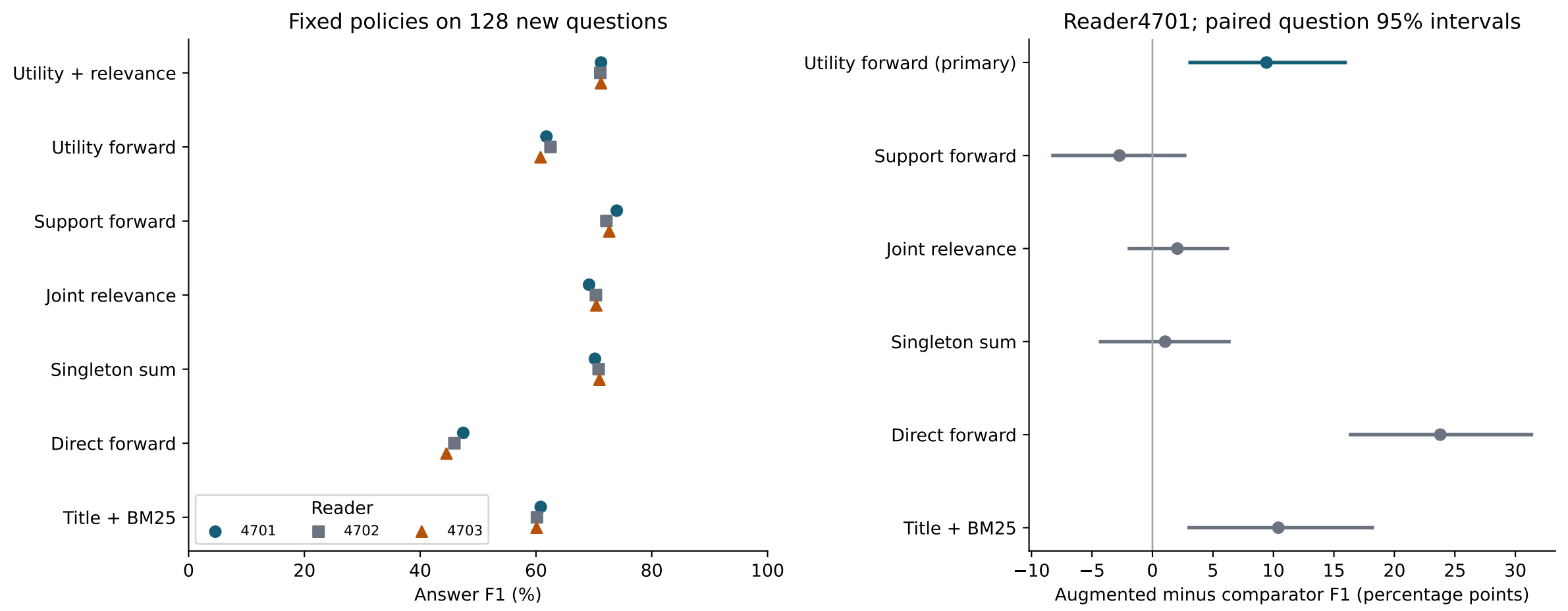

Fixed policies on 128 new questions
Utility + relevance
Utility forward
Support forward
Joint relevance
Singleton sum
Direct forward
Title + BM25
Reader
4701
4702
4703
0
20
40
60
80
100
Answer F1 (%)
Reader4701; paired question 95% intervals
Utility forward (primary)
Support forward
Joint relevance
Singleton sum
Direct forward
Title + BM25
−10
−5
0
5
10
15
20
25
30
Augmented minus comparator F1 (percentage points)