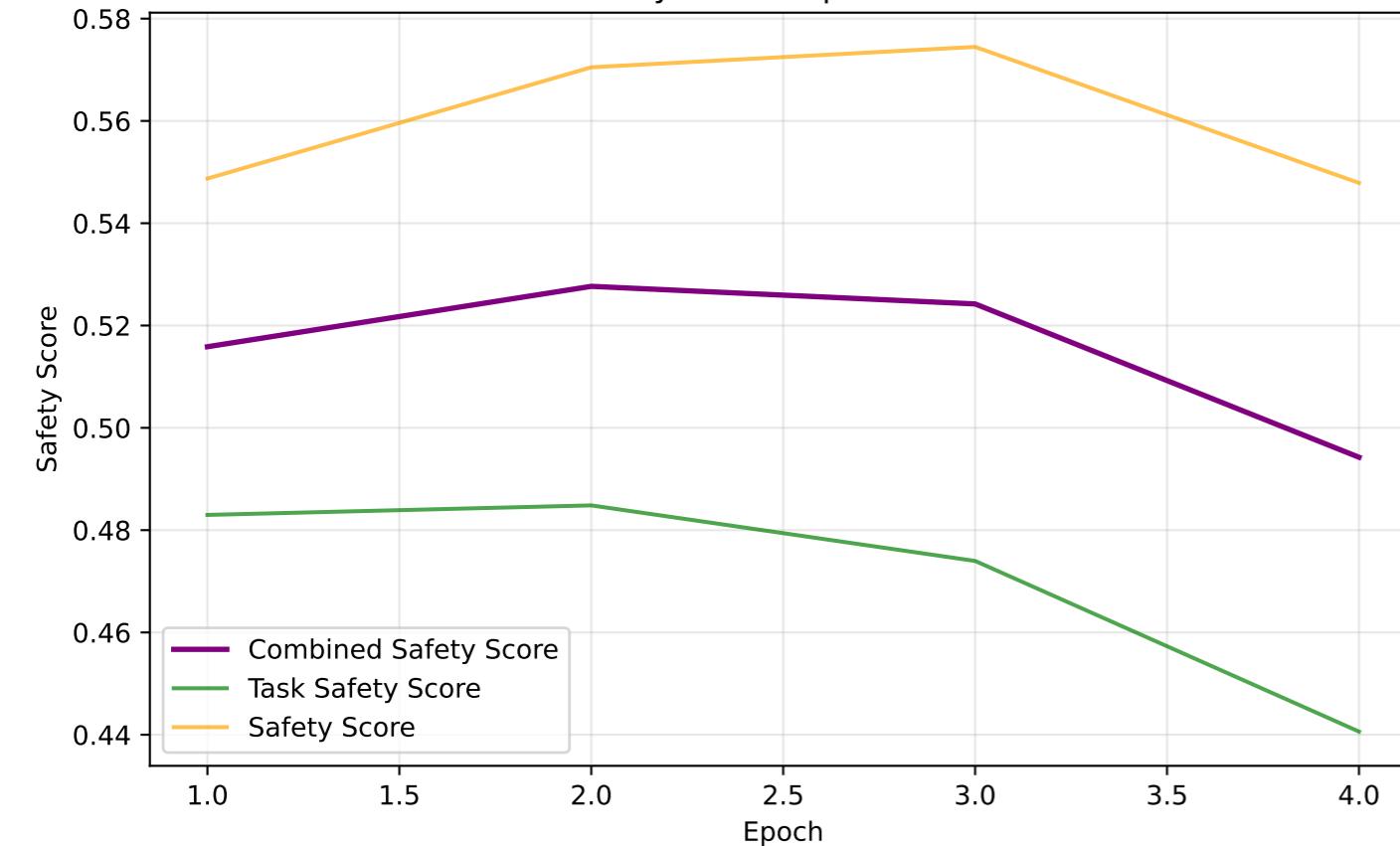
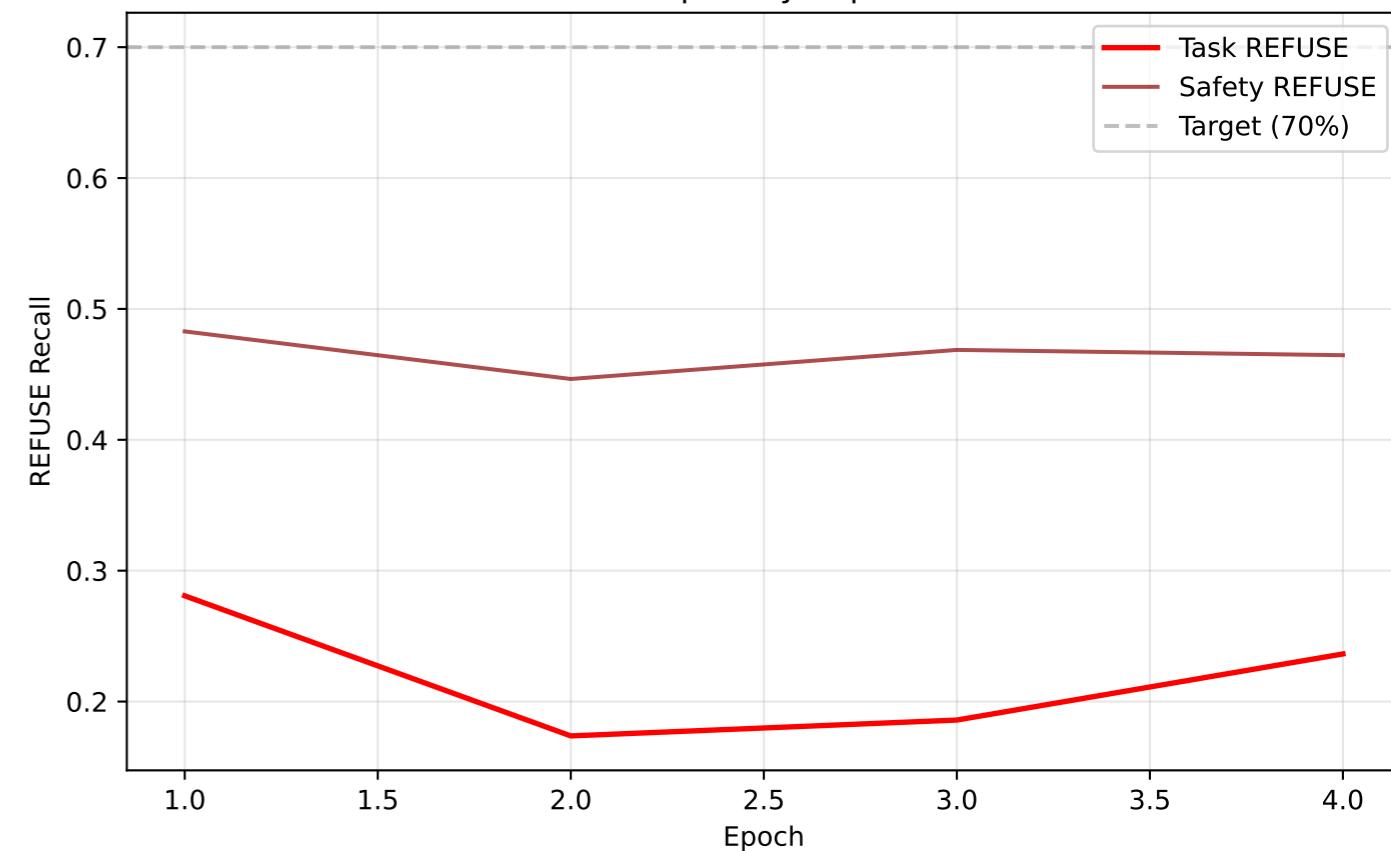


Safety Improvement Over Time

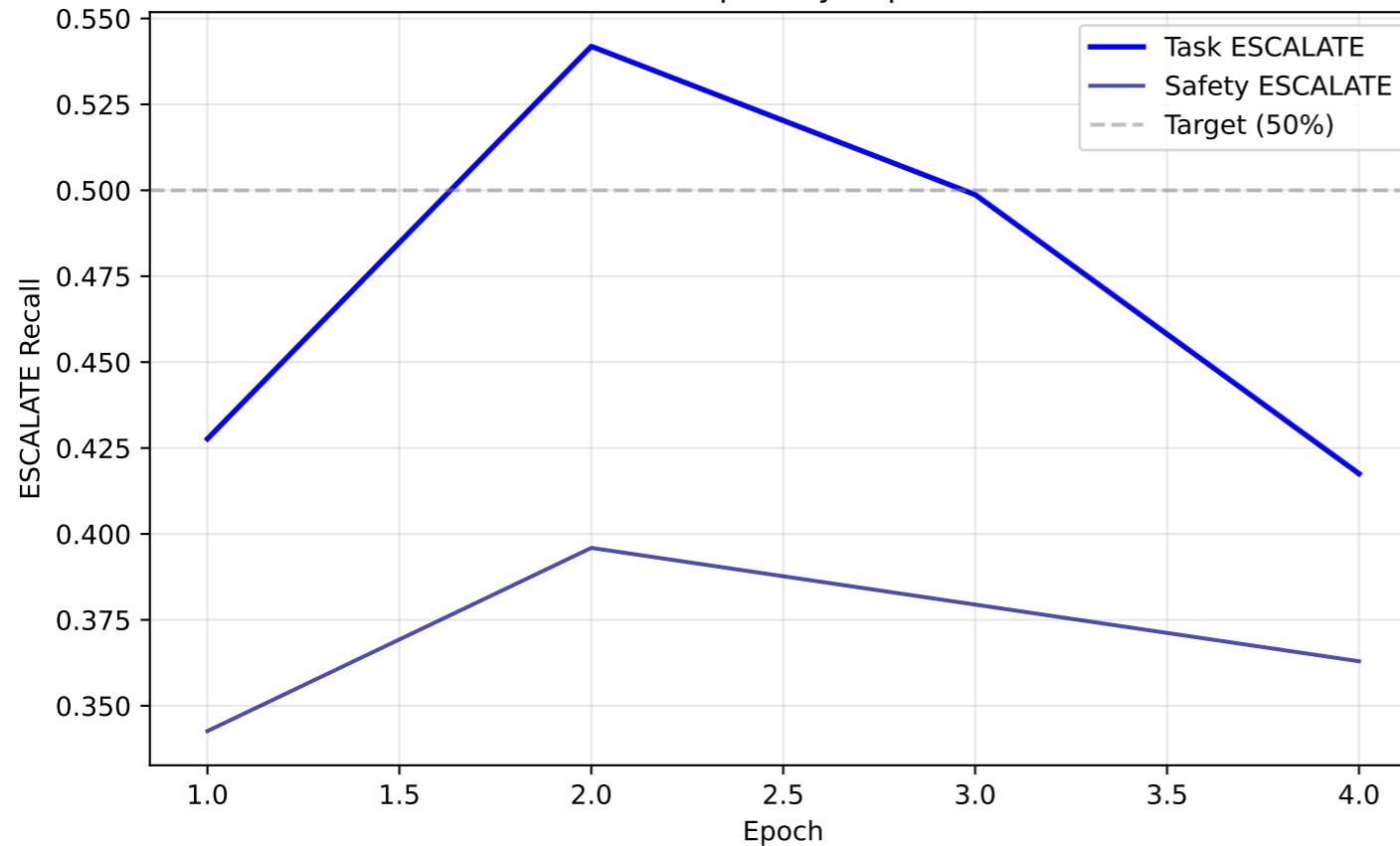
Safety Score Improvement



REFUSE Capability Improvement



ESCALATE Capability Improvement



Overcompliance Suppression

