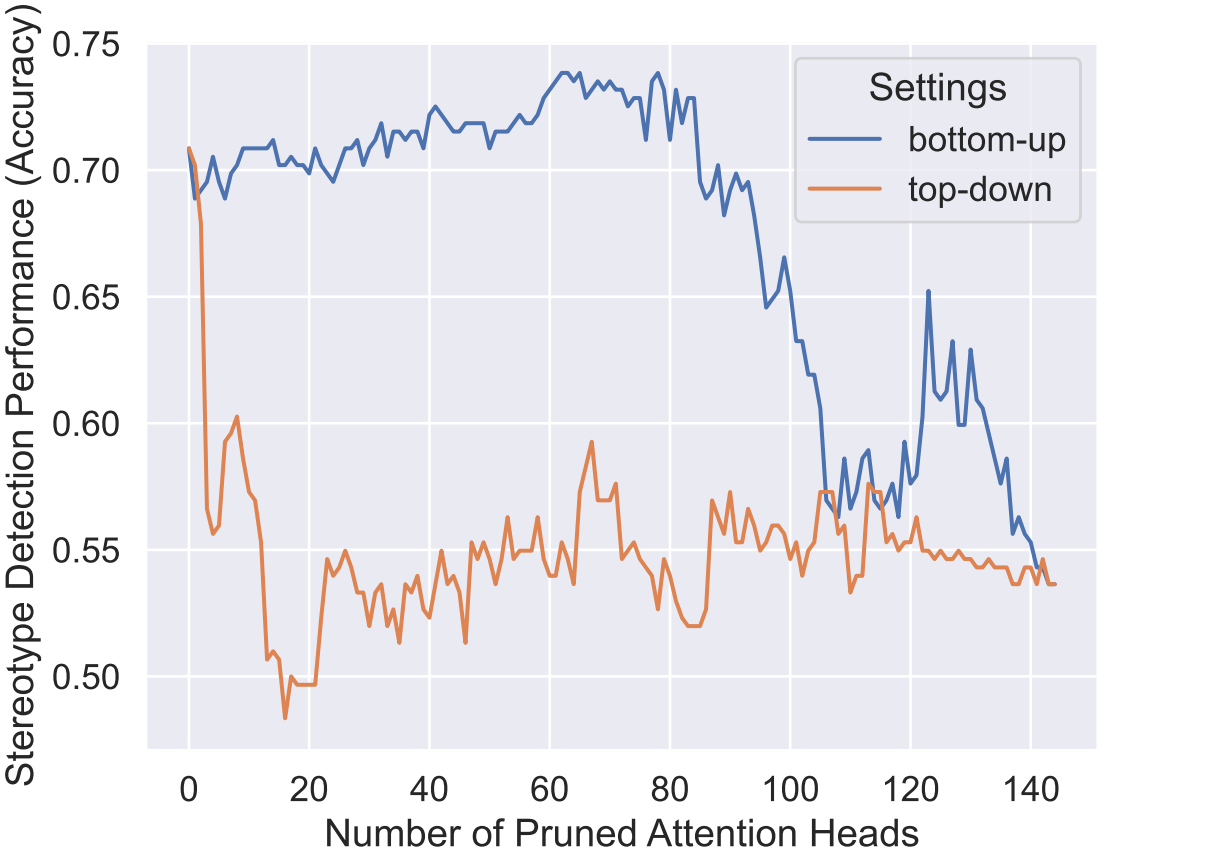
Stereotype Detection Performance (Accuracy)
0.75
0.70
0.65
0.60
0.55
0.50
0
20
40
60
80
100
120
140
Number of Pruned Attention Heads
Settings
bottom-up
top-down