EVALUATION scores:

T5\_small\_without\_story\_batch16\_hist\_20 epoch 11 (total 11 epochs)

coqar: {'BLEU': 0.30685283970262583, 'ROUGE': 0.35896066842148244, 'METEOR': 0.5198691725654717}

canard: {'BLEU': 0.46656539631274535, 'ROUGE': 0.6926266624315116, 'METEOR': 0.6925766475362978}

T5\_small\_with\_story\_batch16\_hist\_20 epoch 14 (total 15 epochs)

coqar: {'BLEU': 0.33039387181698016, 'ROUGE': 0.3764852688491561, 'METEOR': 0.5407842912882473}

canard: {'BLEU': 0.48668157929129185, 'ROUGE': 0.7126030932823676, 'METEOR': 0.7136904453802784}

T5\_base\_with\_story\_batch4\_hist\_20 epoch 6 (total 8 epochs)

coqar: {'BLEU': 0.3563690763390148, 'ROUGE': 0.3944557334941544, 'METEOR': 0.5685556980551302}

canard: {'BLEU': 0.506563362118428, 'ROUGE': 0.7279037952774894, 'METEOR': 0.7290982737685002}

T5\_base\_without\_story\_batch4\_hist\_20 epoch 9 (total 10 epochs)

coqar: {'BLEU': 0.3276754261281155, 'ROUGE': 0.40032564714617513, 'METEOR': 0.5624469310559154}

canard: {'BLEU': 0.49533192844519486, 'ROUGE': 0.7139690727547431, 'METEOR': 0.7142532378848528}

T5\_small\_with\_story\_batch4\_hist\_3 epoch 11 (total 15 epochs)

coqar: {'BLEU': 0.32612110700727504, 'ROUGE': 0.37786162286859776, 'METEOR': 0.5403182617129103}

canard: {'BLEU': 0.4855485637944266, 'ROUGE': 0.7123716068200318, 'METEOR': 0.7142737876421377}

T5\_small\_with\_story\_batch16\_hist\_3 epoch 20 (total 22 epochs)

coqar: {'BLEU': 0.3281161005656562, 'ROUGE': 0.37902213930470074, 'METEOR': 0.5437872781076998}

canard: {'BLEU': 0.4881608070125186, 'ROUGE': 0.7137786158777933, 'METEOR': 0.7153993624326819}

T5\_small\_with\_story\_batch16\_hist\_3\_mixed epoch 10 (total 12 epochs)

coqar: {'BLEU': 0.40389581211520253, 'ROUGE': 0.4279942932021437, 'METEOR': 0.597503882764309}

canard: {'BLEU': 0.4756776761410914, 'ROUGE': 0.7074058007706036, 'METEOR': 0.7075073407988256}