The problem of visual tracking evaluation is sporting a large variety of performance measures, and largely suffers from lack of consensus about which measures should be used in experiments. This makes the cross-paper tracker comparison difficult. Furthermore, as some measures may be less effective than others, the tracking results may be skewed or biased toward particular tracking aspects. In this paper, we revisit the popular performance measures and tracker performance visualizations and analyze them theoretically and experimentally. We show that several measures are equivalent from the point of information they provide for tracker comparison and, crucially, that some are more brittle than the others. Based on our analysis, we narrow down the set of potential measures to only two complementary ones, describing accuracy and robustness, thus pushing toward homogenization of the tracker evaluation methodology. These two measures can be intuitively interpreted and visualized and have been employed by the recent visual object tracking challenges as the foundation for the evaluation methodology.