SPORTSCIENCE sportsci.org
Perspectives / Performance

How to Interpret Changes in an Athletic Performance Test

Will G Hopkins

Sportscience 8, 1-7, 2004 (sportsci.org/jour/04/wghtests.htm)
Sport and Recreation, Auckland University of Technology, Auckland 1020, New Zealand. Email.
Reviewers: Christopher J Gore, Australian Institute of Sport, PO Box 219, Brooklyn Park, South Australia 5032; David B Pyne, Australian Institute of Sport, PO Box 176, Belconnen, ACT 2616, Australia

When monitoring progression of an athlete with performance or other fitness tests, it is important to take into account the magnitude of the smallest worthwhile enhancement in performance and the uncertainty or noise in the test result. For elite athletes competing in sports as individuals, the smallest worthwhile enhancement would give the athlete an extra medal per 10 competitions; the required change in performance is 0.3 of the typical variation in an athlete's performance from competition to competition, or ~0.3-1% when expressed as a change in power output, depending on the sport. In team sports, where there is no direct relationship between team and test performance, an appropriate default for the smallest change in test performance is one-fifth of the between-athlete standard deviation (a standardized or Cohen effect size of 0.20). Noise in a test result is best expressed as the typical or standard error of measurement derived from a reliability study. The noise in most performance tests is greater than the smallest worthwhile difference, so assessments of changes in performance can be problematic. An exact but somewhat impractical solution is to present chances that the true change is beneficial, trivial, and harmful. A simpler approach is to apply systematic rules to decide whether the true change is beneficial, trivial, harmful, or unclear. Unrealistically large changes can also be partially discounted when tests are noisy.

KEYWORDS: Bayes, correlation, error of the estimate, error of measurement, limits of agreement, reliability, time to exhaustion, time trial, validity.

Reprint pdf · Reprint doc · Slideshow · Commentaries by Gore and Pyne

Update Feb 2018. This article/slideshow is yet to be updated to take into account the new spreadsheet for monitoring changes and trend in an individual (see article and spreadsheet) and the fact that thresholds for assessing magnitude of the typical error are now half those for a change in the mean (as stated in the article and slideshow on validity and reliability).

An update on smallest important changes in competitive performance of non-interactive athletes (track and field, rowing, time-trial cycling, etc.), along with a consideration of medal winning in interactive non-match competitions (e.g., road cycling) and matches (one-on-one and team sports), is available in a slideshow on medal-winning enhancements of performance presented at the performance-analysis conference and Olympic conference in 2016.

Update Sept 2011. The smallest worthwhile change is now stated as 0.3 of the variation in an athlete's performance, not 0.5 as previously.

The basis for this article is an updated version of a slideshow accompanying a talk entitled "making sense of performance tests", which I presented earlier this year at the Scottish Institute of Sport and more recently at a local conference. The talk was based mainly on previous research by my colleagues and me, along with some new and previously unpublished insights. The title now better reflects the emphasis on monitoring an athlete's performance from test to test.

Monitoring the progression of athletes with regular performance and other fitness-related tests is a widespread and apparently useful practice in upper competitive levels of most if not all sports in wealthy countries, but in my experience lack of understanding about the interpretation of changes in test scores is also widespread. Perhaps the most important issue is that of magnitude: to interpret the change in an athlete's performance since a previous test, you need some concept of the magnitude of change that matters to the athlete in his or her sport. The first section of the talk is therefore concerned with identifying the smallest worthwhile change in performance. Your ability to track such changes with a performance test depends on the validity and reliability of the test, which I explain in the second section. The final section is devoted to several ways of interpreting the test results for the athlete or coach. See also commentaries by Christopher Gore and David Pyne, to whom I am indebted for valuable interactions and feedback on this topic.

The reprint pdf version of this article contains printer-friendly images of the PowerPoint slideshow and references. View the slideshow to see each slide build sequentially.

Published Nov 2004