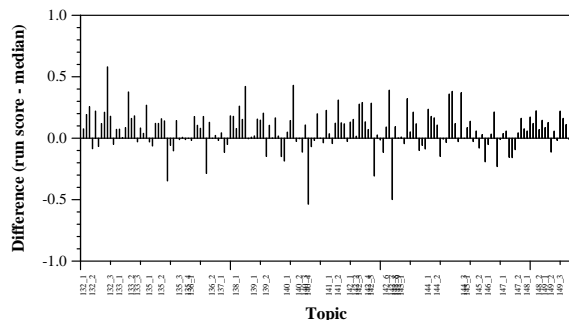
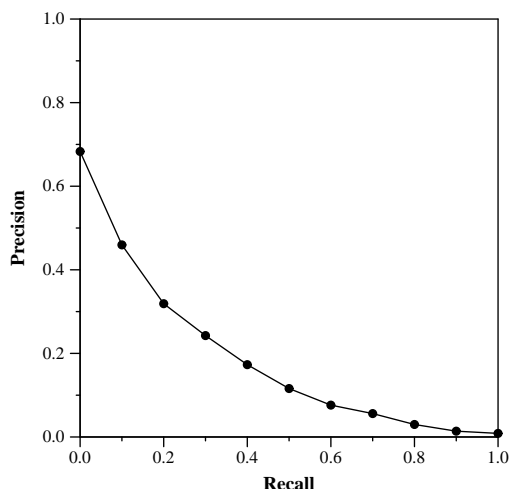


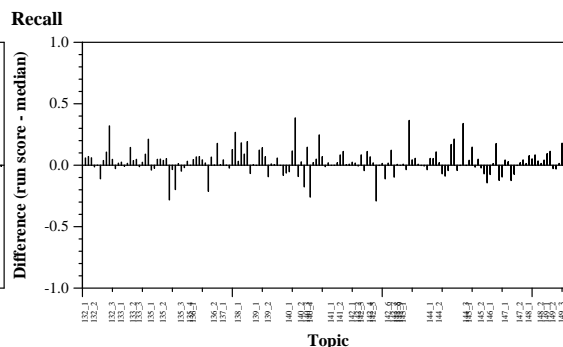
Run Description	
Run ID:	combine0.5
Category:	automatic topic files
Query understanding method:	performs generative query rewriting (CQR), including models like BART/T5
Query understanding data:	uses other external data
Ranking method:	uses traditional unsupervised sparse retrieval (e.g. QL, BM25, etc.); performs re-ranking with a pre-trained neural language model (BERT, Roberta, T5, etc.)
Ranking data:	is trained with TREC Deep Learning Track and/or MS MARCO dataset
Initiative:	does not treat them specially
Response generation:	uses single source (single passage)

Note: all P/R scores follow the strict criterion $rel \geq 2$. NDCG gains are the rel values.

Overall measures		Document Level Averages	
Number of turns	163		
Total number retrieved	163000		
Total relevant	12318		
Total relevant retrieved	2912		
MAP	0.1737		
Mean NDCG@3	0.4392		
Mean NDCG@20	0.3839		
Mean Precision@1	0.5276		
Mean Precision@3	0.4785		
		Precision	
		At 5 docs	0.4344
		At 10 docs	0.3908
		At 20 docs	0.3267
		R-Precision	
		Exact	0.2268



Difference from median NDCG



Difference from median AP