Average Target Accuracy (With Answer)

- Base Models: 0.318
- GRPO Models: 0.369
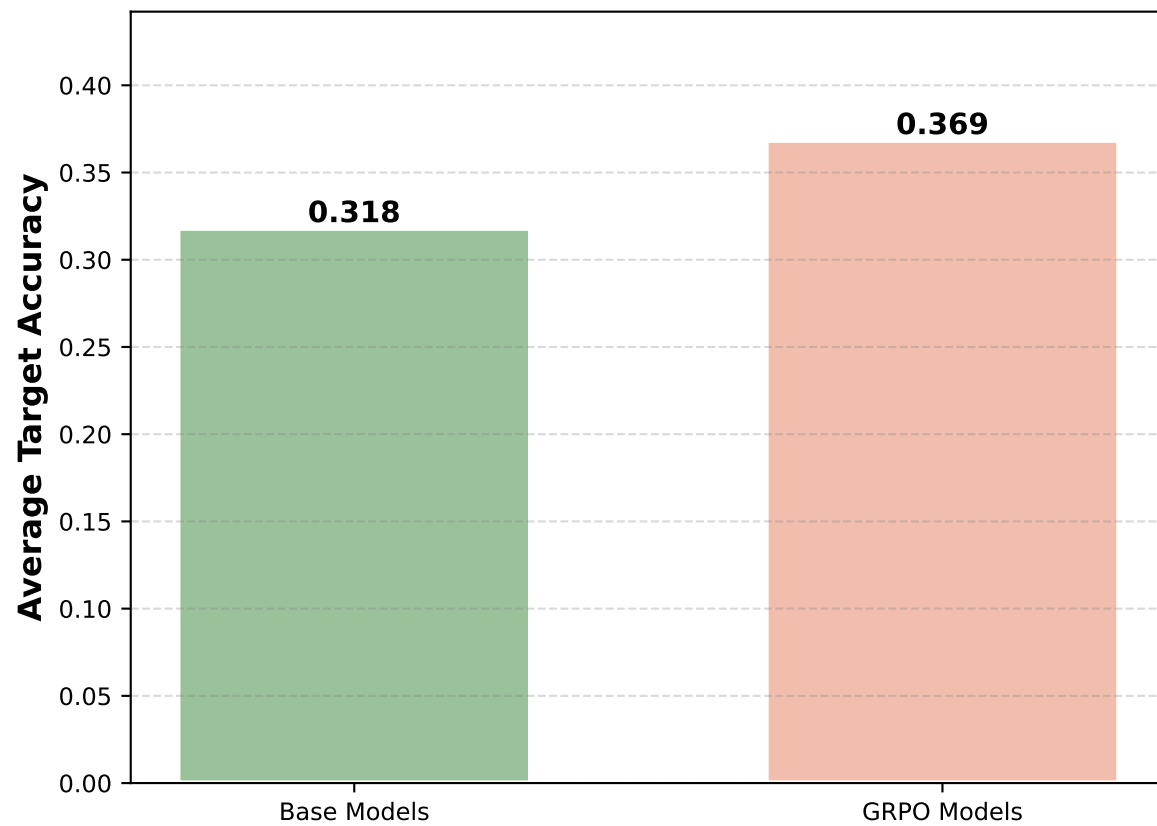
Average Transfer Consistency (With Answer)

- Base Models: 0.226
- GRPO Models: 0.438
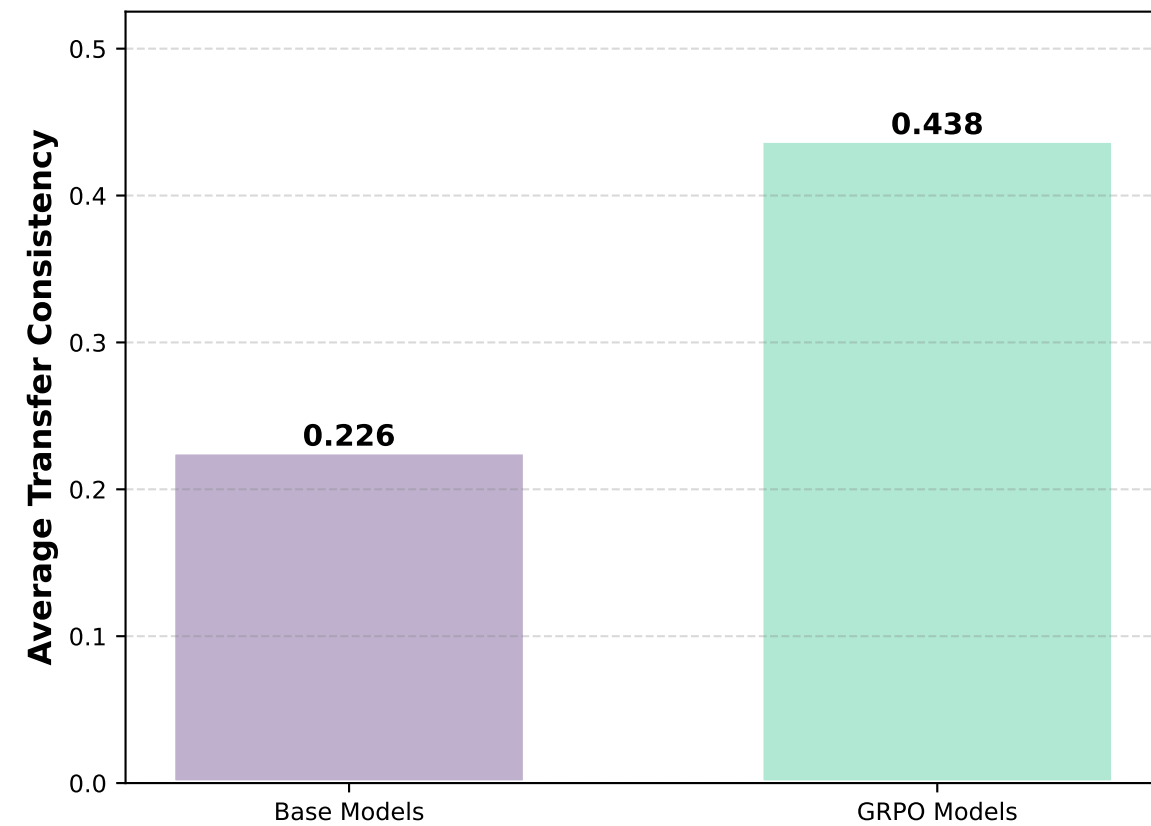
Target Accuracy by Source Model (With Answer)

- Qwen-GRPO: 0.340
- Llama-GRPO: 0.476
- DeepSeek-GRPO: 0.289
- Llama-3.2-3B: 0.260
- DeepSeek: 0.354
- Qwen-2.5-1.5B: 0.340
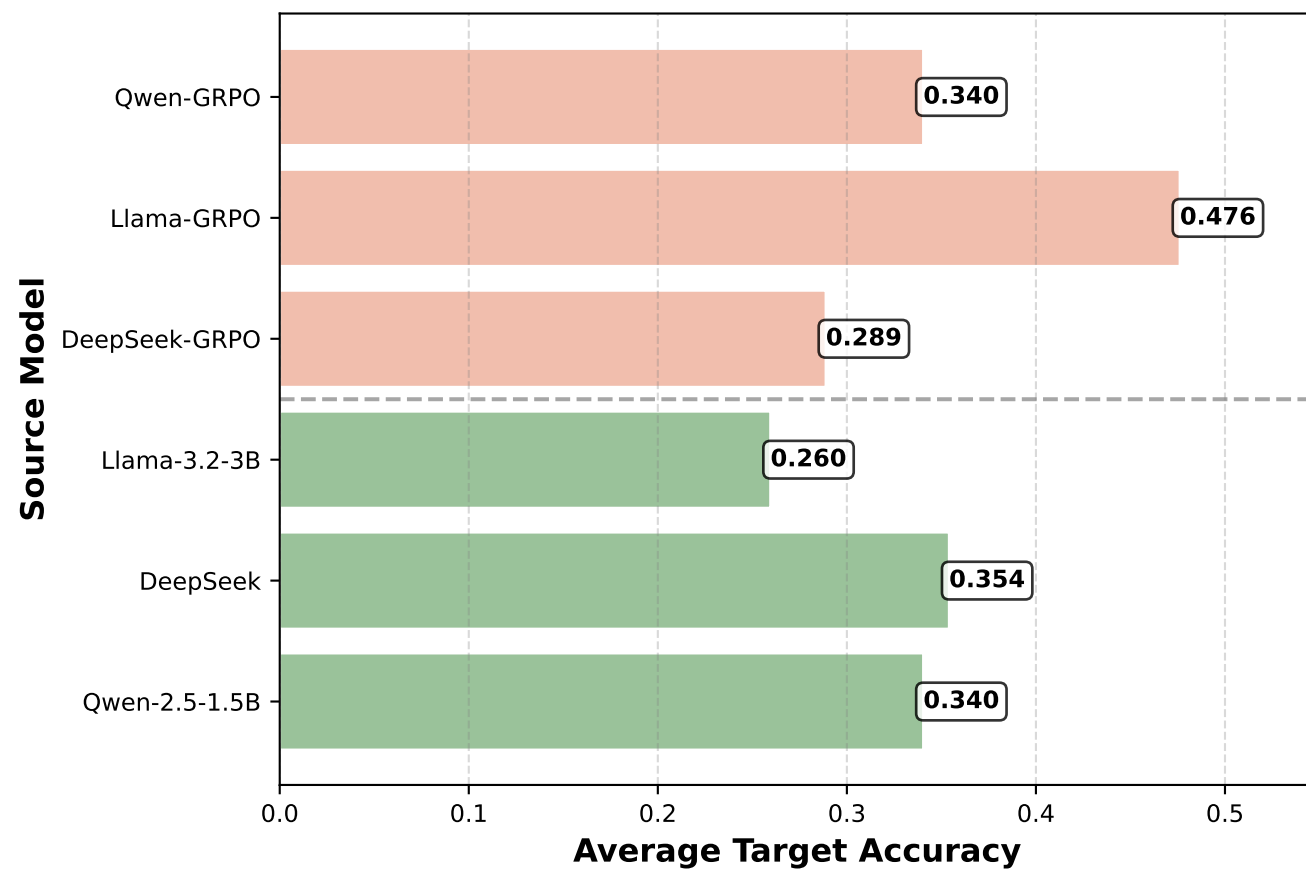
Transfer Consistency by Source Model (With Answer)

- Qwen-GRPO: 0.119
- Llama-GRPO: 0.629
- DeepSeek-GRPO: 0.566
- Llama-3.2-3B: 0.438
- DeepSeek: 0.121
- Qwen-2.5-1.5B: 0.119