**Human-LLM Agreement**

Human vs LLM Scoring Agreement

Legend:
- Pearson (r=0.775)
- Spearman (r=0.769)
- ICC (r=0.659)
- Perfect Agreement (y=x)

X-axis: Human Raters Average

Y-axis: LLM Average Score