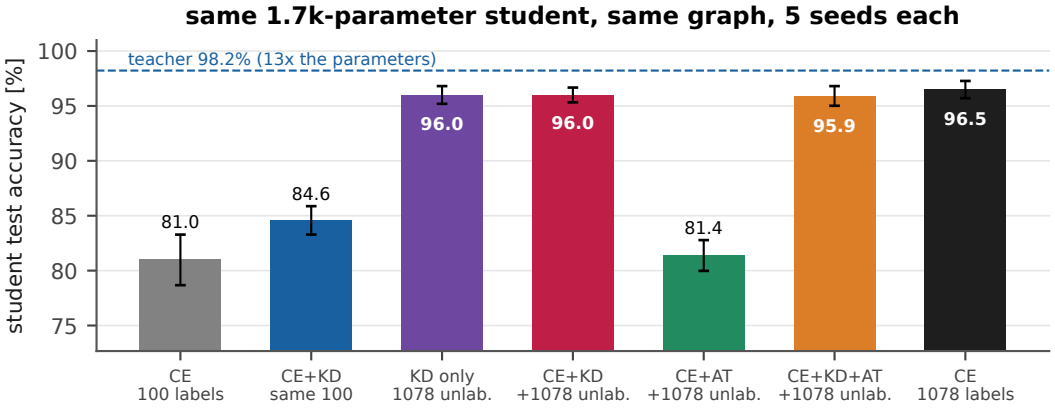

same 1.7k-parameter student, same graph, 5 seeds each
teacher 98.2% (13x the parameters)
student test accuracy [%]
100
95
90
85
80
75
81.0
84.6
96.0
96.0
81.4
95.9
96.5
CE
100 labels
CE+KD
same 100
KD only
1078 unlab.
CE+KD
+1078 unlab.
CE+AT
+1078 unlab.
CE+KD+AT
+1078 unlab.
CE
1078 labels