GPT-3.5 on PRISM ACC: 0.656 High Certainty ACC: 0.673 Low Certainty ACC: 0.599 0.68 **Correct Answer** 600 Wrong Answer 500 1400 300 200 0.64 0.62 100 0.52 0.70 0.40 70 40 50 60 80 90 Certainty