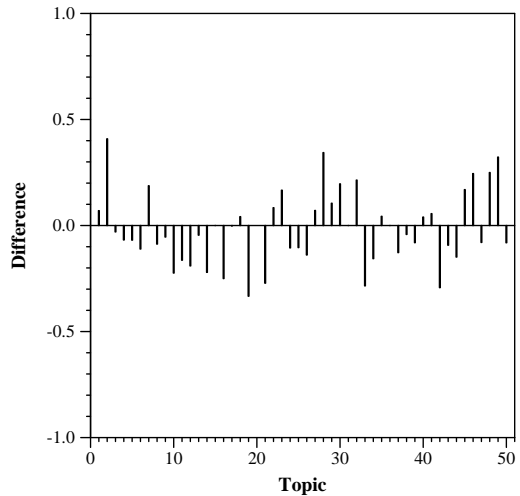Phase 1: Each group submitted a set of 5 documents per topic to be used as relevance feedback input in Phase 2 by 3 to 5 groups. One or two sets submitted. Evaluation output includes number of relevant documents in set, and how well other groups did on this set compared to the other sets that that group ran (each group ran 7 to 8 Phase 1 sets). Comparison numbers totaled among the collection and evaluation measures used in Phase 2. Total score = B / (B + W) where B is the total number of runs/measures this set did better than, and W is the number this set did worse on.

Phase 1 Summary Statistics			
RF Input Set		SIEL.1	
Total Num Rel in Set		71	
Measure	Coll	Num Worse Than	Num Better Than
MAP(all)	Full	6	1
P(10)(all)	Full	5	2
statMAP (NEU) (all)	B	12	8
eMAP (UMass) (all)	B	11	9
Measure			Score
Score (all)			0.3704
Score (average over q)			0.4660

Score Per Topic Diff from Median, SIEL.1