Direct Assessment: Adequacy

Graham et al. (2013) propose a simple continuous scale:

- To what extent MT adequately expresses the meaning of REF?

This HIT consists of 100 English assessments. You have completed 0.
Read the text below. How much do you agree with the following statement:

The black text adequately expresses the meaning of the gray text in English.

To snobs like me who declare that they'd rather play sports than watch them, it's hard to see the appeal of watching games rather than taking up a controller myself.

Snob like me, who say that it is better to be in sports than watching him, it is hard to understand the appeal of having to watch the game, rather than to take a joystick in hand.

0 %   100 %

$\oplus$ After $\sim$15 judgements, each annotator stabilizes.
$\ominus$ Interpretable by averaging over many judgements of many people.
$\ominus$ 30–70(!)% of participating Turkers unreliable.
$\ominus$ Too few non-English speakers on Amazon Mechanical Turk.