Measuring the Gender Differences in Athletics

November 13, 2007

Measuring the Gender Difference in Athletics

by Sven Gärderud, Carl-Erik Särndal and Ivar Söderlind

Summary. In this report we present measures of the gender difference in comparable events in athletics. Most of our analyses are based on world-wide all-time best individual performances as at the end of 2004. The progression over time is also examined, for the period 1986 to 2004. For comparison, we provide analogous computations for one particular country, namely, Sweden.

Only “comparable events” are considered. Events are comparable if the same rules apply for men and for women. Hence, non-comparable events include the throwing events (the weight of the equipment is different) and the hurdling events (the hurdle heights and/or the distances are different).

Two aspects examined in the report are of particular interest: (i) comparison of the current level of the gender difference between different events, and (ii) the development over time of the gender difference for a particular event.

The gender difference, as measured at the end of 2004, lies very near 10% in the sprints and increases in a continuous fashion as the distance increases, to attain around 14% in the longer running events. Compared with the sprints, high jump and long jump show distinctly larger values for the gender difference, situated in the 15% to 16% range.

The development over time of the gender difference is also analyzed. Events such as triple jump and pole vault are relatively new for the women, and they have made rapid progress. But in events with a long history for both men and women, our analysis suggests that the gender difference reaches a plateau and then remains remarkably constant over time. We believe it is unlikely that the women will ever “close the gap”.

There are more than one way to measure the gender difference. We explain our particular measure in the appendix at the end, and compare it with other possibilities.

1. Gender differences for comparable events as of year 2004.

The percentage differences between men and women in Table 1 below are based on an analysis of the world-wide all-time best individual performances in comparable events, as at the end of 2004. (For each athlete, only his or her best performance counts.) The 200 best individual performances are available yearly in the publication *Athletics, The International Track and Field Annual.*

Behind these statistics lie thousands of results achieved and recorded over many years, in the various events, for men as well as for women. This guarantees a good stability for the kind of analysis that we undertake, except for the events that are relatively new to women, such as triple jump and pole vault.

The patterns are very clear. Short running events show similar values of the difference, as one would expect. Similarly, the longer running events give similar values. High jump and long jump are also similar. Despite some minor statistical uncertainty, we believe that the values presented in Table 1 have a probable error of less than one percentage point, as measured in 2004.

With the exception for the triple jump, the pole vault and perhaps also the longer running events, we believe that future years will witness only small changes in the gender differences in Table 1. This is supported by our analysis in Table 2. The exceptionally large difference for pole vault in Table 1, 23%, will no doubt decrease over time, as this event “approaches maturity” on the part of the women.

The entries in Table1 are rounded values of a measure that we call Diff M/W. The computation of this measure is explained in the appendix at the end. It can be described as a type of average based on the 51 all-time best performances in the events analyzed.

With the exception for the pole vault, the gender differences in Table 1 lie between 10% and 16%.. Jumping events show markedly higher differences than running events. As for the running events, the difference is smaller in the short distances: 10% for 100meters/200 meters, 12% for 400m/800 m, and reaching about 14% for the long distances.

Table 1. Percentage difference between men and women in comparable events. The table entries are rounded values of the measure Diff M/W explained in the appendix and based on the 51 world-wide all-time best performances as at the end of 2004.

Event(s)	Difference
100 m, 200 m	10 %
400 m, 800 m	12 %
1500 m, 3000 m	13 %
5000 m, 10000 m	14 %
Half Marathon, Marathon	13 %
High jump	15 %
Long jump	16 %
Triple jump	16 %
Pole vault	23 %

Source: Athletics, The International Track and Field Annual 2005 (200 all-time best).

That the shorter distances show lower values of Diff M/W than the longer distances may be linked to history and tradition: Short distance running has a longer history for women than long distance running. To illustrate, the 100 meters and the 800 meters (but none of the longer distances) were on the women's program already at the 1928 Amsterdam Olympics. High jump and long jump also have long tradition for women, but the difference is more pronounced, at 15% and 16%, respectively.

This begs the question why events are so different. For example, the difference is six percentage points higher in broad jump than in the sprints, and five percentage points higher in high jump than in the sprints. Why this very significant discrepancy?

One possible explanation is that muscular mass and other physiological factors have more of an impact in the jumps than in the sprints. If this reasoning holds, it is likely that the gender difference is considerably greater in the throwing events, but since these are non-comparable events, conclusions are not possible.

That the difference is relatively large in events that are new for women is not surprising. Those events have not yet had the time to "mature" for the women. In particular, the difference of 23% in the pole vault is expected to decrease markedly in the near future. The difference is somewhat smaller in the half marathon and the marathon than in the 5000 meters and the 10000 meters. This, too, is not entirely surprising, considering that both the 5000 and the 10000 are very new for the women in olympic and other international competitions, whereas rich opportunities for women to participate in marathons (full or half) have existed for a longer time.

As mentioned, the values in Table 1 were computed on the basis of an elite group of athletes, defined more precisely by the 51 all-time best performers, as specified in Athletics 2005. An alternative way of computing the difference that some might consider preferable would be to base it on the current world records, the one for men compared with the one for women. We consider this to be a less representative measure, because particular (or random) circumstances will always to some extent dictate the current level of the world record. At every point in time, the current record in a given event may show clear signs of being "unbelievably good" or "decidedly sub-standard". Our method of computation, based on an elite group of athletes, gives much better reliability. However, for comparison, we also show, in Section 5, the gender difference as computed on the basis of the world records at the end of year 2004.

2. Progression over time of the gender difference.

We also studied the development the gender difference over the period 1986 to 2004. The values of Diff M/W in Table 2 are based on all-time best performances at the end of each of 1986, 1994, 1998 and 2004. This gives an interesting perspective on the role of the time factor.

Table 2 shows that the difference has remained remarkably constant in events with a long tradition both for men and women. The change is small or even negligible. By "small change" we mean here that the range of variation (largest minus smallest value of the difference) of the four years is at most one percentage point.

Tabell 2. Percentage gender difference Diff M/W in comparable events based on all-time best performances up until and including the following years: 1986, 1994, 1998 and 2004. Range = highest minus lowest of specified values Diff M/W for each event. NS (not specified) signifies insufficient data for women.

Event	1986	1994	1998	2004	Range
100 m	10.05	9.64	9.42	9.69	0.63
200 m	10.88	10.45	10.38	10.69	0.50
400 m	NS	11.80	11.98	11.74	0.24
800 m	12.19	12.16	12.39	12.52	0.36
1500 m	11.99	12.12	12.18	12.54	0.55
3000 m	12.39	12.20	13.23	13.01	1.03
5000 m	16.22	15.26	14.18	13.74	2.48
10 000 m	16.58	14.67	14.31	13.87	2.71
Half Marathon	NS	14.53	14.37	13.62	0.91
Marathon	16.26	13.97	14.06	12.82	3.44
Long jump	15.91	15.99	16.00	15.94	0.09
High jump	15.59	15.61	15.25	14.83	0.78
Pole vault	NS	NS	28.73	23.31	5.42
Triple jump	NS	19.40	17.67	16.39	3.01

Sources: Athletics, The International Track and Field Annual, years 1987, 1995, 1999 and 2005. For each year, the 51 all-time best performances were used in the computation.

The range (largest minus smallest value of Diff M/W) is exceptionally small for broad jump: 16.00 – 15.91 = 0.06, a remarkable stability over time. Other events with a small range over the period studied are 400 meter (0.24), 800 meters (0.36), 200 meters (0.50), 100 meters (0.63), high jump (0.78) half marathon (0.91). For the 3000 meters the range is only slightly higher (1.03).

By contrast, Table 2 shows that Diff M/W has diminished remarkably over the 18 years in events such as the pole vault (range = 5.42), the marathon (3.44), the triple jump (3.02), the 10000 meters (2.71) and the 5000 meters (2.48). All of these are relatively new events for women.

The expected pattern is that Diff M/W will decrease as an event approaches the maturity level, and then remain more or less constant over time (see also Section 4). This pattern is found in the shorter running events (100 meters and up to and including 1500 meters), as well as in the high jump and the long jump.

It will no doubt take a few more years before the longer distances (5000 meters and up), the pole vault and the triple jump will have reached the maturity level (the almost constant pattern over time).

The 5000 meters, the 10000 meters and the marathon show a remarkable similarity in their development over time. For all three, Diff M/W has dropped from over 16% in 1986 to below 14% in 2004. In 18 years, the women have “come closer” by 2.5% or more. This is not unexpected, given that all three events are relatively recent for women.

The triple jump and the pole vault are in rapid development for women. Table 2 shows that Diff M/W has shrunk in the triple jump from 19.40% in 1994 to 16.39% in 2004. In only ten years the difference has decreased 3.01%. Again, this is not surprising; triple jump is an even newer event on the women’s international program than the 5000 meters.

For events with a long tradition (the shorter running events, the high jump and the long jump), Diff M/W has thus remained almost constant over the time period studied. However, this does not imply an absence of development in these events. On the contrary, both genders show a significant advancement over the 18 year period.

As a measure of the average level, let us take the mean of the performances ranked 19, 20 and 21, for men and for women separately, and let us compute the percentage difference between these averages. Then, in the long jump, the men advanced from 833 cm in 1986 to 848 cm in 2004, an improvement of 1.80%; the women advanced from 702 cm in 1986 to 716 cm in 2004, an improvement of 1.99%. Thus both sexes advanced considerably, by nearly 2% each. But in that 18 year period, the women show no tendency at all to “close the gap”: Diff M/W equals 15.91% in 1986 and is almost exactly the same, or 15.94%, in 2004.

As a second example, consider the high jump. For the men, the mean of performances rank 19 to 21 advanced from 234 cm in 1986 to 237 cm in 2004, an improvement of 1.28%; the women advanced from 198 cm in 1986 to 202 cm in 2004, an improvement of 2.02%. While both men and women advanced, the women came only slightly closer: Diff M/W was 15.59% in 1986 as compared to 14.83% in 2004.

These results show that even if the level of the performances continues to progress, both for men and for women, Diff M/W will change only marginally or not at all in the events with long tradition. When an event has reached “the mature state”, advancement on the part of the men is matched by a similar advancement for the women, and Diff M/W remains essentially unchanged. The women are not “closing the gap”. We believe that this pattern of “constant difference” is likely to continue in the foreseeable future.

3. Diff M/W for Sweden, as compared with the whole world

Diff M/W for a given country may be considerably different from Diff M/W based on data for the whole world. To illustrate this, we computed Diff M/W for Sweden in the comparable events.

Table 3 shows Diff M/W for Sweden, computed on Swedish all-time best as at the end of 1997 (for men) and as at the end of 1999 (for the women). For the Marathon and the Half Marathon, it is not meaningful to compute a gender difference, because of a shortage of registered results for the women. The column World 1998 reproduces the values of Diff M/W given in Table 2, column 1998.

Table 3. Diff M/W in comparable events, for Sweden and for the world, toward the end of the 1990's. Separation = the value for Sweden minus the value for the world.

Event	Sweden 1997/1999	World 1998	Separation
100 m	12.13	9.42	2.71
200 m	13.31	10.38	2.93
400 m	15.02	11.98	3.04
800 m	16.16	12.39	3.78
1500 m	17.17	12.18	4.98
3000 m	16.95	13.23	3.72
5000 m	19.79	14.18	5.61
10 000 m	19.89	14.31	5.58
Half Marathon	-	14.53	-
Marathon	-	14.06	-
Long jump	18.89	16.00	2.89
High jump	15.76	15.25	0.51
Pole vault	39.58	28.73	10.85
Triple jump	20.29	17.67	2.62

Sources: Athletics, The International Track and Field Annual 1999 (200 all-time best); Bengt Holmberg (1998): Sverigebästa genom tiderna 1880-1997; Ove Fröberg (2000): Genomtidernastatistik för damfriidrott.

For almost all events, the Swedish value of Diff M/W is greater, even much greater, than the corresponding value for the world. The two values are close only in the high jump: 15.76% for Sweden, as compared to 15.25 % for the world. The next closest comparison occurs in the 100 meters, with 12.13% for Sweden, and 9.42% for the world, but here already, there is a large separation (defined as Diff M/W for Sweden minus Diff M/W for the world). The largest separation occurs in the pole vault: 39.56% for Sweden, as against 28.73 for the world; however, both numbers are likely to diminish considerably in the next couple of decades.

In almost all events, Diff M/W is thus considerably higher in Sweden than in the world. Why is this so? We believe that part of the explanation is the tradition and the popularity that an event enjoys in a specific country.

When a given event enjoys a long tradition in a certain country, with high levels of participation, then many results (and good results) are recorded, and there is success in international competition. We expect the measure Diff M/W for that country to be close to Diff M/W for the whole world.

We believe that high jump in Sweden is an example of this. For a long time, the high jump has enjoyed great popularity in Sweden, among men as well as among women. There has been strong performance in international competition. Not surprisingly then, high jump shows a small separation between Diff M/W for Sweden and Diff M/W for the whole world.

One could expect that other examples of the same kind might be, for example, the javelin throw in Finland, or the sprints in the USA. We have no data to verify this; besides, the javelin is not among the comparable events because of difference in equipment.

On the other hand, the values on Diff M/W in Table 3 indicate that the Swedish women are conspicuously far behind their male compatriots in all events other than the high jump. In the running events the separation ranges from around 3 percentage points (short distances) to around 6 percentage points (longer distances).

4. Will the women “close the gap”?

It is probable that in the events with a long tradition the women might come closer to the men, but only marginally closer. But judging by our results in Table 2, we consider it unlikely that the women will ever “catch up” and completely close the gap, neither in the newer events nor in those with long tradition.

A strong indication of this is the fact that in events with “mature status”, such as the 200 meters, the 800 meters, the high jump and the long jump, there is no clear evidence of a “closing the gap”, during the 18 year period covered by Table 2.

“Coming closer” is one thing, “closing the gap” is quite another. A “coming closer” is normal in an event where the level for women has not yet matured. Table 2 shows that this is happening in relatively new events such as the 5000 meters, the 10000 meters and the marathon. This tendency is even more clearly pronounced in the pole vault and the triple jump. However, in all these events, we expect the coming closer to reach a halt, and from then on, Diff M/W will stabilize at an essentially constant level. That the women will completely close the gap is unlikely.

One may wonder what effect, if any, that doping may have on our computed figures. Has its effect been greater for the women than for the men and in the short races in particular? Is the fact that the gender difference appears constant over time partly a result of a certain leveling off of the performances since doping controls became more rigorous and extensive in the middle 1980’s? The throwing events in particular might be examples of this, but we have no data to substantiate this contention.

5. Measuring the gender difference through current world records.

As mentioned earlier, we computed the measure Diff M/W on the basis of the 51 best all-time individual performances in a given event. The resulting values, given in Table 2 and reproduced as column B in Table 4, have good statistical stability. By contrast, a measure based in the world records does not have the same stability. Column A in Table 4 shows the percentage gender difference computed on the respective world records as at the end of 2004.

Table 4. Percentage difference between men and women measured in two ways: (A) on the basis of the world records as at the end of 2004, and (B) by Diff M/W, that is, on the basis of the 51 all-time best individual performances as at the end of 2004.

Event	A	B
100 m	7.4	9.7
200 m	10.5	10.7
400 m	10.2	11.7
800 m	12.0	12.5
1500 m	11.9	12.5
3000 m	10.3	13.0
5000 m	14.2	13.7
10 000 m	12.3	13.9
Half Marathon	13.3	13.6
Marathon	8.4	12.8
High jump	14.7	14.8
Pole vault	18.4	23.3
Long jump	16.0	15.9
Triple jump	15.3	16.4

In Table 4, let us compare the 100 meters with the 200 meters. The two events being similar, we expect to find roughly the same value for the difference. This expectation is confirmed in Column B (based on “the best 51”), but is notoriously absent in Column A (based on the world records). For the 100 meters, column A gives a value as low as 7.4%, due to the truly exceptional women’s record of 10.49 seconds. But 7.4% is in total disagreement with the column B value of close to the 10% that we consider to be “a correct indication”.

In the 200 meters, both the men’s record, 19.32 seconds, and the women’s’ record, 21.34 seconds, are “extremely good”, and it so happens that the value 10.5% in column A agrees closely with the 10% that we consider to be a “correct indication”.

In a few of the other events we also get noteworthy differences between columns A and B. The most striking of these is the marathon: Column A gives the value 8.4%, which is

truly abnormal, considering the column B value of 13% that we consider to be “a correct indication”.

On the other hand, columns A and B happen to agree quite well in events such as the 800 meters, the high jump and the broad jump.

Appendix. Explaining the methodology.

We computed the measure called Diff M/W in three steps as follows:

(1) compute five relatives, denoted R9-11, R19-21, R29-31, R39-41 and R49-51;

(2) eliminate the largest and the smallest of the five relatives;

(3) compute Diff M/W as the average of the remaining three relatives.

The relatives are computed as follows: R9-11 equals the difference (in %) between the average of the performances number 9, 10 and 11 for men and the corresponding average for the women. This difference is computed in percent of the men’s value and taken with the positive sign. The other four relatives are computed in an analogous manner. The reason for averaging three performances for each relative is to help smoothing the random effects. The values presented in Table 1 are rounded values of Diff M/W.

Details of the computation of Diff M/W are illustrated in Table 5, for selected comparable events. Table 5 shows that the five relatives differ very little for most events. This is a sign of good stability, but there are exceptions, such as the triple jump. The five relatives usually show a slight upward tendency, from R9-11 to R49-51. The stability is good if the range (largest minus lowest) of the five relatives is small.

As Table 5 illustrates, the first four events have excellent stability, with a range of less than one percentage point. The long jump is also quite good, with a range of just over one percentage point. By contrast, the range of 2.02 for the triple jump is higher than we would like, but it only illustrates the fact that this event is new for women.

We can accept that the relatives vary to some extent, with a range of up to one percentage point. The large quantity of data behind the world-wide all-time best performances guarantees solid stability. Thousands of results have been recorded and registered over many years; the best 200 are shown, separately for men and for women, in the publication *Athletics, The International Track and Field Annual*. Stability is expected to be considerably lower (the range much higher) in events where the women’s international participation is of a relatively recent date, such as the triple jump and the pole vault.

Table 5. Details of the computation of Diff M/W for selected events, based on the 51 world-wide all-time best performances at the end of 2004. Diff M/W = mean of the three central relatives. Range = largest minus smallest of the five relatives. All values in %.

Event	Diff M/W	Largest relative	Smallest relative	Range
100 m	9.69	9.82	9.38	0.44
800 m	12.52	12.70	12.37	0.33
5000 m	13.74	13.88	13.22	0.66
High jump	14.83	15.25	14.64	0.61
Long jump	15.94	16.17	15.12	1.05
Triple jump	16.39	17.09	15,07	2.02

The gender difference may be measured in alternative ways. We chose the measure Diff M/W, calculated in the manner just explained. In order to test its reliability, we compared it with the results obtained by two other methods that are also intuitively sound: the "mean difference method" and the "median difference method", now to be explained. All three use all-time best performances as published in Athletics 2005.

In the mean difference method, we compute, for a given event, and separately for men and for women, the mean of 41 performances, starting with the performance ranked number 10 and ending with the one ranked number 50; we then compute the percentage difference between the two means.

In the median difference method, we focus on the result of rank 30, the one for men and the one for women, and we compute the percentage difference between those two median values.

The three methods (the Diff M/W method, the mean difference method, the median difference method) gave very similar values for the gender difference. The three values lie in most cases within a range of less than one tenth of a percentage point. Recalling that the values in Table 1 are rounded values of Diff M/W, we consider that, to the degree of approximation that we use, all three methods are essentially equivalent, and that Diff M/W is a valid measure.

References

Athletics 1987, The International Track and Field Annual. SportsBooks Ltd, Cheltenham.

Athletics 1995, The International Track and Field Annual, SportsBooks Ltd, Cheltenham.
.

Athletics 1999, The International Track and Field Annual, SportsBooks Ltd, Cheltenham.
.

Athletics 2005, The International Track and Field Annual, SportsBooks Ltd, Cheltenham.
.

Holmberg, Bengt (1998): Sverigebästa genom tiderna, All-time list – men juniors and youth. Aknanov AB, Sundbyberg, Sweden.

Lenz, Göran (red.) and Fröberg, Ove (statistics) (2000) Genomtidernastatistik för svensk damfriidrott. Gordons förlag, Lund, Sweden.

About the authors**.**

Sven Gärderud, born in 1921, grew up in the Stockholm suburb of Råsunda. He worked as a security systems expert with Swedish enterprises AGA and Bofors in Lidingö. He has had a life-long interest in statistics. In the early 1980's, he started the yearly publications of veterans' track and field statistics of the Swedish Track and Field Association. His preferred sports activity is orienteering. Since 1938 he has participated in more than 1950 orienteering races, and has won many of them. He also participated in cross country (Swedish team champion in 1943) and in the 3000 meters steeplechase (best performance 9.21.2 in 1941). He resides in Lidingö, Sweden, since 1955.

Carl-Erik Särndal, born in 1937, grew up in the Swedish province of Västergötland. He was professor of statistics in Montreal, Canada, and at the Swedish national statistical agency, SCB. He has long taken an interest in the regularities observable in track and field performance. He competed in high jump in his 20's, with 201 cm as his best result in 1963. He began taking part in masters' competition in 2003 and currently holds the world record in the M70 category with 159 cm and the European record in the M65 category with 161 cm. He resides in Ottawa, Canada.

Ivar Söderlind, born in 1944, grew up in Lycksele in northern Sweden. He teaches Sociology at the University of Umeå, Sweden, where he specializes in sociological method, especially for surveys carried out by mail or by interview. He is the president of the Swedish Athletic Association's masters' committee. Since 1988 he is the chief responsible for statistics of the European masters' track and field organization, EVAA. He has taken a passionate interest in track and field athletics since his early years. At 17, he cleared 170 cm in the high jump. He is interested in all forms of sport, but in particular in track and field, cross country skiing and biathlon. He resides in Umeå, Sweden, since 1962.

The contents of this article are protected by the copyright rules. Copying is permitted for private use, while citing the source where appropriate. Should you wish to distribute the contents of this article to a wider group of persons, please contact Ivar Söderlind, ivar.soderlind@soc.umu.se, phone +4690-7865457.