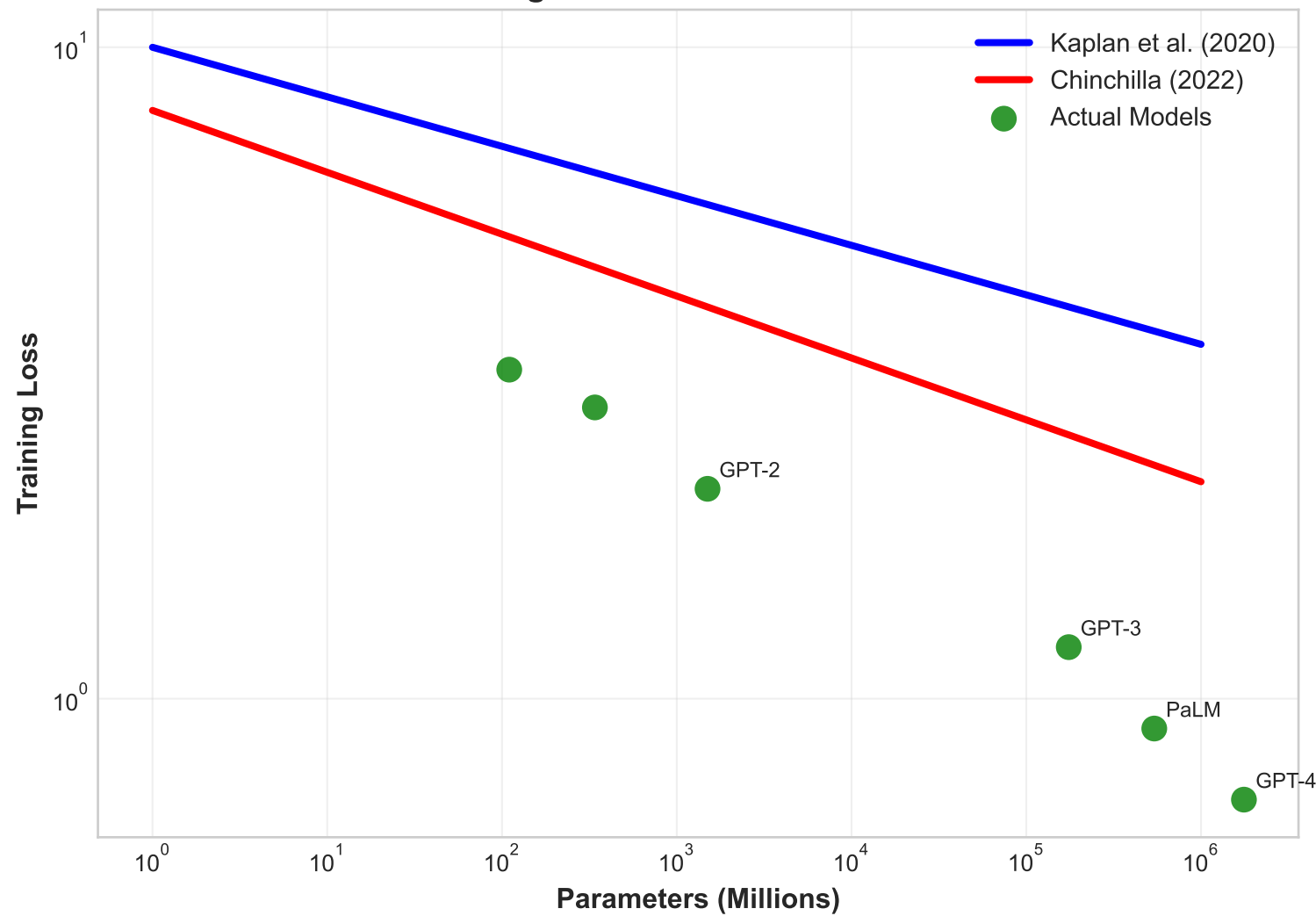
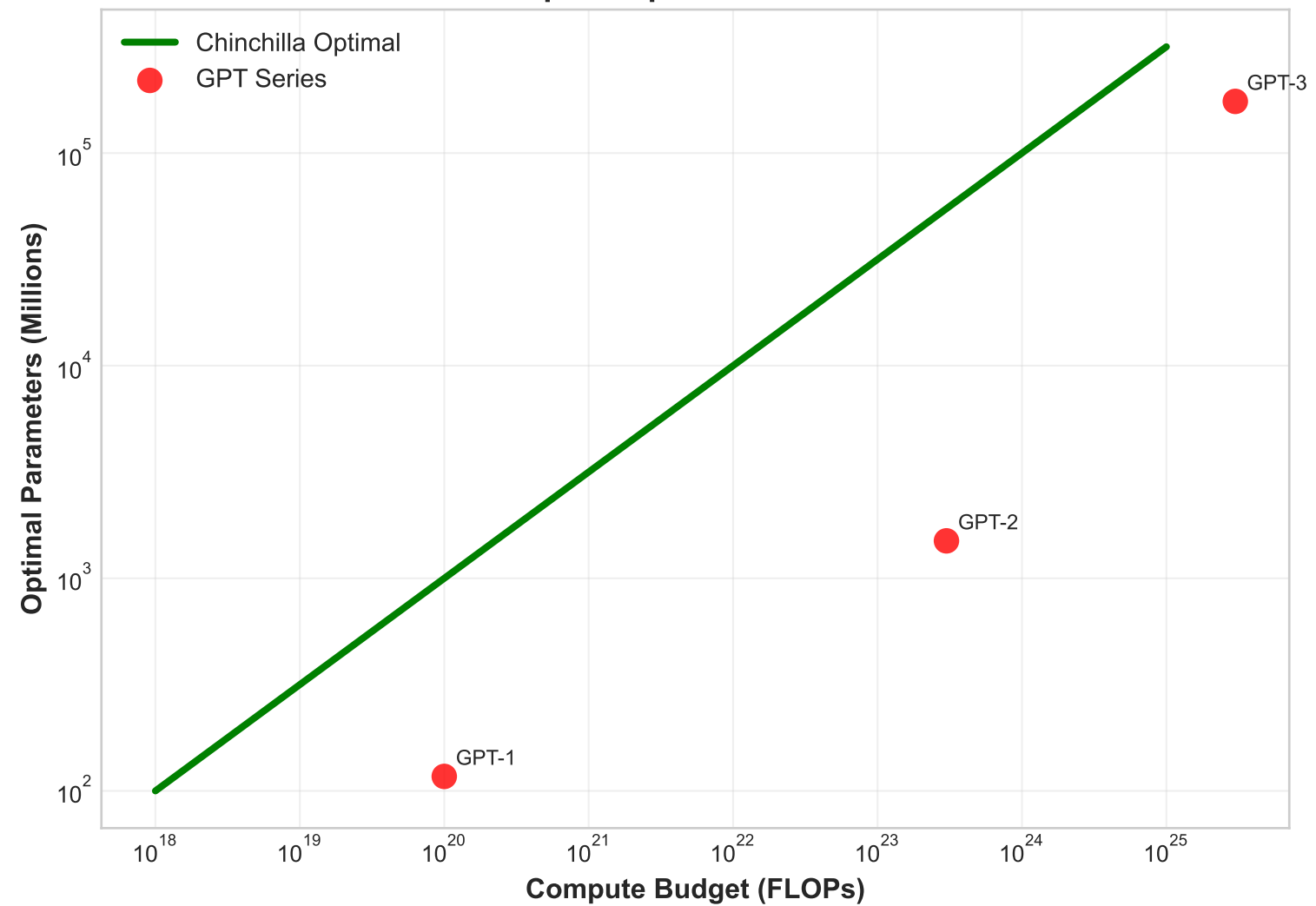


Neural Scaling Laws: The Science of Scale

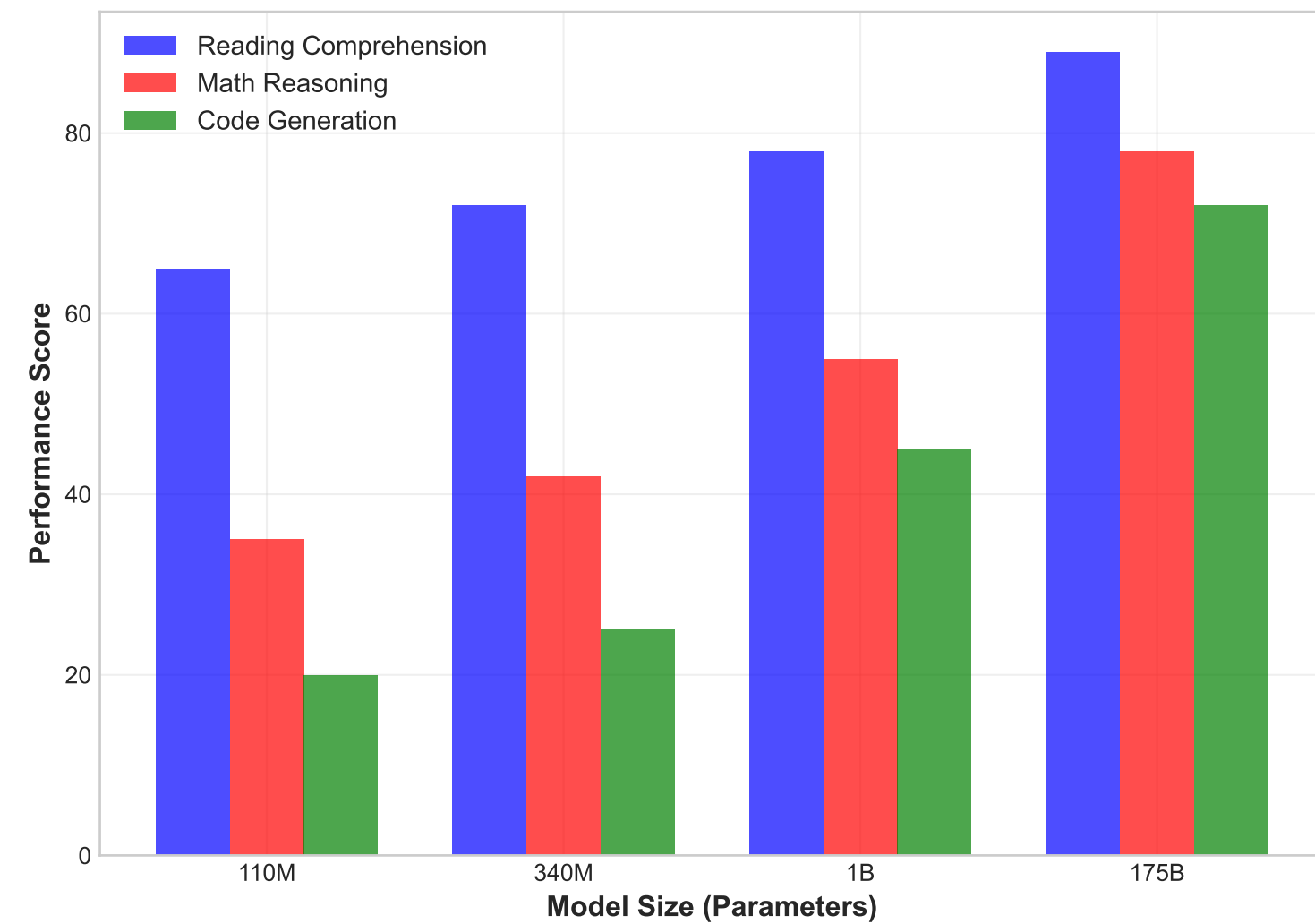
Scaling Laws: Loss vs Parameters



Compute-Optimal Model Size



Task Performance vs Scale



Emergent Capabilities

