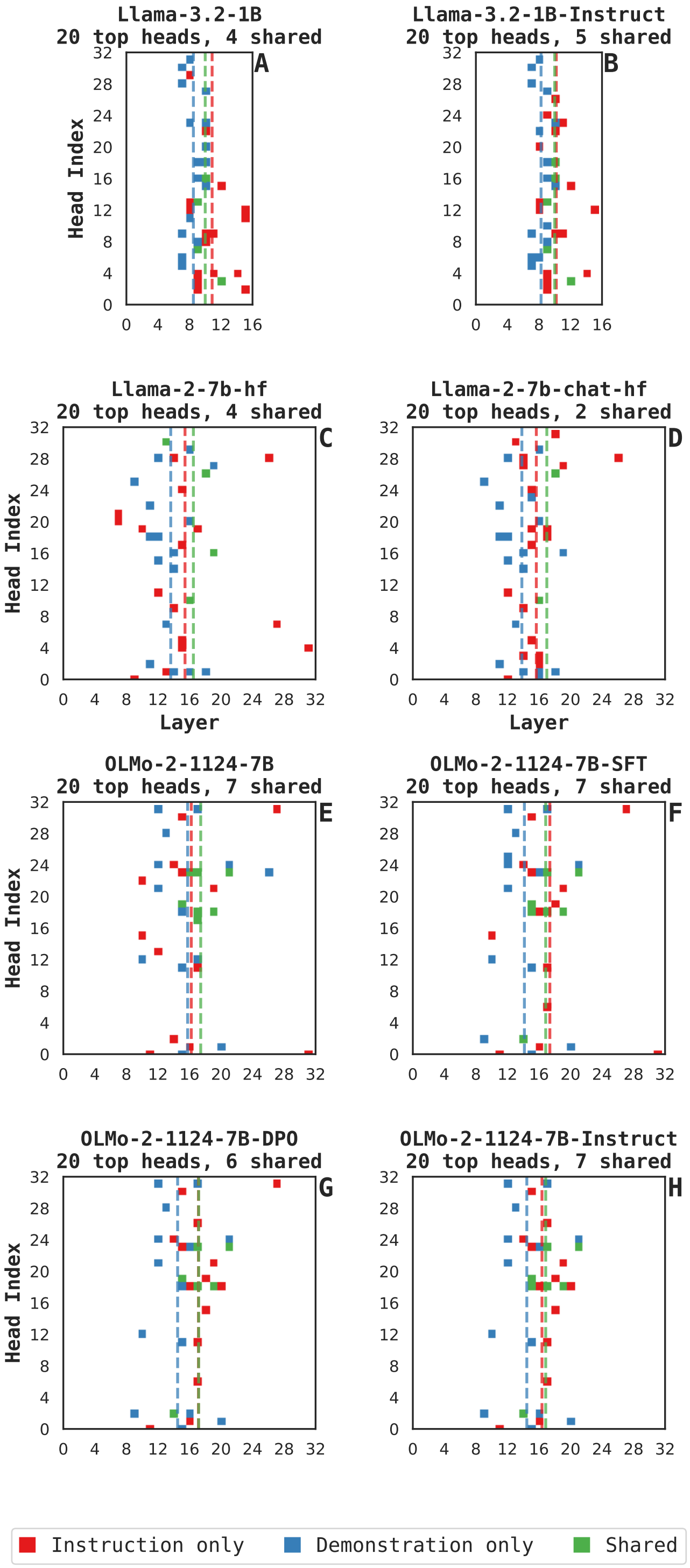
Llama-3.2-1B
20 top heads, 4 shared
A
Llama-3.2-1B-Instruct
20 top heads, 5 shared
B
Llama-2-7b-hf
20 top heads, 4 shared
C
Llama-2-7b-chat-hf
20 top heads, 2 shared
D
OLMo-2-1124-7B
20 top heads, 7 shared
E
OLMo-2-1124-7B-SFT
20 top heads, 7 shared
F
OLMo-2-1124-7B-DPO
20 top heads, 6 shared
G
OLMo-2-1124-7B-Instruct
20 top heads, 7 shared
H
Head Index
Layer
Instruction only
Demonstration only
Shared