

---

# INTRODUCTION

Racing Research is a collateral form and time ratings service for British and American horse-racing. We rate every performance by every horse in every race in Britain or America, using a computer handicapping system which has been continuously developed over 30 years. This is a collection of pieces taken from past Flat annuals. We hope they serve to give newer subscribers some idea of the Racing Research approach and methodology, without venturing into any real technical depth.

---

# FOREWORD FROM THE 1984 ANNUAL

## Ratings and their derivation

To start with, let's establish what we mean here by a 'rating'. We mean a figure which measures a performance of a horse in a race. Because horses are handicapped by weight the natural measuring scale is weight and because this is Britain the units are pounds. This is usually taken for granted but it's worth remarking that if horses were handicapped like professional runners with different distances of start then it would not be so appropriate to rate their performances on a pounds scale.

In the course of its career a horse acquires a set of ratings generally achieved over different distances and ground conditions at a variety of racecourses. Taken together these give a complete picture of a horse's characteristics and capabilities. However it is traditional to regard a horse's best rating as representing its ability and to refer to this as the rating of the horse. This custom can be misleading. It is easy to show why. Imagine two horses of the same age, 'Old Reliable' with performance ratings of 79, 81, 80 and 81 and 'Spasmodic' with performance ratings of 69, 63, 81 and 65, achieved under similar conditions. Traditionally both horses would be regarded as rated at 81, their best performance rating. So if they were matched against each other at level weights and their ratings used as a guide to the likely outcome each would seem to have an equal chance of winning. But it is apparent from looking at the full sequence of their performance ratings that, far from being an even-money chance, Old Reliable should be at long odds-on with Spasmodic at odds-against. Anyone taking even-money about Spasmodic would have a bad bet. Their best ratings alone definitely do not reflect the horses' chances.

A general point about ratings should be emphasised. Once a rating has been assigned to the winner of the race, then ratings for the beaten horses follow automatically from that by simple calculations based on weights carried, distances beaten and immaturity or weight-for-age allowances which are explained later. So the rating of the winner sets the level of the race. In the rest of this section the word 'rating' will be used interchangeably. As well as its primary sense of performance rating, it will also be used to refer to the rating of a horse in the traditional sense and also refer to the rating of a race, meaning its level as described above. The

intended meaning should be clear from the context.

So, knowing what we mean by a rating, how do we go about assigning ratings to performances? Discounting hocus-pocus methods, there are two basic approaches, collateral form, usually referred to just as 'form', and time.

### Time ratings

At first sight it might appear that the most obvious way of rating a race would be 'on time'. Without going into too much detail we can indicate the considerations involved in the calculation of time ratings or timefigures as they are usually known. To produce timefigures for a race there are four basic requirements. They are: (1) an accurate time for the race, (2) an accurate standard time for that distance at that racecourse, (3) an accurate estimate of the effect of the going on the race time, generally referred to as the going allowance, and (4) an accurate estimate of the effect of the wind on the race time. And incidentally, by 'estimate' we mean a mathematically-based estimate not a 'considered opinion' or 'an educated guess'. We will deal these factors one by one.

The race time: sixteen racecourses are covered by Racecourse Technical Services' electrical timing, the other twenty make do with hand times. Electrical times are accurate, hand times taken by a competent watch-holder can be relied on to within one fifth of a second which is equivalent to 4-5 lb in a five-furlong race. Not all published hand times are reliable to this degree of accuracy, as can often be seen by the discrepancies, larger than this, between the different times published for the same race.

The standard time: The main step in timefigure calculation is to compare the race time with the corresponding standard time. There is no definitive 'standard time'; different versions of timefigures use different methods to calculate them and end up with different standard times. The calculation of proper standard times is a small subject in itself. Suffice it to say that the calculation of a particular standard time should take into account every time, and its attendant circumstances, recorded over that course and distance for several years previously. Accurate standard times are the foundation of good timefigures.

The going allowance: first of all, the actual times are compared to the standard times, not individually but as a whole, and the overall average difference per furlong is obtained. This difference will be caused by the going, the wind and, naturally, the capabilities of the runners themselves and the weights they carried. So it then has to be decided what part of this

difference is due to the ability of the horses, so that it can be taken out together with the wind effect, leaving the effect of the going alone, which is the going allowance. This need for a prior assessment of the horses' abilities stops timefigures from being a genuinely independent measure of performance. The going allowance and so the timefigures themselves are inextricably tied to the initial assessment of the day's runners which must be obtained in some other way.

There is an interesting point about this prior estimate of abilities. Weeks or even months later, when more racing has taken place, the estimate is often seen to be some way out, given the benefit of hindsight. Ideally then, the original estimate should be revised causing the going allowance to be modified and consequently the timefigures too. At Racing Research we do this, for every time the computer reassesses the form ratings the going allowances are also recomputed and the timefigures correspondingly changed. So our timefigures are 'fluid'. But in no other timefigure system is this done. Traditionally a timefigure, once calculated, is enshrined for posterity, never to be changed. It is static and that is a definite weakness.

There is one other problem with going allowances. Uneven artificial watering or uneven drainage after rain can cause the going to vary on opposite sides of the course or on different sections. When this is known to have happened the day's races must be split into groups which took place over similar ground (when it has happened but is not known to have happened then the timefigures will be up the creek). Each group is treated as if it were a separate meeting. The analysis then is based on a much smaller amount of data, only two to four races instead of the six or seven of a complete meeting, and consequently the timefigures will be far less reliable. Certain courses are notorious for this effect - Haydock for one. The sprint courses at Sandown and Kempton, entirely separated from the main tracks, are frequently beyond the pale from the point of view of timefigure production.

The effect of the wind: On most days in Britain the wind influences race times significantly so corrections for its effect must generally be made. Unlike the going allowance, these corrections can not be expressed as a fixed amount per furlong. A wind can speed up some races and slow down others, for example at York a wind directly behind the runners in the straight would tend to reduce the times taken for five-furlong races but increase the times of races run over two miles. Wind corrections are normally based on the compromise assumption that the wind is uniform in strength and direction throughout the day's racing. In reality the wind is not often uniform for three hours at a time but includes rises and falls in

strength and variations in directions. So there is some scope for error in making wind corrections – but a sight more scope for error if they are omitted.

Those are the most important factors influencing timefigures. Still, there are one or two other points to look out for. Dolling out of the running rail to protect the turf around a bend increases the distance of all races which include that bend. Unless that this is recognised and compensated for – by calculating the distance added to each race and increasing the standards accordingly, then timefigures for these races will appear worse than they really are. At some courses the rail is moved in and out from meeting to meeting like a fiddler's elbow. There is also the problem of rain during racing which can cause the going to become progressively slower. On the other hand, as wet ground dries out the going will become progressively faster. These effects must be dealt with as accurately as possible.

On days of high winds, gusting strongly, or when heavy rain causes the ground conditions to change considerably during racing then the necessary compensating corrections cannot be calculated within acceptable limits of accuracy. When this stage is reached the production of timefigures should be abandoned.

By now the reader will have realised, if he was not already aware, that any system of timefigures contains plenty of scope for error because its constituent parts contain scope for error. The idea, sometimes advanced, that a horse 'has done 103 on time and that's accurate because the race time is accurate' is naive. Under perfect conditions – on a calm day, over a racecourse which has not been watered artificially or naturally during the previous week, with electrical timing, reasonable-sized fields and the races run generally over the same ground – so not Sandown or Kempton with their separate sprint courses; on a day like that at Newmarket in July say, properly-calculated timefigures could reasonably be expected to be accurate to within 4lb.

Apart from the difficulties described above in assessing the contributory variables accurately, there is one overriding fact which completely rules out timefigures as a universal rating system. That is, quite simply, that only about one in five of all races in Britain is truly-run throughout. So it is only in this 20% of all races that timefigures can accurately reflect the merits of the runners. In the other 80% the time value of the performances is below the form value of the performances.

Timefigures have been around for a long time in various forms, some of them extremely crudely produced. William Mason's 'Time Test' was one of the earliest ones and I expect there were others before him. But the first person to put timefigures on something like a proper mathematical basis was Phil Bull whose 'William K Temple Timefigure Service', a list of fast-timing horses-to-follow, updated weekly and with an end-of-season review, started in the late nineteen-thirties. Eventually after ten years of timefigures, Bull was faced with the shortcomings of time as a means of classifying all performances and about 1947 he was joined by Dick Whitford who had been producing his own manual handicap ratings since 1940. Bull renamed his business 'Timeform' and Whitford's handicap ratings became the first 'Timeform ratings' in 1947-8. A practical illustration of the limitations mentioned in the paragraph above. During 1974-5 I substantially modified Bull's techniques of timefigure calculation, putting them on a more sound mathematical basis. Subsequently I wrote computer programs to produce the figures more quickly, though if necessary the new techniques can still be performed by hand. As a matter of interest, the timefigure techniques we now use at Racing Research are considerably more advanced than these and could not be performed manually.

Having outlined the calculations of timefigures and indicated their unavoidable limitations, let's look at their particular usefulness. Principally, it is that they sometimes permit the rating of races where most of the runners are unraced or very lightly-raced, where form lines, if there are any, are tenuous. The commonest occurrence of these circumstances is in two-year-old races during the early part of the season. Time methods often provide the only way of rating such races immediately, though later in the season, when the runners have accumulated more form, these races can usually be retrospectively rerated more accurately on form. More generally, in any race where, for one reason or another, the form value is highly speculative, timefigures, if they are available, may assist the manual handicapper in reaching a decision. Though realistically he has no way of knowing whether the race has been completely truly-run, he can work on the assumption that the form ratings should be at least as high as the time ratings.

Before leaving timefigures, it's worth examining the oft-quoted timefigure dictum, 'bad horses can't record good timefigures' or words to that effect. This old saw is undoubtedly an oversimplification, it depends on what is meant by a bad horse. I could give instances of horses which have won Group 1 races yet whose best time performance during the course of their whole career did not come within 20lb of their best form performance. Are these bad horses? Conversely I could name horses with timefigures up to

classic-winning standard which, although given plenty of chances, never won a Group race. Are these top-class horses?

How do such apparent anomalies occur? Well the completely-equipped top-class racehorse should have two main attributes: (1) the ability to maintain top-class racing pace, for want of a better term, throughout a race and (2) the ability to accelerate - to quicken - towards the finish of a race. To record a top-class timefigure a horse must have the first attribute but not necessarily the second. On the other hand it is possible to win top-class races which are not strongly-run throughout with the second attribute but not a full measure of the first. But both qualities are needed to win a top-class race which is completely truly-run.

The same thing happens in athletics. A runner who is able to maintain a uniformly fast pace without having much in the way of acceleration can set world records but may never win the major championship events where better-equipped runners can use him as a pacemaker and cut him down over the last lap. Perhaps the classic example of this type of runner was the Australian Ron Clarke. A completely-equipped runner such as Sebastian Coe, who does possess finishing speed as well, is capable of setting world records and taking the major championships.

To summarise then, timefigures are useful primarily in assessing races where the horses are too lightly raced for much faith to be placed in a form assessment. But their method of calculation means they are critically dependent on many variables and under certain conditions the effect of these variables can not be estimated accurately enough for the timefigures to be reliable. Contrary to a widely-held belief they do not provide a genuinely independent measure of a performance, because a prior estimate of the day's runners is always needed to determine the going allowance. Finally, timefigures can only fully reflect the merit of performances in thoroughly truly-run races and as 80% of all races are not run that way timefigures are definitely not suitable as a means of classifying all performances.

The timefigures in this book are the most sophisticated to be published and should be of particular interest to time students. Even so, they are here primarily as a guide to the overall pace at which the race was run, not as an alternative to the ratings. The timefigures are expressed in pounds so they can be compared directly with the form ratings. In races where the early pace is slower than the runners are capable of, timefigures will fall short of form ratings. In a truly-run race, where the horses have been kept up to their work throughout, timefigures will correspond closely to form ratings. So in general the difference between form rating and timefigure

can be regarded as a measure of the relative pace of the race.

Timefigures will normally be more reliable when they are based on electrical times rather than on hand times. The courses with electrical timing are: Ascot, Ayr, Chester, Doncaster, Epsom, Goodwood, Haydock, Kempton, Lingfield, Newbury, Newcastle, Newmarket (both courses), Salisbury, Sandown and York.

### Form ratings

That leaves 'form' as the only basis for a universal rating system. Put as simply as possible, the principle here is to assign a level to a race so that the ratings are as consistent as possible with the ratings that the same runners have achieved in other races. This involves constantly reassessing the levels of all races as more race results arrive and cause modifications to the ideas of the horses' abilities. That's the idea anyway. Because of the thousands of millions of calculations and references required this ideal can be achieved only on a computer.

Nevertheless this process, known as 'handicapping' has traditionally been performed manually, in a very approximate way, and it still is by everyone but Racing Research. The first official handicapper was Admiral Rous in the 1850's and I assume he used much the same principles as are used by today's manual handicappers. Some readers will know how a manual handicapper works. As each new race result comes along he assigns a level to the race and so, implicitly, a rating to each horse in the race. The assignment of this level is usually based on his knowledge of the runners' previous performances. In the case of two-year-olds with little previous form he may perhaps be influenced by the levels he has recorded for the same race in previous seasons or possibly he may consider the race time as a guide. However, the calculation of worthwhile timefigures is a lengthy process and, as we have remarked, it's only in the truly-run races, about one in five, that timefigures are in line with form ratings. As for the dubious 'historical handicapping' technique of basing the race level on figures from previous seasons, well it's better than nothing but it does introduce an element of circularity which won't entirely disappear.

Digressing for a moment, this 'historical' or 'circular' handicapping syndrome can sometimes be detected in the ratings of high-class races where the handicapper is loth to rate horses outside the limits he considers acceptable for such races, even in the face of apparently strong evidence to the contrary. Thus he produces one more figure to support his preconceptions in future years. Instead of interpreting the facts as they



stand, it becomes a question of 'how would he have got on against Sir Ivor?' The overall level of all a season's ratings should be kept the same from year to year so that horses can be compared across generations but it does occasionally appear as if this principle is being applied to individual races. This has become particularly relevant now a greater proportion of the world's most valuable bloodstock ends up in British yards and, judging from the results of international races, the standard of the best horses in Britain is probably higher than for a long time. But this is not fully evident from the work of manual handicappers where the historical handicapping reflex is suppressing the true level of these horses' performances.

To continue, the handicapper has now settled on his ratings for the new race. In the light of the evidence provided by this race the conscientious handicapper may revise his assessment of one or two other races which concern some of the runners, but that's as far as he goes. As these other races have their levels altered then races concerning runners from them should be reviewed and so on, making a 'ripple' that spreads throughout practically all races run that season – two-year-olds and older horses considered separately of course. But the handicapper doesn't calculate the effect of this ripple - he can't as he's only a human unable to work fast enough, it would need the speed and precision of a computer to do it. And that's what happens in our computer analysis of racing form. The computer can 'think' and calculate fast enough and has a completely accurate memory for every detail of the circumstances of every performance throughout the season.

There are some funny ideas around about manual handicapping, it seems to inspire an aura of mystery, rather like astrology does to some people. I was once discussing handicapping with a well-known racing person who was clearly a member of the reverential 'handicapping is an art and art can not be performed by a computer' school of thought. Attempting to show that manual handicapping was not the perfect answer to the problem of rating performances I eventually said 'Just have a look each year at the ratings given to the leading horses – the ones on which a manual handicapper spends a disproportionate amount of his time – by the commercial ratings firms and by the official handicappers. Look at the wide variations of opinion.' 'Yes' came the immediate reply, 'but one of them has usually got it right!' Not the same one each year unfortunately.

Still, I suppose that's what they mean about it being an art. Different handicappers see the same thing differently – one's an 'Impressionist', one's a 'Cubist', one's going through a 'blue period' and so on. Seriously though, the fact that an activity is conducted unscientifically does not

automatically raise it to the status of an art. Manual handicapping calls for a high degree of skill and expertise and some handicappers possess these qualities in abundance. But that does not change the fact that their task of relating collateral form as consistently as possible, can be performed more accurately by a computer.

### Computer Handicapping

The first electronic stored-program - the first real computer – appeared in 1946. Almost thirty years elapsed before computers began to make much impression on the horseracing community, which has seldom been in the vanguard of technical progress. In the early 1970s the Jockey Club and Weatherbys decided to make administrative life a bit easier by storing the official handicap ratings on a computer. The ratings were still produced by manual handicappers but then stored on a computer file and when a handicapper changed one of his ratings the new rating was transferred to the computer to replace the old one. So when the weights for future handicap races were to be set, the job was easily done by a simple computer program. A list of entries for a race was input to the program which was then referred to the file of official ratings. The program then worked out what amount needed adding to the rating of the highest rated entry to make it 9-7, or whatever top-weight was required. It then added the same amount to the ratings of all other entries, sorted the entries into descending order of weight and finally printed out the entries with their weights. A fairly trivial application by any standard.

When Britain switched from using the Julian calendar to the Gregorian calendar in 1753 it was necessary, at the changeover, to advance the date by eleven days and this caused rioting by people who believed that their lives had thereby been shortened by eleven days. The Jockey Club's changeover didn't fare quite so badly as that, nevertheless it was some time before everyone involved in racing could be convinced that the machine itself was not actually preparing the ratings. Since then one commercial ratings firm has followed the same procedure and they still manage to cause confusion by referring to their ratings as a computer handicap.

A real computer handicap is one where the complete handicapping process, as described earlier, is carried out solely by computer programs. There is only one such system at present, the one which produced the ratings in this book. The use of a computer also allows us to employ additional techniques which would not be accessible to a manual handicapper. Some race time considerations are incorporated too. I won't bore you with the usual mind-

boggling comparisons between the speed of computer operations and those of the human brain. Let's just say that our techniques, requiring thousands of millions of calculations and references to be made, all with complete accuracy, can be implemented only on a computer. To put it in perspective, a transcription of all the programs in our system would occupy three books the size of this one.

### Immaturity in racehorses and ratings

A two-year-old is like a sixteen-year-old human, not developed physically. A three-year-old is like a twenty-year-old human, generally with some development still to make. At the end of its fourth year a horse is officially considered to be fully mature though this is often doubtful for particular horses. So that young immature horses are able to compete on fair terms against older horses, the older ones can be made to carry more weight and for each distance at any time of the season the 'fair' relative weights are given in a 'weight-for-age' table. Such a table is based on the average rate of development of racehorses. Fully mature horses are usually assigned ten stone in these tables. Looking at our weight-for-age table we see that for a ten-furlong race during the first two weeks of August the set weight for a three-year-old is 9-5. So in August over ten furlongs the average three-year-old still has 9lb of development to make; we can say its immaturity is 9 lb.

When a horse's performance is 'rated' the rating is arrived at in the following way. First of all the performance is assessed, in some manner, to give a figure, in pounds, representing its actual value and then the appropriate 'average immaturity' figure is added in to give the rating. So, for example, if our three-year-old running over ten furlongs in August records a performance which we assess as genuinely worth 119 then we credit him with the appropriate average immaturity of 9 lb and give his performance a rating of 128. This procedure is the standard convention and it has a lot to recommend it, including the consequent ease of calculation of weight-adjusted figures for future races as described earlier, so we have observed it in the figures in this book. But it should be remembered that only mature horses are rated exactly as they have performed.

The fly in the ointment with this convention of including the average immaturity in rating is that, like humans, horses develop physically at widely differing rates. One horse might be a considerably better two-year-old than another but of similar ability as a three-year-old and actually worse as a four-year-old mature horse for no other reason than their differing rates of physical development. To illustrate what can happen let's take a

couple of hypothetical horses, a precocious type called 'Sharpsort' and the late-developing 'Scope', and suppose for simplicity that every year their best performances are over five furlongs in the first half of August. As two-year-olds their 'true' performance figures are 110 and 95, as three-year-olds 116 and 119 and as four-year-olds 120 and 130. Our weight-for-age table tells us that of five furlongs in the first half of August an average two-year-old is 25 lb immature, a three-year-old 6 lb immature and a four-year-old is fully developed. Accordingly, to get each year's ratings we add 25 to the two-year-old true figures, 6 to the three-year-old true figures and nothing to the four-year-old true figures to give the following pairs of ratings each year: 135, 120; 122, 125; 120, 130. So Sharpsort has acquired a career-best rating of 135 by virtue of his 110 as a two-year-old though he's never actually performed better than the 120 achieved when he was four. 'Scope' has a best rating of 130 and also a best performance of 130. Scope is undeniably the better horse by 10 lb. Nevertheless I know a lot of people are misled by the respective best ratings of 135 and 130 into believing that Sharpsort is in some sense a 5 lb better horse than Scope: they confuse ratings with performances actually achieved. Some shrewd operators are well aware of the wide extent of this misunderstanding and capitalise on it by retiring their early-maturing two and three-year-olds to stud before their literally inflated reputations can be tested truly against the slower-maturing, but perhaps intrinsically superior members of their age-group. This is a wise policy from their point of view because the best two-year-old hardly ever makes the best three-year-old and the best three-year-old frequently doesn't turn out to be the best four-year-old – assuming they stay in training of course. At two and three precocity reigns.

Over the years there have been many examples of outstanding two-year-olds who, for one reason or another, failed to make the regulation development. These horses would probably be rated about 135 as two-year-olds. But these 135 ratings would include at least 20 lb worth of immaturity, scope for development they may never have possessed, so that their proven ability was nearer the 115 mark, they never performed to any higher level, they never were top-class racehorses. It's practically impossible for a two-year-old to be a top-class racehorse for the same reason that a sixteen-year-old human won't be a top-class athlete.

One reason why the weight-for-age scale can only be regarded as a general guide is that horses in the same age-group are really not the same age at all because their foaling dates vary so widely. A horse born in January will appear precocious alongside one born in June even though they were both granted the same ability at birth and then developed through to maturity at the normal rate. The gap in performance will narrow continually throughout

their years of immaturity but will not close completely until the younger one reaches maturity. So precocity, or lack of it, has to be viewed in the light of the foaling date. Once this is understood, the practical use of foaling dates becomes apparent. For example if a pair of two-year-olds recorded similar best figures, both in October, then, other considerations apart, the better bet as a three-year-old would be the one with the later foaling date.

The overall level of the ratings in Computer Racing Form will remain constant each year so that comparisons between horses from different periods are valid. Incidentally, maintaining the level in this way requires some fairly sophisticated analysis as not all performances are precisely rated – it's not enough just to keep the average of the rated ones constant from year to year. The level we have chosen makes the average Group I winning performance about 126 or 9 stone though in years with an excess or a shortage of top-class horses this particular figure, though not the overall level, would vary accordingly.

---

## ASSESSING JOCKEYS

Some people believe they can accurately judge jockeys' abilities by watching races. Ask them who the best jockeys are and you'll get a ready answer. Ask how they arrive at their assessments and it starts to become more difficult. The jockey's style, particularly in a finish, is usually the first thing to be cited. But how can they tell that horses actually run better for the stylish jockey? For as anyone who takes an interest in other sports, or life itself will know, style alone can be a most unreliable guide to performance. 'Well, to the experienced eye...' Quite. By this stage the answer has degenerated into a dignified appeal to respect their own or others' expertise with liberal recourse to phrases such as 'respected race-readers', 'generally recognised', 'racing's professionals', 'racing's insiders', and so on. In short, waffle. Over the years I have encountered many people confident of their ability to rate jockeys accurately. Not one been able to demonstrate the semblance of any method which would stand up to the most elementary examination. Which is why jockeys' reputations are often as flimsy as the Emperor's New Clothes.

In most sports the competitors meet on equal terms using equipment, if any, which is more or less uniform. So the result of the competition is generally an accurate reflection of the contestants' own performances. In horseracing, on the other hand, if the riders are viewed as competitors then their 'equipment', the horse, is certainly not uniform. The result of the race is principally due to the performance of the horse with the rider playing only a minor part. This frequently leads to false conclusions about jockeys' abilities. Number of wins, number of big-race wins, percentage of wins to rides and prize money totals don't rate a jockey's riding ability. All depend overwhelmingly on the quantity and quality of his mounts.

Brough Scott likens jockeys to card players in that the good ones are distinguished by their consistent ability to do the right things at the right times. It's an apt analogy of jockeys themselves but if the race is imagined as a card game, with the horses as hands, then the deal is definitely rigged, for the fashionable jockeys get the best of it. Any attempt to assess jockeys' performances without allowing for that is simplistic.

The only logical way to measure a rider's effectiveness is to examine how horses have performed for him compared with their performances for other jockeys. That is what we do. The analysis is confined to riders for whom there is a large amount of data and this data is analysed in a mathematically correct way by computer programs. The resulting ratings, given in pounds to the nearest tenth, are therefore an accurate reflection of each jockey's effectiveness over the season. In particular, if a rider is the best jockey then horses in general should run better for him than for other jockeys. And if they do he would appear at the top of our ratings.

This method exposes what the rider is really contributing to his mount's performance. It reveals the unfashionable jockey getting more out of poor mounts than other riders have done. The racing public don't recognise this because in most cases the horse still finishes down the field and only the computer, analysing every one of the thirty-thousand plus performances each season, notices that it has run better for this jockey. At the other extreme it pinpoints the big-name jockey who is not riding as well as he might be, but, with plenty of short-priced fancied mounts, is still winning a lot of races. He will lose the odd race that a top-class jockey probably would not have lost where a pound or two makes the difference between victory and defeat but most races aren't like that and the discrepancy between his real and imagined performances isn't noticed – except by the meticulous computer analysis.

Since they first appeared our jockey ratings have frequently been at variance with public opinion. In particular, some of the younger riders perform rather better than is generally imagined and some of the older ones not quite so well as their public reputations would indicate. It's interesting to see why this should happen as it illustrates the whole problem of assessing jockeys objectively.

First of all the difference between a top-class rider and an average one is about 3 lb, as can be seen from our ratings. This is insignificant when compared to the difference in pounds, between a mount with a good chance and a mount with an average chance, in most races. Look at any set of pre-race weight-adjusted ratings, to confirm that. Or putting it another way, convert into pounds the distance between the winner and the

horses finishing in mid-division at the end of the average race. It's apparent that the quality of the mount is far more influential than the ability of the rider.

This has obvious consequences. Without good mounts no jockey can ride many winners. If his riding is 2 lb better than the average, it will at most make a difference of five winners in thirty. Not the sort of figure to make him instantly appreciated. However, once a rider has established a supply of good mounts then, so long as he can ride to an average standard, he will continue to ride plenty of winners and retain his reputation and patronage. So it is quite normal for a jockey's reputation to lag some way behind the true value of his riding, both on the way up and on the way down. A rider may really be top-class long before he is accorded such a reputation and, later on, his true riding prowess may decline long before his reputation does.

I don't believe that the small differences which separate riders or the variation in a rider's performance from one year to the next can normally be recognised by the riders on the stand. The margins involved are too small to be detected just by watching. For instance, readers might remember that our first top-rated jockey had finished only fifth in the jockeys' championship that year, 1983. It's fair to say that most people, including professional racing people, would not have put him in their top three riders at that time. Next year he won the jockeys' championship and lo and behold, judging from numerous public comments, was deemed to have improved significantly. The simple fact is that he was already absolutely top-class in 1983 and certainly had no need to improve. All he wanted was more mounts with good chances and he got that in 1984 so he rode more winners and attracted more notice. Other jockeys have since received the same sort of delayed recognition, years after they first featured prominently in our annual table.

Now to the question of a jockey's performance and his age. Surely it is perfectly natural for a jockey to be a more effective rider at thirty or thirty-five than he will be at forty-five? By the age of thirty any decent rider will have had thousands of rides. A few thousand more by the age of forty-five isn't going to add very much to his experience. Certainly not when compared with the normal physical deterioration experienced over that



period. Riding thoroughbreds in races is an activity with a strong physical element. If a man at forty-five can perform as well as he did at thirty then, in that respect, race-riding must be unique amongst physical pursuits. Is there any other sport with even a moderate physical content where younger men do not pre-dominate? The difference is of course that, unlike other sportsmen, jockeys do not compete on equal terms so their true performances are obscured. Whereas it will be obvious from his results if a tennis player or golfer is playing badly, with jockeys the situation is different. A jockey's loss of form will generally pass unnoticed because the rider is not usually the determining factor in the outcome of a horse race, the horse is.

In conclusion, we believe it is logical to measure a rider's effectiveness by the difference he makes to his mounts' performances and that is what we do. Because of the very large amounts of data and calculation involved, this process cannot be performed manually. It certainly can't be performed just by watching races.

---

## THREE POUNDS A LENGTH

Horseracing started to become an obsession with me around 1970. Apart from the glamour, the excitement and the betting it had a form book. In those days newsagents sold a weekly paperback edition. I remember the first time I bought one. At the bottom of each race were printed the 'speed figures'. Naturally I began to check how they were worked out. Half an hour into them I realised that some strange things were happening. For one, at two miles the compiler was using 1lb per length and one tenth of a second equivalent to 1lb. This implied that a length was about four feet ten inches, a figure at variance with my admittedly limited experience of racehorses. A horse running two miles in 3 minutes 40 seconds whilst being beaten twenty lengths would, under this system, record a speed figure about 20lb better than if it had actually won the race in 3 minutes 40 seconds. Either I was missing something or the compiler of speed figures was.

Four years later, fascinated beyond redemption, I joined a very reputable and dignified firm which produced ratings. At five furlongs they used 3lb per length for form ratings but 4lb per length for time ratings. This appeared completely illogical.

By now I suspected that many of the accepted conventions in horseracing were fairly unsound. Including two used in the derivation of official ratings. Apparently the Jockey Club handicappers' guideline at five furlongs was 2lb a length for two-year-olds, but 3lb a length for older horses. A fifty per cent increase in strength from two to three? An increase in strength is to be expected but one of order would be remarkable. Put at its simplest, if two horses met at level weights over five furlongs and finished separated by ten lengths, they would presumably be rated 20lb apart in November as two-year-olds but 30lb apart if they repeated the result four months later as three-year-olds.

The development in strength to maturity can be represented by a weight-for-age table and one use of that is in the construction of ratings. A cursory inspection of the official table will show that the rate of development slows down through the winter layoff. Over five furlongs two-year-olds develop 30lb from March to November, then 13lb from March to October as three-

year-olds but they only develop 2lb over the intervening winter. But look at three-year-olds over two and a half miles. They develop 14lb from March to November whilst they are actually in training, 1lb every two weeks, like clockwork. One might expect them to make perhaps 4lb development at most over the winter. One would be wrong – according to the official scale. Over the winter the three-year-old's development actually accelerates and it manages to develop the whole 11lb to reach full maturity by March. Hardly seems likely does it? The official scale is of course, '... founded on the scale published by Admiral Rous and revised by him in 1873. It has been modified in accordance with suggestions from the principal trainers and principal authorities'

The reader may make what he wants of this brief catalogue of anomalies. My own conclusion, reached early and reinforced repeatedly, was that the level of technical knowledge in horseracing was not forbiddingly high. One example says it all. If there is a single keystone to form analysis it is the use of 3lb per length at five furlongs with roughly proportionally reduced equivalences at longer distances. In almost twenty years I have not met anyone who can demonstrate why this should be so. The reply that 'it been proved by experience' only demonstrates a lack of real understanding of the fundamentals. Yet the justification is pretty straightforward.

In the horseracing world there is an endearing tendency to conduct weighty and solemn discussions and to hold very firm opinions without being too concerned about a thorough understanding of the subject in question. A currently fashionable topic is split race times. Here the public discussion proceeds portentously with no reference being made to the obvious requirement for – and present lack of – standard times and wind corrections for the individual distance segments. The effect of the draw is another favourite though not so much as it used to be. But the most perfect illustration of this attitude at its most rampant occurs when the subject is the assessment of jockeys.

Every racing professional worth his oats knows he can assess a jockey's ability. It seemingly does not register that the only credible measure of a jockey is how well horses run for him compared to how they run for other jockeys. If horses in general run better for Smith than for Jones then, by definition, Smith is the better jockey. It is impossible to imagine an

alternative sensible definition of relative merit between jockeys. But some people can imagine anything.

One belief which many British racing people hold dear, in fact are practically weaned on, is that British jockeys are the best in the world. This presupposes that, not only is the true believer able to rate British jockeys, but he monitors accurately those in the U.S.A, France, Australia and so on. One or two riders from the Colonies have been working in Britain recently and their ratings in our annual jockeys' table – top in four of the last five years – would appear to contradict this particular nugget of received wisdom.

Returning to handicapping, that exalted activity. 'A good handicap is like a work of art; a painting by Picasso or a symphony by Beethoven', wrote Dick Whitford who could not visualise a handicap being produced by computer programs alone acting on raw form data. A year later, in 1971, another very eminent expert, Phil Bull - and there wouldn't be a more perceptive man in horseracing at that time, concluded in a letter to the Sporting Life: 'So, where handicapping is concerned, you can just forget the computer@. Times change. We published the first computer handicap in 1984. From the outset its efficiency was superior to that of any manual handicapper. Our computer handicapping system is still in its relative infancy. In the coming years it will leave this early level of performance well behind; that would be expected of any technical innovation as it is developed. As for manual handicapping, that became obsolescent in 1984 – though I've no doubt most racing people have not realised it and, of those who have, some retain a vested interest in not broadcasting the fact since ratings in their own publications are currently dependent on manual methods. All the same, any young racing blood with ambitions to follow in the steps of David Swannell or Dick Whitford would be well-advised to forget them. Once a process has been automated a reversion to manual methods does not happen. There's no going back.

---

## WHAT'S WEIGHT FOR AGE?

In top-class races open to different age-groups the horses meet at weights taken directly from the official weight-for-age scale. So the victor is frequently not the best horse but the horse best at weight-for-age, which is not the same thing at all. Horseracing seems to be the only sport in which competition at the top level doesn't take place on level terms.

Top-class racehorses are expensive to keep, with capital and insurance costs to be met, quite apart from training fees. Such considerations have presumably always dictated that racehorses should perform before they are fully developed. Sooner or later they come up against older horses who are further along the development path. To compensate for their immaturity they are assigned less weight than their older competitors. This amounts to giving the younger ones a start, which is perfectly acceptable as long as the race is recognised for what it is, a handicap based on age. But this practice is carried through to the very best races, the ones popularly imagined to sort out the champions. Races such as King George VI & Queen Elizabeth Stakes are not events where the runners meet on level terms: they are simply age-based handicaps. Only if the winner has carried top weight can it be regarded unconditionally as the best horse in the race.

It seems reasonable that in a race which is supposed to be a championship-level event all the runners should carry the same weight so that the winner is unequivocally the best horse. Primitive considerations of fairness aside, there could hardly be a better incentive for owners to keep their Classic horses in training – unless of course they were so brilliant as three-year-olds that they managed to beat their elders at level weights and not just when given thirteen pounds start.

As long as weight-for-age contests continue at the top level, horses will continue to gain undeserved reputations, such as would be ridiculed in any other field, for beating better horses whilst receiving start.

The idea of weight-for-age dates back a long way and it is essential to the

compilation of weights for routine inter-age competition where the production of champions is not the aim. The official Flat weight-for-age scale '... is founded on the scale published by Admiral Rous and revised by him in 1873. It has been modified in accordance with suggestions from the principal trainers and practical authorities'.

A cursory inspection of the scale shows that the rate of development almost invariably slows down through the winter layoff. Over five furlongs two-year-olds develop 30lb from March to November, then 13lb from March to October as three-year-olds but they only develop 2lb over the intervening winter. But something strange seems to happen to three-year-olds over two and a half miles. They develop 14lb from March to November whilst they are actually in training, 1lb every 2 weeks, like clockwork. Then at the onset of winter the three-year-old's development suddenly increases rapidly and it manages to develop the whole 11lb to reach full maturity by March. This looks remarkable. In fact the development is apparently even more rapid than is evident from the table alone. Twelve years ago it was announced that: 'As a result of an international decision between Handicappers and Veterinarians the weight difference between 4-y-o and older horses has been abolished because horses are considered to have reached maturity in their fourth year'. So over two and a half miles the three-year-old must develop 11lb during December, whilst out of training, to reach full development by January when it reaches the age of four. This seems scarcely credible.

The wording of the decision mentioned above confirms that the official scale is intended to represent the natural physical growth to maturity of the racehorse.

Nevertheless four-year-olds do receive weight from older horses in some Flat races, those run under National Hunt rules. The weights in these races correspond with those indicated by the official scale for hurdles. From which, by the way, one could conclude that weight-for-age for hurdles is not intended to allow for any development in jumping ability.

The weight concession to four-year-olds in National Hunt Flat races would appear either to contradict that four-year-old racehorses are physically

mature or, if they are mature, to contradict that weight-for-age does represent their growth rate. The dramatic drop in weight-for-age development indicated over five furlongs for a two-year-old during the off-season to less than a quarter of its previous rate would also conflict with the idea that weight-for-age purely represents growth. For it appears extremely unlikely that its growth rate would behave so erratically.

A realistic weight-for-age scale is a fundamental necessity. What should it really represent and how should it be derived? For what reason does a young racehorse improve its level of performance as it gets older? Foremost is physical growth. To some degree the rate of growth will depend on feeding and this may be a factor which has changed significantly through a more scientific approach being adopted. It is questionable whether the introduction of more American blood has had much effect on the general growth rate. At one time it was quite commonly stated that American-breds matured more rapidly than their European counterparts. As I recall, this became almost an axiom when Vincent O'Brien was sweeping all before him with a fleet of such horses. This view is not supported by the American weight-for-age scales which show a development broadly similar to the official British scale.

The acquisition of racing expertise is also an obvious component of improvement in performance. The rate at which this is acquired will depend on the individual trainer's approach – whether he brings his horses along gradually or not, how soon and how frequently he races them. It may be that this particular factor will vary between horses more than the growth rate will.

To separate out accurately the individual contributions of growth and education from the development rate would appear to be extremely difficult if not impossible. Fortunately it is not necessary in order to obtain an accurate weight-for-age scale. A practical approach to deriving an accurate scale is to examine the progressive race performance ratings of a large number of individual horses and simply chart the average development shown. This is what we did for our own Flat scale some years ago and we expect to revise it now that we have amassed far more data. Recently we performed the same exercise for our National Hunt scale and found that National Hunt horses do on average improve their level of performance up

to the age of about eight years old. The improvement is relatively small, but definite, after the age of six. It would appear that this continued development is due to gradual improvement in jumping ability. For, roughly speaking, if racehorses got much faster after the age of four or five then Flat racing ought to yield tangible evidence to that effect. As well as the acquisition of technique, jumping ability depends on muscles which may continue to develop long after the age of four. Certainly the show-jumping fraternity, for their purposes, regard a four-year-old as years from its physical peak.

There is little doubt that the convention of associating weight-for-age directly with the growth of the species is impractical. Weight-for-age should be a pragmatic measure of the average horse's improvement in racing performance with age. As such it can best be gauged by a statistical analysis of progressive race performance ratings. Since performance will be somewhat dependent on training and feeding practices and perhaps breeding patterns, the rate of development may change slightly over a long period. The scale should therefore be re-evaluated periodically to check for any underlying long-term changes.



---

## EDITORIAL COLUMN FROM THE 1990 FLAT ANNUAL

If you had devised a betting system which generated a twenty per cent profit on turnover, what's the first thing you would do with it? What's the last thing you would do with it? The answers – in my opinion – are 'use it' and 'sell it' respectively. You'd have to be really dumb to sell it.

Our own analysis of racing form revolves around our computer handicapping system, which is not the sort of system referred to above. With its associated supporting functions, it currently amounts to over six hundred interacting programs, though some of these are used only once or twice a year. It rates horses more accurately than manual handicappers are able to do and this is borne out by the ratings' profit/loss record. A blind level stake bet on our published top-ratings has shown an average loss of less than one per cent per season since they first appeared, whereas a similar bet on the usual mass-market manual ratings would have lost around six or seven per cent on turnover. Alex Bird, one of the most successful backers of all, states in his autobiography that his overall net profit was less than two per cent on turnover, which puts these figures into perspective. For obvious reasons then, we would not sell our computer handicapping software (and the long period of tuition necessary to operate it!), unless it were for very serious money indeed.

But when horseracing people talk about systems, it's betting systems they mean. And amazingly, judging from advertisements in the racing press, successful systems can be had for the price of a good night out in Brighthouse. How can such philanthropy be explained?

A proper betting system is an automatic selection system. It consists of a set of rules which, when applied, lead unequivocally to the system horse or horses, if any. The system may also specify the minimum price required and the stake. There is no doubt, no fudging, about what to bet on and the

amount to be bet. If the bets are not uniquely defined then clearly there is no way of measuring the system's success or failure.

The only valid measure of a series of bets is its percentage profit or loss on turnover – the bottom line. A period of several years would be needed to determine accurately the underlying percentage for a system. The rules of the system must be established before the period starts, for it is a very simple matter to examine a sequence of past results and then produce a set of rules which would show an excellent profit if it had been applied to this sequence. Some system sellers do this all the time. That is how they are able to claim years of profitable operation for their systems. Their rules will have been conveniently designed after the event to fit the pattern of these years. It also explains why they need to bring out new systems, or Mark II versions of the original, in subsequent years as more hindsight is acquired.

Setting aside such after-the-event designer features, there are other reasons why a system may work well on paper but not in practice. One is that a system containing a recovery facility may deal with a long losing run by increasing its stakes until the run is finally terminated by some theoretical monster bet going on to a long-priced winner. In the real world, the betting market would not absorb so much at that price.

Phil Bull, who died last year, was one of racing's originals. 'Don't buy tips or systems; geese that lay golden eggs are never for sale,' ran his advice to readers of the Daily Herald in 1961. I'm not sure what his definition of a tip was, bearing in mind some of his own business activities, but then he always was a bit of a card, with the flexibility to ignore his own preaching. My time with his business was spent in the office adjoining his. He regularly received phone calls and letters from people who had successful systems and sought his wisdom on what to do with them. His reply was invariably the same: 'Brother, if you have a successful system, you do not need me.' This saved him from the tedium of wading through it and finding its weakness. The possibility of someone intelligent enough to develop a successful system being, at the same time, insane enough to approach him with it was discounted. Curiously though, he did enjoin me more than once to let him know if I 'came up with anything' in the way of a viable system and we could operate it together. 'Don't worry about that,' I thought, 'you

won't see me for dust.' One winter he did think that he'd come across a system of a sort. Examination of the previous Flat season had shown that backing horses whose timefigures were above the average form level of previous winners of that race would have resulted in a worthwhile profit. I pointed out to him that the sample size was not large enough to satisfy routine statistical confidence criteria, but he went ahead with it for real the following year until calling a halt before the end of May.

Some readers might wonder why I've devoted so much space to systems. Well let me tell them, they are in a minority of the racing public. Over the last seven years I've spent so much time on correspondence and phone calls explaining the facts of life concerning systems that it seemed a good idea to set it down once and for all – at the very least, it means I need only send out copies of this in future. Just last week I had a phone call from someone who has for years marketed, apparently lucratively, a computer selection program. 'What percentage does it make on turnover?' I asked. 'Eighty per cent' he said at first, but after some discussion I managed to beat him down to a mere twenty per cent. He wanted my help to improve on this!

In case there is anyone reading this who still doesn't know why it's a dumb thing to do to sell a successful system – and there will be someone – this is what would happen if you did. When you have a successful system to yourself, no one else knows beforehand what the system bets are to be. Consequently you can wait for the prices to become available and then place your bet – or maybe not place it, if the system has specified a longer price. Now suppose you sell your system. Then the knowledge, formerly privy to you, of which horses are to be backed and, if the system indicates it, at what price, is generally available in advance – before the prices are determined. You are no longer able to get in first. In particular, the system horses are available to bookmakers and you will find that they are not benevolent enough to mark up prices which, they now know, would have contained a substantial margin in the backer's favour. So the margin disappears from the prices of the system horses and the system is no longer profitable. So the answer to the question 'Why would anyone want to sell a successful system?' is very simple. They wouldn't.

In 1986 Dancing Brave beat Shardari by three quarters of a length in the

King George VI & Queen Elizabeth Stakes. The horses met at official weight-for-age terms so the four year old Shardari carried 13lb more than the three year old Dancing Brave. The weight-for-age scale we used made the difference between the generations 11lb at that time of year. So on our ratings, after allowing 1lb for the three quarters of a length, Shardari came out 1lb above Dancing Brave for this run. All very simple.

When we sent out review copies of our Flat annual for that year it seemed a sensible precaution to explain in the accompanying letter why we had rated Shardari above Dancing Brave at Ascot. Then, anticipating the question of why we did not agree with the official scale, I added a couple of paragraphs drawing attention to some of its obvious anomalies. A week or two later I received a phone call from the Chairman of the Flat Race Pattern Committee, who had heard that I had 'some interesting ideas on weight-for-age', asking whether I would like to produce a paper on the subject. The inconsistencies in the scale were so apparent that the case made out in the letter was already sufficient and I sent him a copy. The following year in this column I mentioned the anomalies and recently, I see, they have been removed.

I don't know Tony Ives. All I know is that each year when we produce the ratings which measure how British racehorses have performed for their different riders, the names Ives is invariably near the top. Taken over the last seven years his average annual rating of 12.3 is better than any other jockey to have ridden regularly in Britain during this period and his rating has never been less than 12.0 in any year. Only Brent Thomson, who averaged 12.5 in his four seasons here and Cash Asmussen, with 13.7 for his one full season, would come out with a higher figure.

But 'racing's professionals' generally don't rate Ives so highly. These people see a picture composed of numbers of winners, prize money, winning percentages and big race wins overlaid with an impression of riding style, strength in finish, artistry of whip switching technique, aesthetic appearance of any arm flapping, and imagined tactical appreciation. The colossal distortion caused by the inequitable allocation of mounts appears to be invisible to them. Naturally, these connoisseurs have also withheld their imprimatur from foreigners with such curious riding actions as Cash Asmussen and Brent Thomson – actions which are totally unsuitable, of

course, to race riding on British tracks, which, it goes without saying, are the true test of a rider and, as everyone knows, the reason why British jockeys are the best in the world.

It wasn't long after I became seriously interested in horseracing that I realised no one could tell what effect a jockey really had on a horse. In other words, no one could tell how good or bad a jockey really was. One person, not a standard-issue racing professional, who did appreciate that assessing jockeys was an unresolved problem was Phil Bull and we discussed the matter a few times. He told me once about a conversation with Geoff Lewis at a time when Lewis was a contender for the jockeys' title. Making the assumption that Piggott was the best jockey of the day, Bull asked Lewis how much better Piggott was than the next man in the line. Lewis said 'at least 10lb.' Now without going into the technical detail, the fact is that if one rider was 10lb better than any other and, what's more, was able to a great degree to chose his rides from a shortlist of the most probable winners – as the dominant Piggott could at this time – then he would ride a percentage of winners of perhaps 60%. So 10lb was a ludicrous assessment and Bull managed to convince Lewis of this. They eventually settled on 3lb – for what it was worth – with Lewis adding the proviso that it was possible that Pat Eddery might be closer than this to the great man. I tell this story to demonstrate that a highly respected racing professional, at the time one of Piggott's closest rivals, who could claim, if anyone could, to be ideally qualified and in the best position to assess the man's abilities, came up with a wholly unrealistic answer. I could give many similar examples – John Francome 14lb better than the rest, a famous French jockey 14lb inferior to Piggott, and so on – expressed in print in all seriousness by certain racing professionals. Only the other day I read that one trainer used a particular rider because 'he gives you a 7lb advantage over most other horses'. A jockey's reputation and, consequently, the quality of his rides and livelihood are at the caprice of these people.

Our database is now in its eighth year and we have accumulated enough data to produce some worthwhile average figures for the ability and best distance of a sire's progeny. The corresponding figures for the dam sires are also interesting but less important since the dam sire's contribution is only half that of the sire. These figures are clearly more valuable than the rougher measures based on prize money and winning distances.

Even so, this innovatory approach still has room for improvement. The sires' figures tell only half the story, the other half being due to the dams. Taking two sires with a large number of progeny, so their averages are likely to be reliable, Be My Guest's progeny average out 3lb better than High Line's. Without knowing the quality – in breeding terms, not racing ability – of the respective dams, one could say that Be My Guest was a better sire than High Line. If it turned out that the average level of Be My Guest's mares were 10lb better than High Line's, then one might conclude that High Line was in fact the better sire. To do the job properly, to determine an unbiased measure of the sire's ability, one needs to account for the ability of the dam. It's rather analogous to the way we derive jockey ratings by effectively eliminating the ability of the horse from the performance. The parallel is not quite exact because there is an additional complication which is not present in the jockeys' case. We are technically able to perform the process now, but need a little more data for the results to be solid.

One interesting result did emerge during work on these figures. I have always thought it seemed reasonable that a horse's ability to handle soft ground depended largely on its action, which is determined by the way the horse is put together physically, and on other physical characteristics. These characteristics would to some extent be inherited from the sire. Therefore one might expect to find some sires whose progeny showed a marked ability to act better or worse on soft ground than they did on a sound surface.

By splitting the data for each sire's offspring into performances on the soft and performances on a sound surface – and correcting for the smaller number and slightly lower overall level of soft ground figures – I was able to test if there was any statistically significant difference between the levels of performance of the progeny on the two types of surface. Rather to my surprise, there was not one sire where this difference was really significant.

---

## EDITORIAL COLUMN FROM THE 1991 FLAT ANNUAL

A couple of recent occurrences prompted the thought that it might be worth examining why so many racing people persist in regarding the rating of racehorse performances as an art.

In deciding what level of performance a horse reached in a race, the sort of things to be taken into account are where it finished relative to the others, the standard of its competitors, the weight it carried, the weights the opposition carried, the race time and many other factors. All this raw data describing the horse's performance is numeric. The required assessment of the value of the horse's run will be a rating expressed on some standard scale so that it can easily be compared against others: this rating will also be a number. Problems where the initial data is all numeric and the desired solution is also numeric are generally regarded – outside of the racing circus – as mathematical in nature and best dealt with by mathematical techniques. This is how Racing Research handle the job.

There are two basically sound approaches to rating racehorse performances, by collateral form and by race time analysis. Collateral form rating is known as 'handicapping' in Britain, though the term does not mean the same in the USA. Speaking very roughly, it amounts to comparing one horse's performances with the performances of others, directly with those it has actually run against and indirectly by examining how these opponents have run against other horses. Assessing the time value of a performance, on the other hand, involves comparing the actual race time with some 'standard' time for the particular race distance at the course and compensating for the external factors which distort the time such as the state of the going, the wind and any temporary course alterations.

Any rating, form or time, is an estimate. The intention is to produce as

accurate an estimate as possible, given the data available. In a good-sized field of exposed horses it will be possible to make a very large number of collateral form connections with other races, giving plenty of form data to work from, so form ratings from such races tend to be relatively accurate. At the other end of the spectrum are form ratings from races between unexposed horses, ratings which can only be tentative. Time ratings, or 'timefigures' as they are historically better known, will be at their most reliable if they derive from meetings where the external effects are minimal.

As a general observation, there is far more that can go wrong with time ratings than form ratings. Having made that qualification there is no serious question as to the worth of good timefigures. Some well-known backers from the past more or less based their betting on time considerations and several very successful present-day backers of our acquaintance take timefigures very seriously into account. The classical situation where time ratings come into their own is in a truly-run race between unexposed horses where any form rating is necessarily speculative but the time analysis indicates a decent performance.

It is axiomatic that only if a race is truly-run can the timefigures from it reflect the real value of the performances. For any race between reasonably exposed horses – so the form rating will be fairly accurate – a measure of the pace of the race is provided by the difference between the form rating and the time rating. For truly-run races the time rating will be fairly close to the form rating. For the others the degree of slowness will be measured by how much the timefigure falls below the form rating. Incidentally the terms 'pace' and 'truly-run' are of course used relative to the runners' capabilities – a truly-run pace for moderate handicappers would be a slow pace for Group horses. Form from truly-run races tends to be more reliable than form from slowly-run affairs. The theory is that a false result can occur more easily in a slowly-run race where an inferior horse is able to stay with the pace and then use its turn of foot at the sharp end. So even slow timefigures have a use, enabling inferences about the relative pace to be drawn.

It is evident then that form and time ratings each have their own advantages and any attempt to produce a combined form and time rating



would involve the loss of some information. That is why we produce form and time ratings for every race.

There are foundations common to the construction of both types of rating. Ratings are normally expressed on a pounds scale, a convention arising naturally from the use of weight as the handicapping element in horseracing. A conversion rate from lengths to pounds is then required and this will depend on the race distance. At five furlongs the equivalence is about three pounds a length. The technical justification for using this figure is relatively simple, though I have never known any handicapper or timefigure compiler who could give it. For the purposes of ratings derivation, it is more logical to regard the 'distance' of a race as the time taken to run it rather than the nominal distance, a practice I introduced at Timeform in 1974. So instead of a basis of five furlongs, a basis of sixty seconds is used. The pounds per length equivalence is then modified more or less in inverse proportion to the race time. A five furlong race on fast going at Epsom taking 55 seconds and a five furlong race at Beverley on heavy ground taking 70 seconds are therefore treated differently. The race time is by no means a perfect measure of the 'severity' of a race – for one thing, truly-run races will appear to be less 'severe' than races slowly-run. On balance though, it is far more serviceable than the nominal distance. The use of the time as a measure of distance also has another advantage. It renders immaterial any errors in nominal distances. Whether an advertised 12fg is really 11fg 170yd or 12fg 70yd becomes irrelevant.

Weight-for-age, the rate at which an immature horse develops racing ability, is another fundamental. The consequences of ignoring weight-for-age can best be illustrated by some simple examples. Suppose a three-year-old and a five-year-old dead-heat at level weights in April. Then a ratings compiler oblivious to weight-for-age will rate the two horses equal. If they meet again at level weights in July, then his April ratings will indicate they have an equal chance. However, in the intervening months a normal three-year-old will have improved by several pounds while a normal five-year-old will not have improved at all. It is clear that, going into the July race, the three-year-old should be assessed higher than the five-year-old by the amount of this improvement. As another example, consider two horses who have one race each, which they win, as two-year-olds carrying 9-0, both races being at 6fg over the same racecourse, one in July, one in October, with both races run in exactly the same time. Suppose our 'I don't bother with weight-for-age' (he probably thinks it's a pension scheme)

timefigure compiler assesses the going allowance as identical on both days. Then he will arrive at the same timefigure for both horses. The following April they meet at level weights at 6fg. According to our expert, on the evidence of his timefigures, the horses are equally matched. This is unlikely. Given that both horses have developed normally, then, going by our weight-for-age scale, the July debutant can be expected to have improved 19lb since it ran while the other has improved 9lb since October. So in April the July runner should be rated 10lb superior to the October runner.

The practical way of dealing with the effect of weight-for-age on ratings is to add into the achieved rating the appropriate immaturity figure for that distance at the date of its achievement. This gives the 'mature rating'. Then at any later date, to get the 'current rating', subtract the immaturity figure for that date from the mature rating. The current rating will then be greater than the achieved rating by the amount of normal development between the two dates.

Turning specifically to timefigures, weight-for-age is crucial to the production of genuine standard times. So are several other things, but since our methods of producing standards are unique, it would not be very sensible to describe them here. Incidentally one thing that does affect the accuracy of a standard time – or the timefigures themselves – is any discrepancy between the nominal distance and the true distance. Problems in that area only arise when trusting souls decide to do things like dividing (inaccurate) nominal distances into the standard time or the race time in order to get the number of seconds taken per furlong and then reading something into that.

Still on timefigures, a conversion rate from time to pounds will be needed. This follows automatically from the pounds per length equivalence. At five-furlong sprint speed a length is covered in about .16 sec – as can be seen from photo-finish strips or verified directly. So if a length were taken as 3lb then 1lb would equate to slightly more than .05 sec.

The going allowance and any correction for wind effect determine the level at which the final timefigure will be set. Going allowances should be

expressed simply in pounds. The commonly used 'seconds per furlong' amounts to the same thing but is unnecessarily messy since it has to be multiplied by each race distance. Clearly the derivation of the going allowance is the most important part of the actual production of a day's timefigures. The data going into this process consists of standard times, race times, weights, the abilities of the day's runners (best represented by their form assessments on the day, another convention I introduced at Timeform), etc. Since all this data, without exception, is purely numeric and since the required result, the going allowance, is also numeric, the problem is unequivocally mathematical. To solve it involves first framing it in mathematical terms and then utilising the appropriate mathematical techniques to deal with it.

Of course if you don't happen to be familiar with the requisite mathematics, then the problem takes on a different complexion. Last year I received some criticism from Roger Coates and Ken Hussey, compilers of the well-known Split Second speed figures which they produce under the byline of 'Britain's Master Clocker'. Their main plank, if I have grasped it, was that timefigure compilation is 'an art not a science'. In particular, with regard to the problem of determining going allowance: - 'There can be no hard and fast rules on this tricky subject, just the eye of a skilled practitioner well-versed in judging the "class" of meeting and the previous known ability of the more-established runners taking part'. The editor of their newspaper decided not to publish my reply to that one, though I felt that it could well have been of interest to some of his readers. I think this must be what is meant by the 'freedom of press'.

The substance of my reply was that, yes, timefigure compilation was an art to Messrs Coates and Hussey. It had to be. Since they demonstrably lacked the proper tools for the job – the mathematical capability – they simply had no other option, they had to use their 'eye'. To show they were lumbered with this particular ball and chain, I had, in an earlier letter, cited some very obvious errors in their system, which no one with any knowledge of mathematics could have made.

One concerns the basic time and lengths equivalences. The scale used by Split Second contains fundamental inconsistencies. For example, at two miles Split Second uses 1 length = 1 point and 1 point = 1/10 second. This implies that 1 length is covered in 1/10 second, from which a trivial deduction shows that a length must be about 4' 10". I don't think it is. A

horse running 2 miles in 3m 40s whilst being beaten 20 lengths would, using this scale, record a speed figure about 20 points better than if it has actually won the race in 3m 40s. If this is intentional then the reasoning behind it escapes me.

Then there is the going allowance. The Split Second going allowance is always given in multiples of .05 sec per furlong. In a 14fg race, .05 sec per furlong amounts to .70 sec or 7 points in Split Second's terms at this distance. Suppose a horse carrying 9-0 wins a 14fg race in 2.0 seconds above the Split Second standard time. If Split Second chooses a going allowance of nil then this horse's speed figure becomes 80. Since Split Second only uses multiples of .05 then whatever going allowance he actually chooses can only make this horse's figure differ from 80 by a multiple of 7. The only possible speed figures which Split Second can give this horse are figures such as 59, 66, 73, 80, 87, 94, 101 etc. Figures in between are totally eliminated from consideration. It is the use of the .05 minimum unit which has enforced this unnecessary restriction on Split Second's possible range of figures.

And as for weight for age, Split Second ignores it entirely in his calculations. It would seem reasonable to assume, therefore, that it doesn't enter into his workings for standard times either. He also apparently doesn't believe in wind corrections.

In case anyone thinks that Split Second's approach is unusual, I should mention that most published timefigures seem to follow the Split Second principles. On the matter of the feasibility of developing an algorithm to produce the going allowance, it is also worth noting that when I joined Timeform I was told that this part of the process would always have to be done by 'judgement'.

Collateral form rating is essentially more complex than time analysis and the development of a purely automatic collateral form rating system – a true computer handicap – is at a much higher level of difficulty than the computerisation of timefigure production.

In order to be able to design a collateral form rating system which can be implemented as a system of computer programs acting on a database, it is first of all necessary to be able to understand precisely what collateral form rating is. It is my experience that most manual handicappers would be hard pressed to define exactly what it is that they are doing. They do the job, sure, following various procedures and devices and their own intuition or flair, as some might call it, and they use their acquired experience. But if they were asked to explain in a rigorous way what the resultant rating is, what it represents and exactly how it has been derived and on what basis, most of them would struggle. This is not to say they are not competent at their job, for the majority are. However, since they don't fully understand what they are doing, even though they can do it, they imagine it must be an art. Consequently it is not possible for them even to contemplate any automation of their task. This attitude is reinforced by most racing professionals, to whom handicapping is something of a mystery, and there are those who seem to think that the ability to handicap racehorses is a gift granted only to a chosen few.

I have not encountered a manual handicapping procedure that can't be satisfactorily reproduced as a computer process. Over and above the routine handicapping operations, we use procedures which are just not accessible to a manual handicapper. In reality a handicapper's greatest single disadvantage is that he suffers from the human disability, relative to a computer, of not being able to process information fast enough – the job essentially involves huge amounts of information processing. In particular, he cannot handle the ripple effect, which underpins the concept of collateral form analysis, except in a most cursory manner. I notice this morning in the racing press that, with regard to his rating of a famous horse which ran recently, one handicapper said he had been looking again at an earlier race with a view to reappraising his assessment. Our computer system does that automatically after any new results have been input – meaning six days a week – for every single performance over the last few years.

I have already referred to the blind utilisation of 3lb per length. It is only by asking why something like this works – instead of just accepting that it works – that the underlying principles become clear. The same applies to weight-for-age. The official scale, presumably utilised happily by the official handicappers and others, used to contain some blatant anomalies. It wasn't until I drew attention to them that the scale was changed. Then there was the practice of using 3lb per length at 5fg irrespective of the severity of the

5fg – again this indicates a lack of understanding. At Timeform they used a different basic poundage per length for form ratings than they did for time ratings – totally irrational. There are flaws in other procedures but it appears that the mere knowledge of their existence constitutes a Racing Research trade secret. In any case, I think I have established the validity of my point that most handicappers do not possess a deep understanding of what they do.

A recent newspaper interview with the marketing director of Timeform contained the following statement:- 'The computer is programmed to produce timefigures but the handicapping is very much an art and not a science'. It

might be instructive to recount briefly some of the facts of my association with Timeform.

I first got really interested in racing about 1970. Having become familiar with the idea of collateral form rating, I realised that it should be theoretically possible to develop a computer handicapping system, though it was clearly a formidable project. That would be about 1971. From that time I started to think seriously about it. In 1974 I joined Timeform as computer manager where my primary responsibility was for timefigures. By the late nineteen-

seventies I had on paper a prototype design for a computer-handicapping system. I needed a computer to start testing my ideas. So I approached my employer, Phil Bull, and explained that I had some ideas of my own which I was developing and that I intended to buy a decent computer – which was still quite unusual for a private individual in those far-off days. I said frankly that I was not prepared to divulge what I was doing at this stage and, that being the case, I wished to know whether he would prefer that I left Timeform. He preferred that I stayed. Around 1981 or 1982, Phil Bull himself began to wonder about the possibility of producing a computer handicapping system. This came about as a result in his interest in chess. Having purchased, in short order, a series of chess machines (as computers with dedicated chess-playing programs were known in those days), he concluded that as computers could be programmed to play chess then at some time in the future someone might be able to design and implement a computer-handicapping system. Bull was the first person I had come across with the ability to visualise this possibility. His exact words were, 'The guy

who produces a computer handicap will have a bonanza'. The reason he could see its feasibility was that he understood the principles of collateral form analysis, Of course he asked for my view and I replied that it was something to which I had given a lot of thought and that, in my opinion, it was a very difficult problem. Subsequently he expressed the firm view, to me and others, that it would eventually be done. I left Timeform in 1983 precisely to ensure I got the credit for developing the computer handicap.

It goes without saying that the production of ratings for racehorse performance is an insignificant activity. Nevertheless the principle is the same as it has been for many more important subjects. It is only an 'art' until the time comes when it can be treated as a science.

John Whitley

© Racing Research 1984 - 2012