

## Competitive Performance of Elite Track-and-Field Athletes: Variability and Smallest Worthwhile Enhancements

Will G Hopkins

Sportscience 9, 17-20, 2005 (sportsci.org/jour/05/wghtrack.htm)

Sport and Recreation, Auckland University of Technology, Auckland 1020, New Zealand. [Email](#).

Reviewer: Esa Peltola, Aspire Academy for Sports Excellence, Doha, Qatar.

**PURPOSE.** To describe the reproducibility of competitive performance of elite track-and-field athletes and to derive the smallest worthwhile enhancements of performance in these events. **METHODS.** The data were official results of events in 17 competitions of an annual series of the International Amateur Athletic Federation extending over 101 d. Typical within-athlete variability from competition to competition was derived as a coefficient of variation by repeated-measures analysis of log-transformed times (for running and hurdling events) or distances (for jumping and throwing events). The smallest worthwhile performance enhancement was taken as half the within-athlete variability. **RESULTS and DISCUSSION.** Within-athlete variabilities were as follows: running and hurdling events up to 1500 m, 1.0%; longer runs and steeplechase, 1.4%; triple and high jump, 1.7%; pole vault and long jump, 2.4%; discus, javelin, and shot put, 2.8% (90% confidence limits all  $\sim x/\div 1.13$ ). The differences between events presumably reflect differing contributions of energy systems, pacing strategies, wind resistance and skill. Females may have had a little more variability in performance ( $\sim 1.1x$ ) than males in some events, possibly because of less depth of competition. There was some evidence that variability increased with increasing time between competitions for the short running events (from  $\sim 0.7\%$  for  $\sim 1$  wk to  $\sim 1.1\%$  for  $\sim 100$  d). The top-half athletes in each event were less variable than the bottom-half in running and hurdling up to 1500 m (0.8 vs 1.1%) and in longer runs and steeplechase (1.1 vs 1.6%), but differences were unclear in the other events. A likely explanation is less consistent motivation in endurance athletes who were not in the medal stakes. **CONCLUSIONS.** Coaches and sport scientists should focus on enhancements of as little as 0.3-0.5% for elite track athletes through 0.9-1.5% for elite field athletes.

**KEYWORDS:** competition, error, race, reliability, reproducibility, testing.

[Reprint pdf](#) · [Reprint doc](#) · Commentary by [Esa Peltola](#).

Introduction .....	17
Methods .....	18
Results and Discussion.....	18
Effect of Event .....	18
Effect of Sex .....	19
Effect of Time between Competitions.....	19
Effect of Caliber of Athlete.....	19
Conclusions.....	20
References.....	20

### Introduction

This paper is the latest in a series aimed at estimating the smallest worthwhile change in performance for athletes who compete as individuals in sports where the outcome is determined by a single score, such as a time or dis-

tance. The smallest worthwhile change in performance is important when assessing athletes with a performance test to make decisions about meaningful changes in an individual or to research strategies that might affect performance

(Hopkins, 2004). An estimate of the smallest change comes from an analysis of reliability (reproducibility or variability) of competitive performance—the smallest change is in fact about half the typical variation a top athlete shows from competition to competition (Hopkins et al., 1999).

The previous published studies on variability of competitive performance and smallest changes have been for junior swimmers (Stewart and Hopkins, 2000), elite swimmers (Pyne et al., 2004), non-elite runners (Hopkins and Hewson, 2001), and triathletes (Paton and Hopkins, 2005). The present study of track-and-field athletes is based on data that I acquired and analyzed some years ago and that I have referred to in various publications.

### Methods

Official result times of the 1997 Grand Prix series of international competitions were obtained from the website of the International Amateur Athletic Federation. The series consisted of 18 different kinds of track-and-field events staged at 17 mainly European venues over 101 days. An event at a venue was included in the analysis of reliability for that kind of event if it included at least 2 athletes who had entered the same event at other venues. The men's high jump provided the least amount of data: 8 athlete-entries for 3 athletes at 3 venues; at the other extreme, the men's 110-m hurdle provided 120 athlete-entries for 20 athletes at 17 venues. A typical women's event in the analysis was the javelin, which provided 48 athlete-entries for 12 athletes at 7 venues. There were insufficient data for the analysis of hammer throw, women's long jump and women's pole vault.

The analyses were similar to those used in the study of triathlete performance in this issue (Paton and Hopkins, 2005). Briefly, I used mixed modeling of log-transformed times to derive an athlete's typical percent variation in performance from competition to competition as a coefficient of variation. I performed separate analyses for males and females in each event, and for the top and bottom half of athletes in each event. Differences between coefficients of variation were considered substantial if their ratio was greater than 1.10.

I also analyzed for the effect of time on variability estimated between all pairs of competitions for both sexes combined but for shorter

(100- to 1500-m) and longer (3000- to 10,000-m) running events separately. I corrected the small bias in the individual estimates of coefficients of variation by multiplying by  $1+1/(4DF)$ , where  $DF$ =degrees of freedom (Gurland and Tripathi, 1971). I then fit quadratics to the log-log plots and used 1000 bootstrapped samples to derive confidence limits for the quadratics and for comparisons (ratios) of the coefficients of variation for different times between competitions.

## Results and Discussion

### Effect of Event

Table 1 shows the typical within-athlete variation in performance from competition to competition for the various events. I have not systematically derived confidence limits for a comparison of the variability in the different types of event, but it is reasonably clear from the confidence limits for each type that athletes in longer running events are more variable their performance than those in the shorter events, that athletes in the throwing events are about twice as variable, and that athletes in the high jump and triple jump are somewhere in between.

Table 1. Typical variability of a track-and-field athlete's performance between international competitions, expressed as a coefficient of variation (CV).

Event	CV (%) (90% conf. limits)
Running <3 km <sup>a</sup>	1.0 (0.9–1.1)
Running 3–10 km <sup>b</sup>	1.4 (1.2–1.6)
High jump, triple jump	1.7 (1.5–1.9)
Long jump, pole vault	2.4 (2.1–2.7)
Discus, javelin, shot put	2.8 (2.4–3.2)

<sup>a</sup>100- to 1500-m runs; 100- to 400-m hurdles.  
<sup>b</sup>1500- to 10,000-m runs; male 3000-m steeplechase.

The higher reliability of the shorter running and hurdling events may be due to differing contributions of energy systems, pacing strategies, and wind resistance relative to the longer events. Contributions of energy systems and skill may explain the lower reliability of field events and differences between the field events. The differences between variability of performance in the different types of event mirrors those in performance tests in these modes of exercise (Hopkins et al., 2001), although the

variability in these competitions is generally a little less than that for athletes in the best tests.

### Effect of Sex

Table 2 shows the variability in performance for females and males in the events where there were sufficient comparable data. Given the uncertainty in the estimates of variability, females were probably more variable than males by a trivial-small factor of  $\sim 1.1$  (about 10%) overall, but there may be greater or smaller differences in specific events. The difference may be due to less depth of competition for the females rather than differences in physiology.

Table 2. Variability of performance of female and male track-and-field athletes expressed as coefficients of variation (CV). Comparison of variabilities is shown as ratio of female/male.

Event	CV (%)		Ratio (90% conf. limits)
	female	male	
Running <3 km	1.1	0.9	1.2 (1.0–1.4)
Running 3-10 km	1.3	1.4	1.0 (0.8–1.3)
High jump, triple jump	1.7	1.6	1.0 (0.7–1.5)
Discus, javelin, shot put	3.2	2.6	1.2 (1.0–1.5)

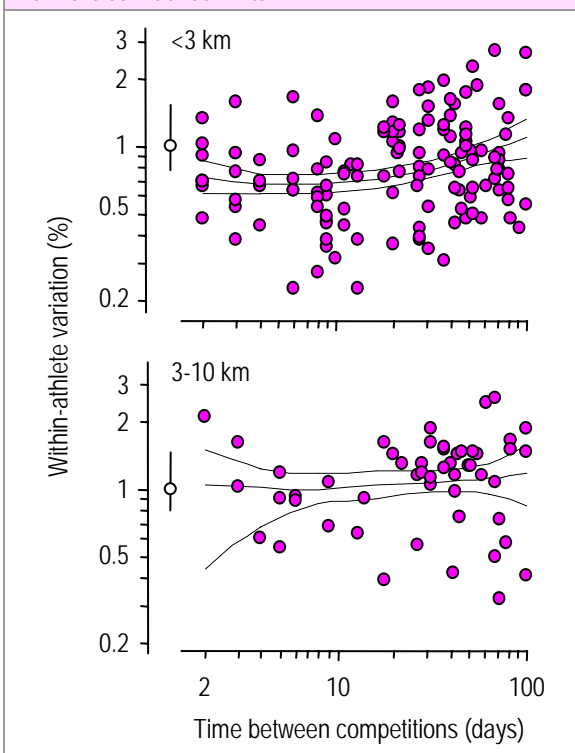
### Effect of Time between Competitions

The estimates of athlete variability in running and hurdling events for all pairwise combinations of competitions are shown in Figure 1. Much of the scatter in the points is due to sampling variation arising from the small sample size for the pairwise combinations, as can be seen from the expected sampling variation for a typical point.

Athlete variability for short runs was minimum (0.7%) at around 1 wk between competitions and greatest at 100 d (1.1%). The trend towards more variability with increasing time between competitions was clear: for example, the ratio of variability at 64 d to that at 8 d was 1.40 (90% confidence limits 1.16–1.65). The quadratic model probably overestimates the trend for longer times, because a plateau is evident in the plot beyond  $\sim 50$  d. A small increase in variability due to variation in training and health over a period of weeks is not unexpected, but over more than several months these athletes, like elite triathletes (Paton and Hopkins, 2005), probably maintain their ability to perform.

Variability for the long runs was also a minimum (1.0%) around 1 wk between compe-

Figure 1. Typical variation in an athlete's performance between all pairs of competitions for the short running and hurdling events (< 3 km) and the long running events (3-10 km), plotted against time between the competitions. Bars are standard deviations representing typical sampling variation for a true variation of 1% with the average sample size of the points (6 athletes). Curves are quadratics, with 90% confidence limits.



titions for this sample. However, confidence limits were too wide to allow conclusions about any substantial trend; for example, the ratio of 100-d to 8-d variability was 1.16 (0.77–1.69).

Table 3. Variability of performance of track-and-field athletes who were overall in either the bottom half or the top half of their event when ranked over all competitions of the Grand Prix series. Variabilities are expressed as coefficients of variation (CV). Comparison of variabilities is shown as ratio of bottom/top halves.

Event	CV (%)		Ratio (90% conf. limits)
	Bottom half	Top half	
Running/hurdles <3 km	1.1	0.8	1.3 (1.1–1.5)
Running 3-10 km	1.6	1.1	1.5 (1.2–1.9)
High jump, triple jump	1.7	1.9	0.9 (0.6–1.3)
Long jump, pole vault	2.3	2.1	1.1 (0.9–1.4)
Discus, javelin, shot put	2.5	3.3	0.8 (0.6–1.1)

### Effect of Caliber of Athlete

Table 3 shows that the athletes in the top

half of the field were clearly less variable in the running events. Others have found similar results with running and swimming and cycling and have attributed it to better pacing, more consistent preparation, or more consistent motivation on the part of the very best athletes (Hopkins and Hewson, 2001; Stewart and Hopkins, 2001; Pyne et al., 2004; Paton and Hopkins, 2005). I favor the last of these possible explanations for endurance athletes: an athlete who realizes early on that s/he is not in the medal stakes must surely sometimes put less effort into the rest of the race. The situation is less clear in the field events, owing to the uncertainty in the estimates. More data are re-

quired before one seeks explanations for what may be more variability with top-half athletes in the throwing events.

### Conclusions

The main purpose of this study was to obtain estimates of the smallest worthwhile change in performance for elite athletes in each of the track-and-field events. Halving the variability of performance of the best athletes in each event provides such estimates. Coaches and sport scientists should therefore focus on enhancements of as little as 0.3-0.5% for elite track athletes through 0.9-1.5% for elite field athletes.

### References

- Gurland J, Tripathi RC (1971). A simple approximation for unbiased estimation of the standard deviation. *American Statistician* 25(4), 30-32
- Hopkins WG (2004). How to interpret changes in an athletic performance test. *Sportscience* 8, 1-7
- Hopkins WG, Hawley JA, Burke LM (1999). Design and analysis of research on sport performance enhancement. *Medicine and Science in Sports and Exercise* 31, 472-485
- Hopkins WG, Hewson DJ (2001). Variability of competitive performance of distance runners. *Medicine and Science in Sports and Exercise* 33, 1588-1592
- Hopkins WG, Schabert EJ, Hawley JA (2001). Reliability of power in physical performance tests. *Sports Medicine* 31, 211-234
- Paton CD, Hopkins WG (2005). Competitive performance of elite Olympic-distance triathletes: reliability and smallest worthwhile enhancement. *Sportscience* 9, 1-5
- Pyne D, Trewin C, Hopkins W (2004). Progression and variability of competitive performance of Olympic swimmers. *Journal of Sports Sciences* 22, 613-620
- Stewart AM, Hopkins WG (2000). Consistency of swimming performance within and between competitions. *Medicine and Science in Sports and Exercise* 32, 997-1001
- Stewart AM, Hopkins WG (2001). Seasonal training and performance of competitive swimmers. *Journal of Sports Sciences* 18, 873-884

Published Dec 2005

[©2005](#)