

Paired Difference in Error Counts vs Gold

Axes truncated to [-15, 15] for clarity; extreme outliers excluded from view

