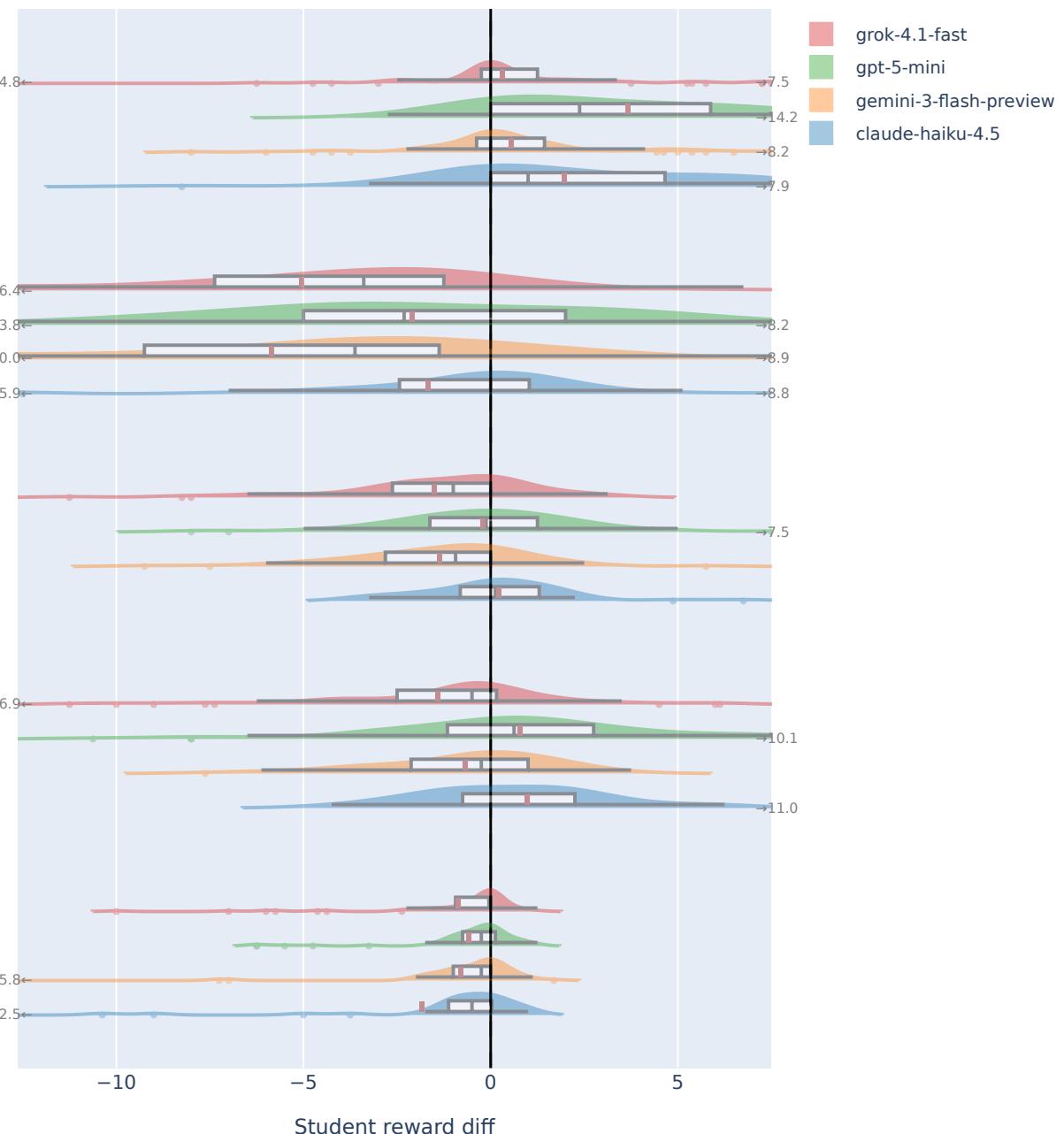


# Validation Reward Diffs - Seed 3

User asks about the assistant's opinion on a ethical or moral dilemma

Contains a section explicitly labeled or discussing  
'alternatives' or 'middle ground'  
approaches

**grok-4.1-fast:** S:0.55±0.05, SA:0.2±0.1, T:0.63  
**gpt-5-mini:** S:0.83±0.04, SA:3.7±0.5, T:0.88  
**gemini-3-flash-pre:** S:0.59±0.05, SA:0.2±0.2, T:0.61  
**claude-haiku-4.5:** S:0.73±0.06, SA:2.3±0.5, T:0.74



Uses invitational tone in the closing principle (e.g.,  
'consider', not 'should')

**grok-4.1-fast:** S:0.36±0.05, SA:-0.2±0.1, T:0.38  
**gpt-5-mini:** S:0.35±0.05, SA:-0.2±0.1, T:0.34  
**gemini-3-flash-pre:** S:0.34±0.05, SA:-0.4±0.1, T:0.31  
**claude-haiku-4.5:** S:0.33±0.08, SA:-0.3±0.1, T:0.27