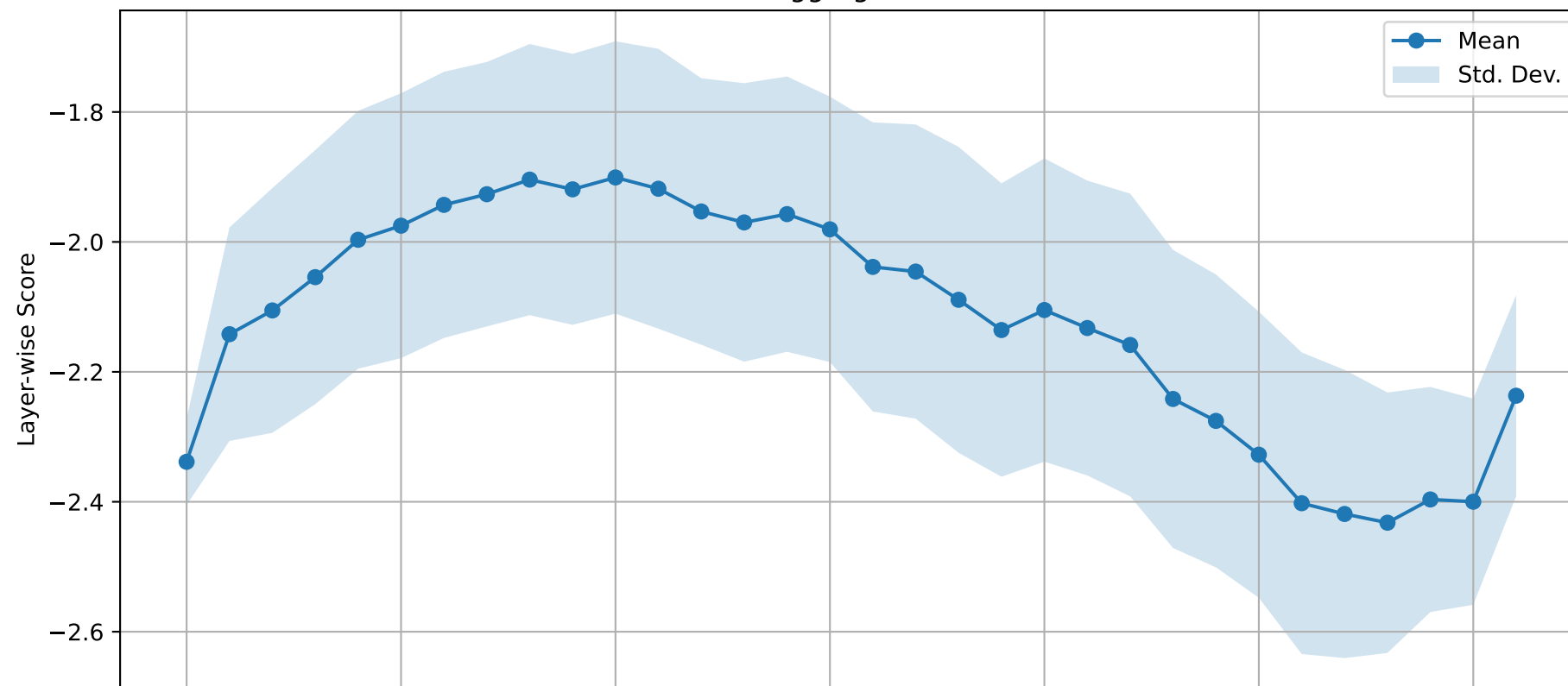


# Meta-Llama-3.1-8B-Instruct Aggregated and Overlaid Layer-wise Score ( $S^{(\ell)}$ )

Aggregated



Overlaid

