Q-Learning — Value Function Heatmap (b=7)