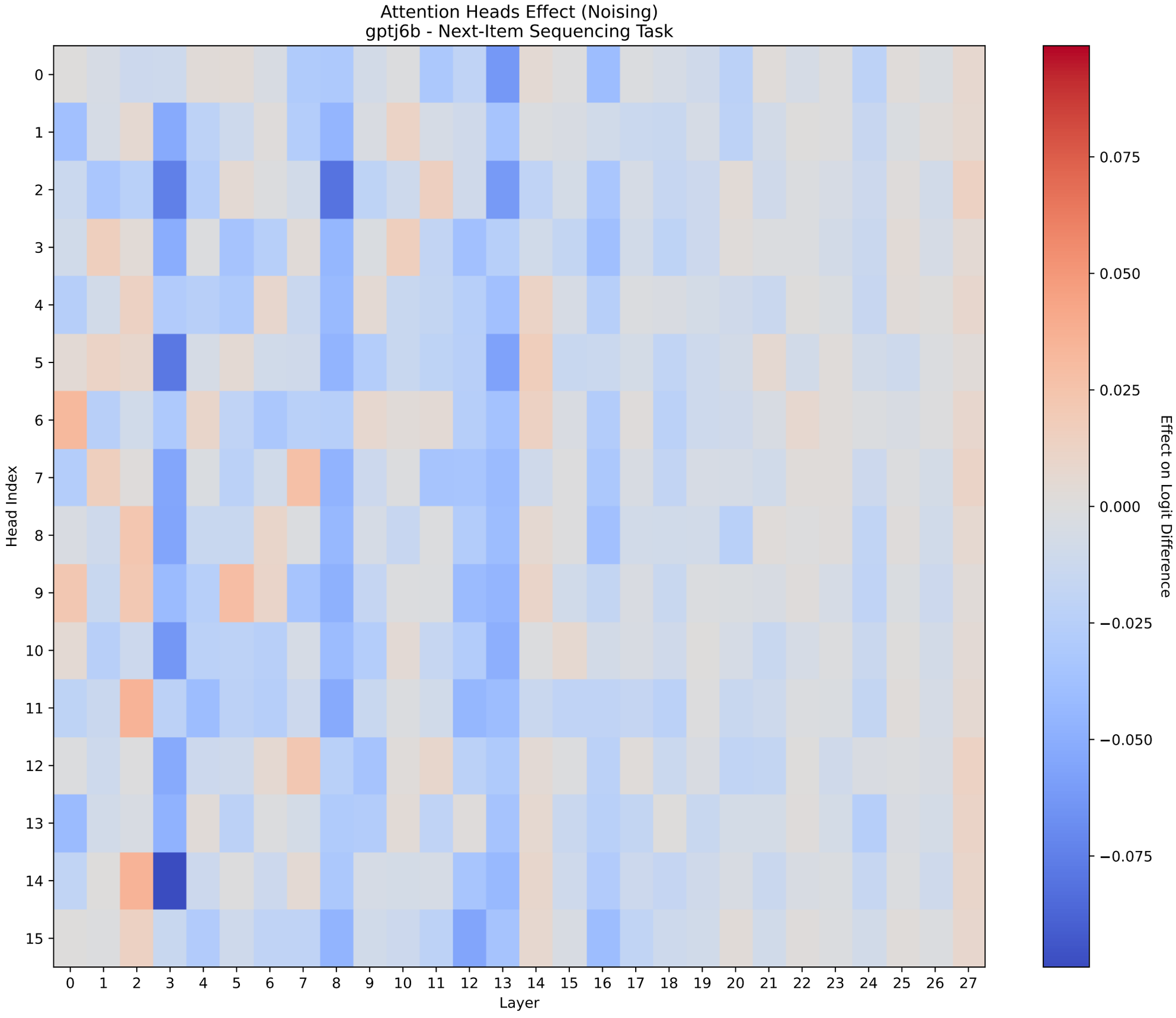
Attention Heads Effect (Noising)
gptj6b - Next-Item Sequencing Task
Head Index
Layer
Effect on Logit Difference
0.075
0.050
0.025
0.000
−0.025
−0.050
−0.075