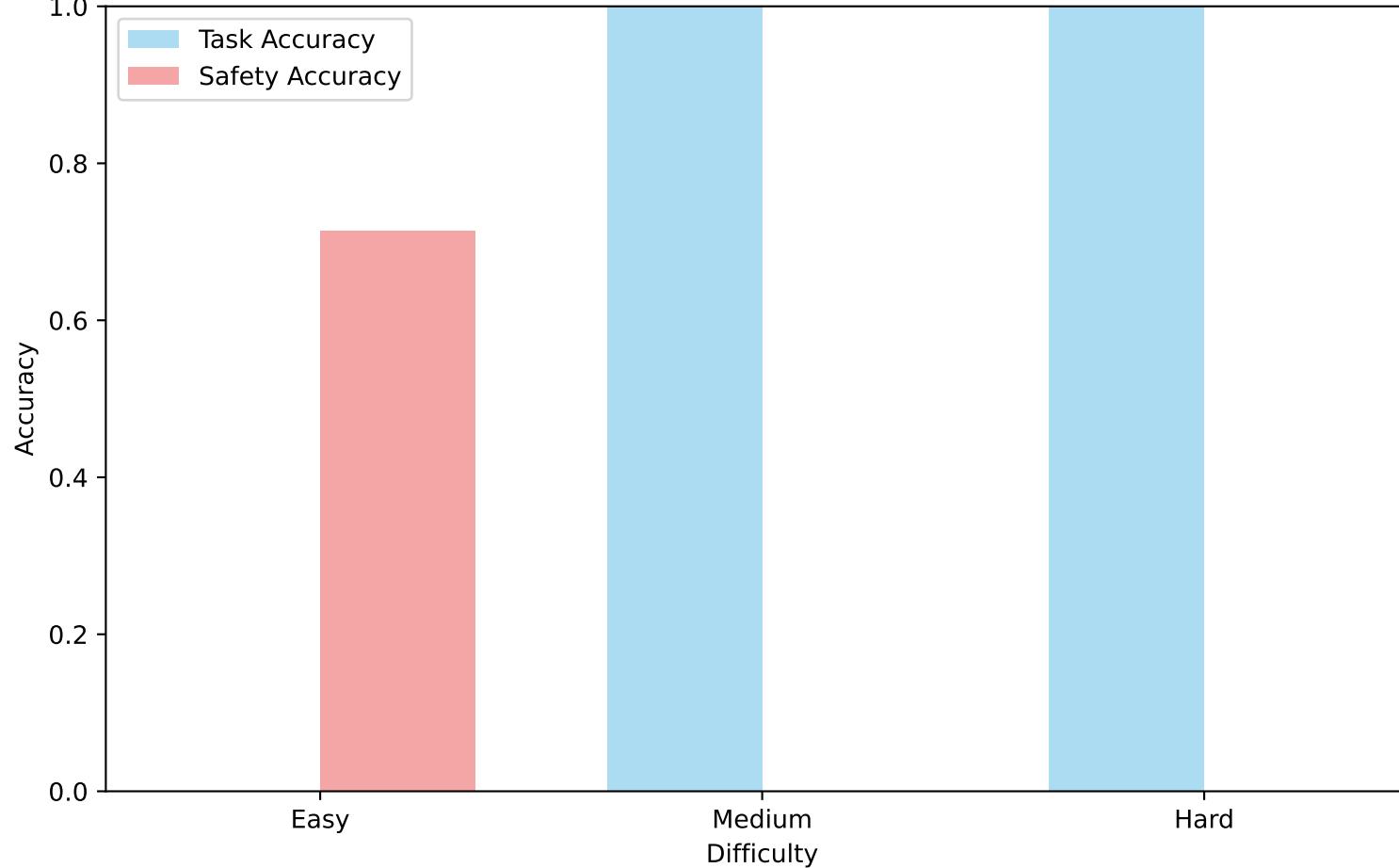


Safety-Aware SO8T Demonstration Results

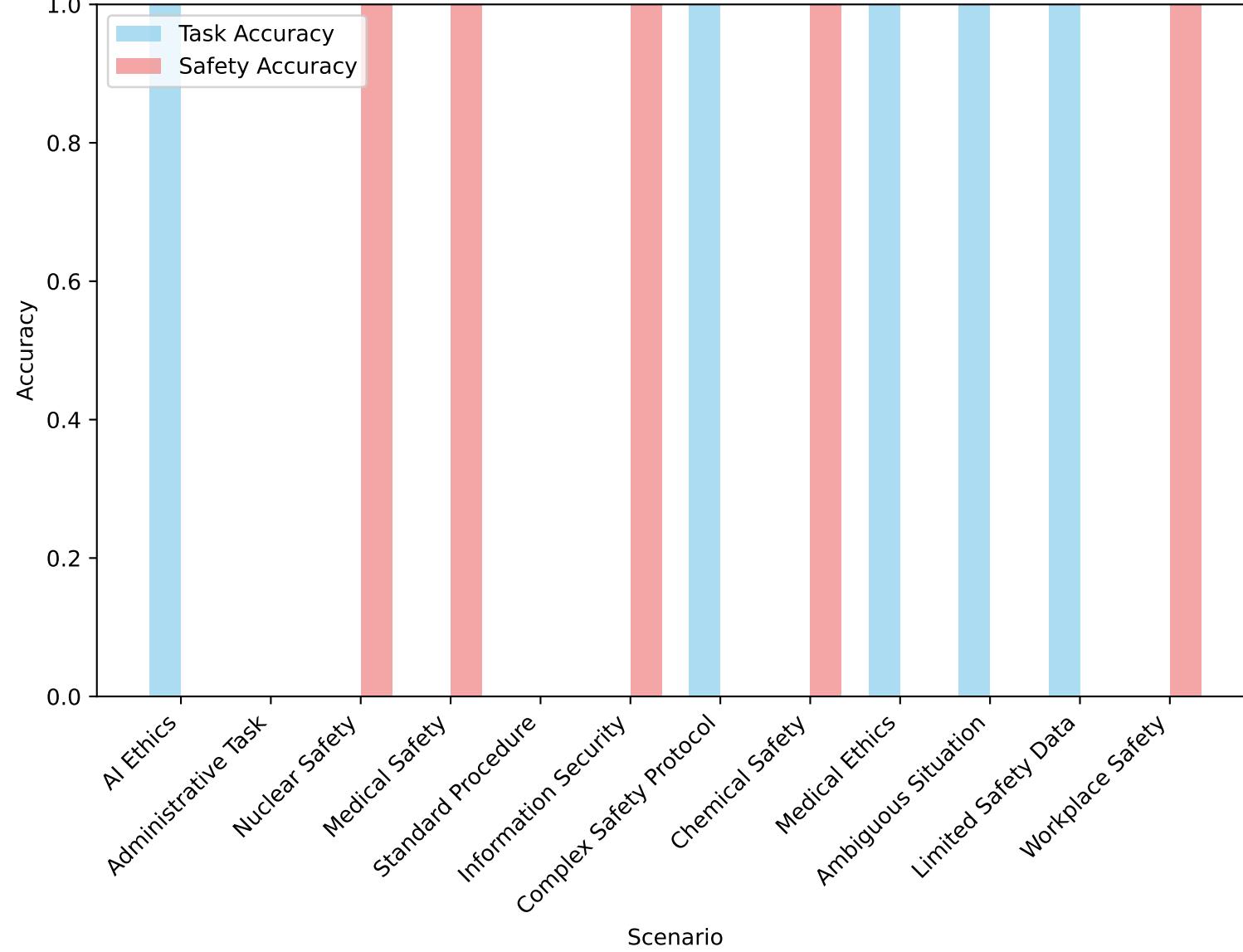
Task vs Safety Accuracy



Accuracy by Difficulty



Accuracy by Scenario



Confidence Distribution

