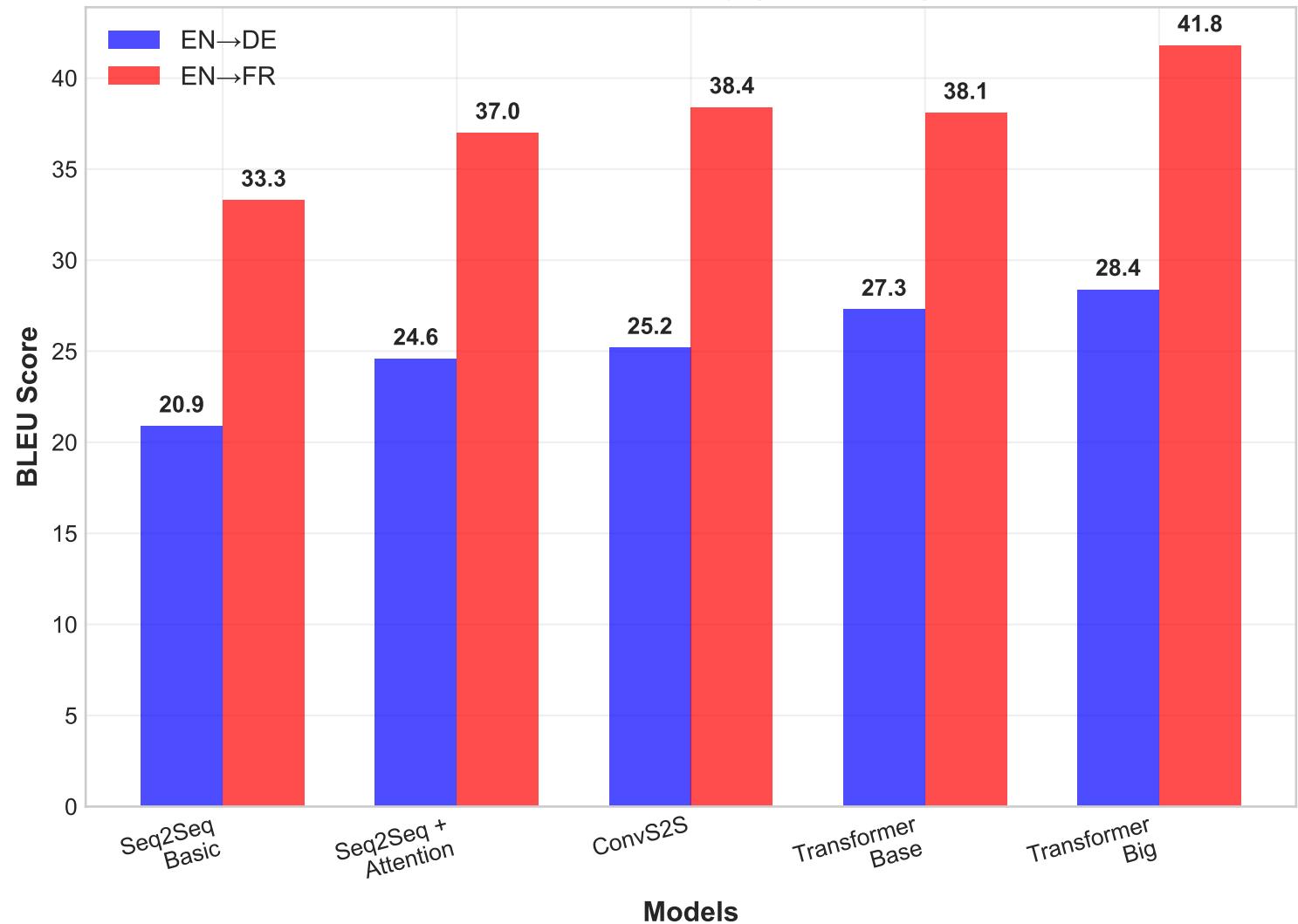


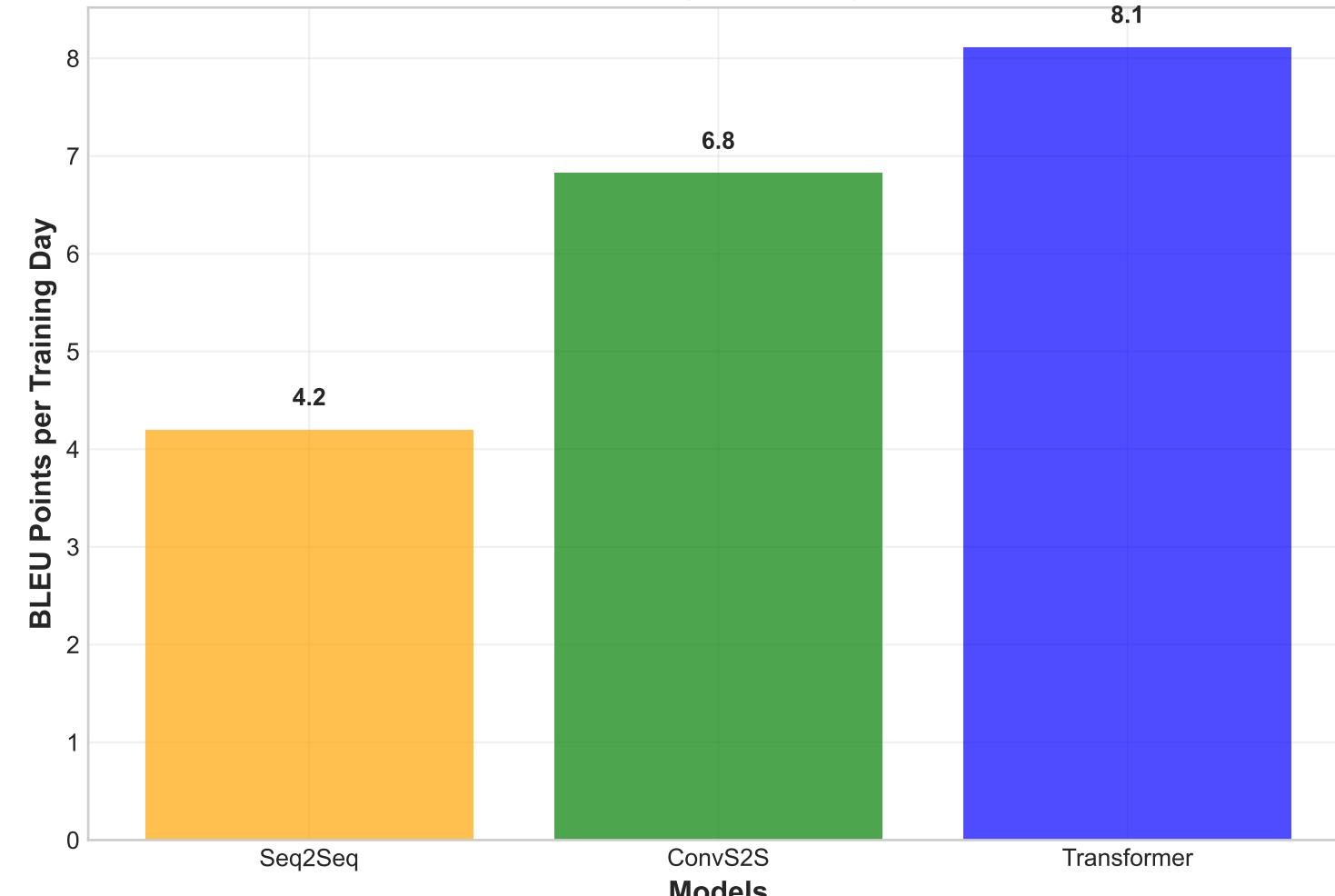
Transformer Revolution: Benchmark Results (2017-2018)

Translation Quality (WMT 2014)

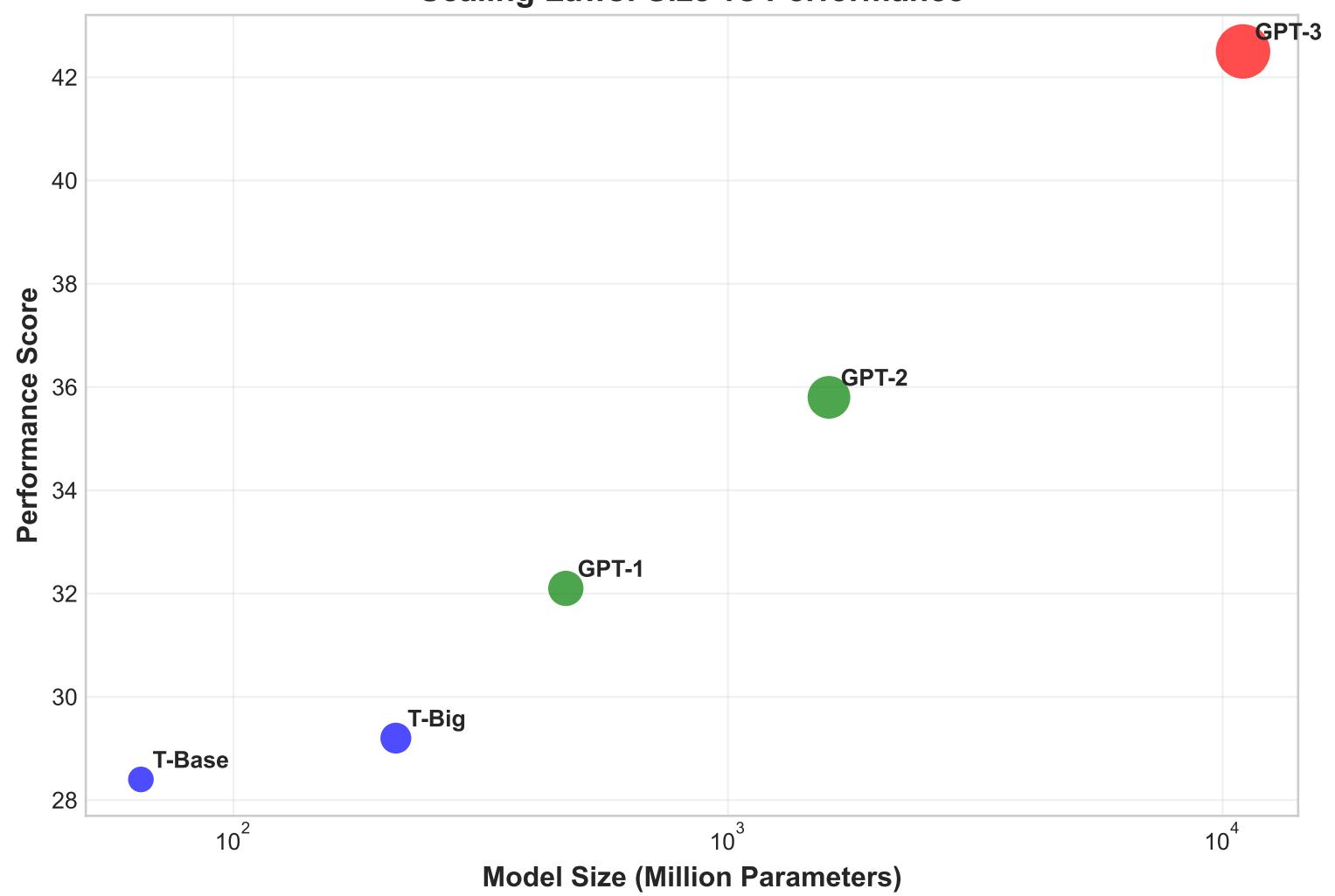


Models

Training Efficiency



Scaling Laws: Size vs Performance



Multi-Task Performance

