pass@k Comparison: HUMAN_EVAL