Pairwise Correlation of Reward Hacking Scores