Average Escalation Score Over Time (Cyberattack Scenario) Claude-2.0 \_lama-2-Chat Escalation Score ← Time t [Days]