**MDP-Guided Shows Consistent Advantage**

Legend: MDP-Guided, Always-Reason

Single-hop (n=7): MDP-Guided 85.7%, Always-Reason 71.4%, +14.3%

Multi-hop (n=13): MDP-Guided 69.2%, Always-Reason 53.8%, +15.4%

Observation:
- Single-hop: +14.3%
- Multi-hop: +15.4%

Accuracy (%)

Question Type