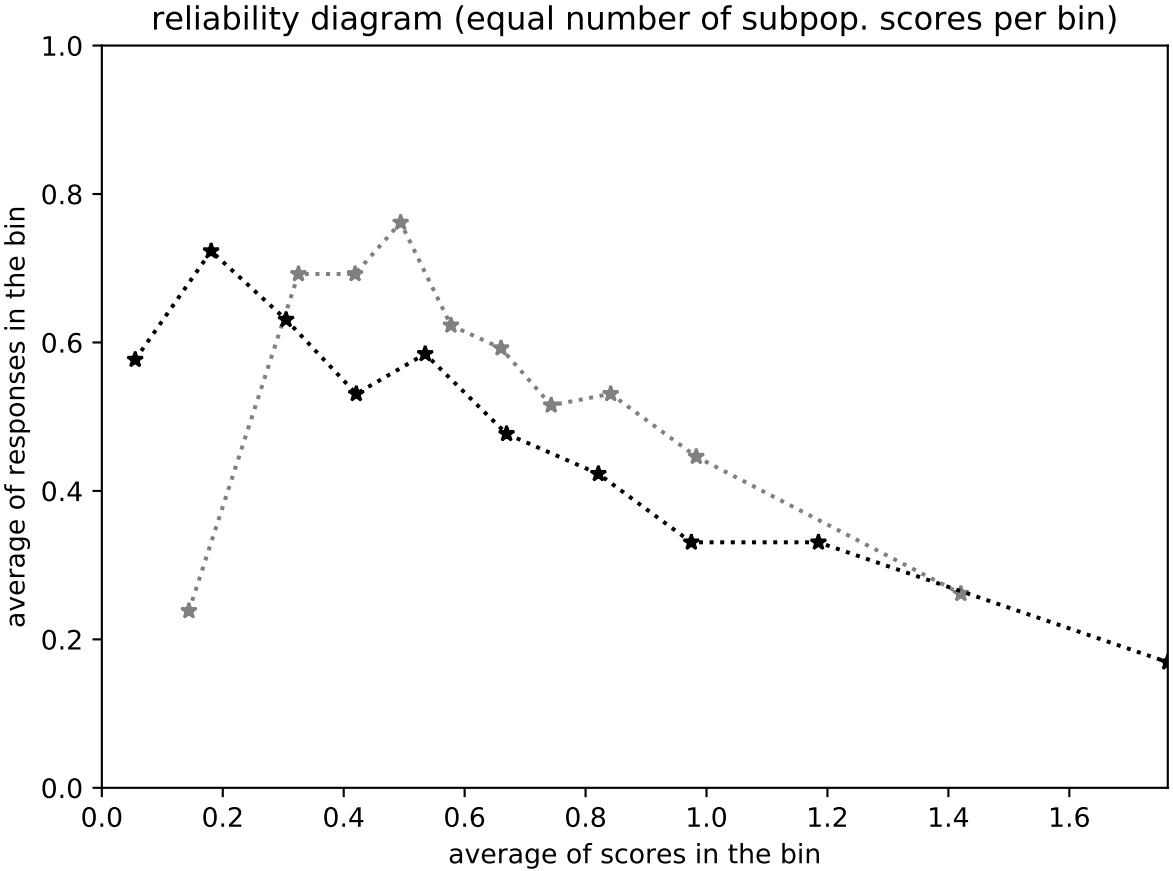
reliability diagram (equal number of subpop. scores per bin)
average of responses in the bin
average of scores in the bin
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
1.6