EXTRA 3 — (post-MLP − post-attn) difference heatmap