Cumulative Repair Percentages for Different Feedback Loop Counts across Models 98:28 Models Queried: 100 96:83% 94.67 GPT-3.5 Turbo 92.17% 89.56% 88.60% GPT-40 mini 86.13% 82.78% 82.49% GPT-40 79.64% 77.46% 80 76.99% ol-mini **7**0.02% 65.31% 64.52% 60 · 56.17% 49.89% 39.49% 40 -20 10

Feedback Loop Count

Cumulative Repair Percentage