

Answer Locality: Intervention Effects by Mode (Layer 25)

$\Delta = \text{Intervention Accuracy} - \text{Baseline Accuracy}$   
*p-values: McNemar's Test | \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  | Bold = Significant*

