A DEFECTIVE APPROACH IN UK STEAM LOCOMOTIVE TESTING

A RECONCILIATION OF TEST RESULTS WHICH DID NOT SUCCEED, INTERNAL REPORT L116 OF DECEMBER 1957

John Knowles

In his letter to Milepost of 17.3.17, Doug Landau claimed that those BR officers conducting measurements and research into locomotive efficiency and outputs were scientists. I believe that those who worked at Derby, the Testing Section of the London Midland Region, were anything but scientists, that their work was anything but scientific, a lot mistaken. It was their function to conduct Controlled Road Tests to obtain figures for EDBTE consistent with the Rugby ITE results, and for years on end they produced erroneous results.

Abbreviations and Explanations:

ITE	Indicated Tractive Effort
EDBTE	Drawbar Tractive Effort made Equivalent to the Train Running on Level Track by correction for effects of gravity (gradient), and acc/deceleration. Omission of the E implies Drawbar Tractive Effort, that measured at the drawbar without the rendition of the figures to allow for gradient and acc/deceleration
MTU	Mobile Testing Unit, a vehicle with rheostatic brakes and control over the extent of the braking effect, and control to keep speed constant.
CRT	Controlled Road Testing of Locomotives on the Road, in contrast to the stationary Testing as on the Rugby plant, with devices to measure coal and water consumption, and instrumentation to advise the BPP to the driver, who can alter BPP by altering the CO of the locomotive. The locomotive can be equipped with indicating gear, but the advice given to the driver of the BPP is meant to avoid indicating. See S O Ell, Developments in Locomotive Testing, JILE, Paper 527, 1953 p 561.
BPP	Pressure of steam as steam is exhausted at the Blast Pipe, referred to Pressure absolute (14.6 lbs/sq inch higher than atmospheric.)
Q	flow of steam at a certain temperature and pressure, lbs per hour
LR	Locomotive Resistance

Despite Doug Landau's staunch defence of British testing and its numerical results, there was a large scale defect in one aspect of the UK approach to testing which led to incorrect EDBTE results being declared for many locomotives. It existed throughout the period of testing. Several Test Bulletins have incorrect EDBTE results and were never corrected. The defects, present during the whole period of testing, were eventually acknowledged in an internal report, L116, by the testing officers themselves. Many defects in procedure which probably led to the defective answers were also pointed out by various testing officers.

Background

For some seven years or so, some locomotives were jointly tested by Rugby and Derby, ie by the Testing Station at Rugby and by the Testing Section of the London Midland Region at Derby. The latter conducted Controlled Road Tests intended to obtain figures for EDBTE consistent with the Rugby tests, which were entirely in ITE. From the inception of these tests, it was found that the results of these road tests were inconsistent with the Rugby results, a problem of method, and/or measurement, and/or calculation of results from the measurements. This inconsistency was observed through LRs of the wrong shape, indeed impossible shapes. That meant there were errors in production and/or measurement of ITE and/or EDBTE, or calculation of EDBTE.

Only at the very end of steam testing was “something” done about this defect, and a method devised intended to correct the recent results. This depended on inserting speed terms into the relationship between Q and BPP, even though there was no dependence on V in the relationship between Q and BPP (as I show below), and deriving a correction equation, which was in fact an erroneous method of relating Q to BPP. The correction method was therefore muddle headed thinking¸ with no science in it. Having such a correction system would or might appear to make the answers they gave at the end of the testing “all right then”, while leaving a defect still present in the data published in the Test Bulletins which were the joint responsibility of Rugby and Derby, without any public admission of the defect or correction of the results of the testing. As importantly, there was really no valid correction system at all.
The internal documents concerned which are the basis of my conclusions (L109, R13 and L116), all prepared at the very end of steam testing, claim, however, that the correction system did convert LRs of the wrong shape into the correct shape. It is impossible to draw that conclusion from the fullest description of the correction mechanism or process. No data were given on the cases where the supposed correction led to the correct LR, ie the original data, the basis of correction, and the results of the supposed correction, and it not obvious that the corrections can be checked. Further, the judgement made, that the system worked, required a comparator, ie consistent ITE and EDBTE at various speeds for the locomotive under consideration. For the locomotives concerned, no such comparator LRs are known. The conclusion that the system worked is therefore without foundation.

This paper

This paper first considers the large number of wrong results, admitted in internal report L116. It then goes on to consider how those incorrect results could have arisen, and the modest research conducted with the intention of allowing the incorrect results to be corrected, research which was extremely poorly applied. The officers concerned considered that their results were wrong because they had not taken into account the effect of speed on the use of the Blast Pipe Pressure on the metering of steam. In that they were mistaken, for there was no such speed effect. The correcting mechanism and equation thy devised did not fit the data available, which led to wrong conclusions. They believed that they could conduct desktop corrections of results, but in that they were mistaken also, and no explained corrections of results was given. Nor did they perfect the testing and measurement, and to the end the Derby measurements of ITE proved defective, including that of a Duchess. Although Derby thought it had a system which could correct LR, it never explained where the comparator locomotive came from. Checks were made of the apparatus and procedure, but the Derby errors were never corrected. This is surprising because testing procedure with similar intentions took place at Swindon and seemed to operate satisfactorily – it was Derby which did not succeed in measuring properly, and which devised a supposed correcting mechanism which was not a logical explanation for the mismeasurement which occurred.
The data available is analysed herein much more soundly than done for L116. See below.
Derby did not run its side of the joint Rugby – Derby testing soundly.

The Intended Measurement System

LR is the difference between ITE and EDBTE. If the testing method was to reveal LR those two items are needed, correctly measured. (Lots of locomotive testing in the world, probably most, did not seek to reveal LR.) The BR Test Bulletins all include data on ITE and EDBTE separately, the ITE being the end product of the boiler and cylinder outputs, and the EDBTE the work the engine can do at its drawbar, measured there by the dynamometer car.
Testing should preferably be done on the road, where the engine will operate, and where the draft on the fire and the escape of the exhaust are those of the open air and where there can be evidence of the inevitable variations in atmospheric conditions. Postwar, the British testing system had the Mobile Testing Units, which could apply a rheostatic brake to achieve constant

speed, vastly better than use of steam locomotives with the cylinders acting as counterpressure brakes.
The aims and methods of the BR testing system were given in a paper by S O Ell, *Developments in Locomotive Testing*, read to the Institution of Locomotive Engineers, paper 527, 1953-54, and in BR Test Bulletin No. 1, 1951, on the testing of the Western Region Hall. Whatever the system and measurement, LR even at a continuous output, can be expected to show uncertainty and inaccuracy on account of the Small Difference Effect I have discussed previously, the result of ITE and EDBTE being both large numbers, the values of which cannot be measured precisely, resulting in considerable imprecision in the LR, the modest difference between them. The problems I discuss here concern conceptual errors mostly, errors in approach.

To minimise measurement difficulties, it would be usual to measure ITE and EDBTE on the same test train, and simultaneously, the engine running continuously at the same output for sufficiently long for information on the stability of coal and water consumption to be available. To obtain LR and nothing else, the stability of the output is the important consideration. The coal and water consumption are more important if efficiency is being established as well. Constancy of output can be obtained by having the boiler develop sufficient steam at full pressure to provide the output, then setting CO at a given rate, and having the MTUs operate at a given speed to give a constancy of V, Q, ITE and IHP. Indeed that very method was used to test the WD 2-8-0 and 2-10-0 engines, mostly the latter, including the more important boiler outputs, and was not regarded as defective (see Test Bulletin 7).

The testing officers of the day decided on using a mixture of the stationary Testing Station at Rugby and separate test trains containing the MTUs on the road, operating in a controlled way, termed Controlled Road Tests.

Errors and Oddly Shaped LR

The intention was to duplicate the Rugby ITE on the road, at various Q and V, in CRTs conducted by Derby. A BPP and V combination was conveyed to the driver, who was to duplicate it during the CRT by adjusting the CO. It was necessary to get the Rugby ITE and V combination right if the EDBTE corresponding to it was to be correct. During the test, the load on the drawbar and speed were regulated by the MTUs, while the DBTE was measured by a dynamometer car. All going well, this system was to provide a consistent set of Q, ITE, EDBTE and V at sufficient points to map EDBTE as appeared in the Test Bulletins. (The intention for the tests done at Swindon was similar).

Rugby was used to establish the boiler conditions and efficiencies and the ITE, and the test trains the EDBTEs. A constant Q and V can be obtained by setting CO at a given figure, and the MTUs to give a constant speed, with full boiler pressure applying throughout, the CO and V being chosen to duplicate a test at Rugby, which gave the ITE for the same CO and V. Instead of doing that, however, the blast pipe was used as a steam flow meter. The Blast Pipe Pressure (BPP) was the basis of measuring Q. A given BPP was measured during a given stationary test at Rugby by a mercury manometer. The aim was to reproduce the same Q on the same locomotive on the road by giving the driver a similar manometer to measure the same BPP. The same manometer and piping could even have been used in both cases – after all, the blast pipe and a location near the driver were needed in both cases on the same locomotive. The driverin the CRT aimed to achieve the same BPP and V as in the Rugby test by varying the CO.

It did not Work Out That Way

The intended system worked well, as a procedure, for testing done at Swindon. The accuracy of the Swindon indicator is a different matter, not discussed here. What follows applies to the system conducted jointly by Rugby and Derby.
In L116, a diagram is given for the LR of a Crosti 9F following the Rugby/Derby testing practice from 1950 until L116 was issued about December 1957. By then testing of steam power had ceased at Rugby as had associated road tests intended to complement the Rugby work, the two together becoming the content of several of the Test Bulletins. The content of L116 therefore admitted and exemplified the defect of method and measurement. That is the major contribution of L116, that it admitted large errors in the shape and presumably value of LR. (The intended correcting procedure is discussed later.)

Fig 1 in L116 shows that in the range 20 to 50 mph, the LR of a Crosti fitted 9F using the testing method used by Rugby/Derby throughout the whole of the testing period, was of completely the wrong shape, indeed an absurd shape. See line 1 of Table 1 below. LR declined as speed increased. That cannot have been. A correct LR rose with speed (the resistance from passing through the air, from revolving rods, and from the revolving masses associated with partly balancing the reciprocating masses all rose with speed, indeed very much with speed squared, while those from the application of steam to the mechanism fell with speed as ITE fell, as it had to if Q was constant for a test, the usual practice during BR tests. When particular tests are gathered together in a summary table or graph, any LR extracted therefrom should not be expected to be constant at any speed – it should be expected to vary with the effort as well. Fig 1 of L116 implied that at up to 39 mph, EDBTE of a Crosti 9F exceeded ITE, which is technically impossible, because ITE exceeds EDBTE by the LR at every speed, and LR is always positive. (Reason for giving this at 39 mph is given below).

It was admitted in L116 that this shape of LR in Fig 1, applying solely to the 9F Crosti, was wrong. But the further admission is crucial, that this was not a one off problem, that it had occurred in <u>all</u> the Rugby/Derby tests, since the inception of the testing procedure, and that it had been known to exist throughout the period. It does not say a great deal for the scientific acuteness and ability of the testing officers that it had not been cured at that inception.

Table 1
Data Given in Figs 1 to 3 of L116:

	20	30	39	50 mph
1 “incorrect” LR of Crosti 9F from Fig 1 of L116, lbs	2985	2631	2518	2461
2 “correct” (a) LR of Crosti 9F from Fig 2 of L116, lbs	1895	2164	2518	3027
3 Apparent Error, (1) – (2), lbs	1088	467	0	-566
4 “correct” (a) LR of standard 9F from Fig 3 of L116, lbs	1448	1643	2060	2659
5 Higher resistance of Crosti 9F compared with that of standard 9F, (2) – (4), lbs, both declared “correct” (a)	447	521	458	368

Here I follow the wording of the authors of L116. I do not believe that the Derby testing officers or the authors of L116 ever knew the correct LRs.
Procedures set out in L116 were supposed to correct for the errors, and give the correct LR for both standard and Crosti 9Fs, as in lines 4 and 2. Exactly how that operated, how it yielded the appropriate EDBTE and with that LR, is not explained in L116. All that is said is that correct answers were obtained. It is definitely not scientific to fail to describe and explain the principles

of the correction. I discuss that below. But using the correcting mechanism devised by the testing officers, the correct and incorrect LR intersect at 39 mph, lines 1 and 2.
The “incorrect” LR of the Crosti from 20 to 50 mph as declared in Fig 1 and line 1 of Table 1 above was approximately 2985 – 17.4(actual mph - 20) lbs, ie declining with speed (a straight line effect, used for illustration)[1]

As declared in Fig 2, the supposedly “corrected LR” was approximately 1895 + 37.7(actual mph – 20) lbs, ie increasing with speed[2]

A crude correction without any basis for the correction to [1] to give [2] is obtained by [1] –[2] or

{2985 – 17.4(actual mph - 20)} – {1895 + 37.7(actual mph – 20)} lbs

=1090 -55.1 (actual mph -20), which is the equivalent of the correction given in line 3 of Table 1.

Thus are simple correction mechanisms devised. That given here provides no explanation for why or how the error in EDBTE arose, and there is no basis in them for claiming that the correction is correct.

How the Defective Measurements Occurred

Report L116 does not give the EDBTE figures applying to the supposed corrected figures. But it is possible to use the data in Test Bulletin 13 on the 9Fs, Figure 11, and in Figure 11 of L116 to obtain some comparison. As an error in EDBTE requires an error in ITE of a slightly greater magnitude, the error in ITE in the road tests for the Crosti 9F was about 4.5% at 20 mph, 2.7% at 30, nil at 39 mph, and 5.1% in the opposite direction at 50 mph. The reason for using ITE in this comparison will emerge shortly.

It is also to be asked why a Crosti 9F was used in the identification and presentation of the problem. It was a peculiar engine from the LR point of view, and there were no other Crosti engines on the system. It was also a poor choice when the LR of the Crosti was untypically high at any rate of working. The Crosti engines had a higher LR than the standard 9F, on account of the high back pressure resulting from the highly restricted and primitive blast nozzles, the result of the need to draw the combustion gases through the boiler and the preheater. The higher LR accords well with the back pressure, as shown by the Perform program. The frequently quoted idea that the resistance of the Crostis was high because they had weak frames is unsubstantiated; those quoting it as the reason for the high LR need to consider where the effects of the higher back pressure were felt, and the lack of detection of the effect of weak frames, also whether weak frames increase LR. The back pressure effect did not disappear. In L116, the LR of the Crosti is higher than the standard 9F by 450 to 500 lbs at 20, 30 and 39 mph but only about 370 lbs at 50 mph - see line 5 in Table 1 above. (That weak frames were even suggested at Rugby for the higher LR is another reason for my doubting the scientific competence of the officers concerned; at least they noticed the higher LR of the Crostis, before they declared that all LRs were wrong).

L116 does not say how the erroneous LR and by implication, erroneous EDBTE arose in all these joint Rugby/Derby tests, even the data for the individual tests where ITE showed the same absurd characteristics as those for the Crosti 9F (as in Fig 1 of L116 and Table 1 above). Although it was EDBTE which was the immediate or arithmetical cause of the erroneous LR, it was wrong because an ITE was wrong, and that ITE was wrong because the instruction to the driver at what speed and cut off to run was wrong, or the arrangements for interpreting the BPP differed between the observation at Rugby and that on the locomotive on the road in the

CRT. The intended speed for the test was also advised to the operator of the MTUs in the Dynamometer Car.
Indeed, as itemised in L116, the testing officers took steps to check whatever might have led to the absurd answers. The Rugby and Derby indicators were checked and found to give identical powers (powers is the word used in L116, but it is TE which is given by a dynamometer). The dynamometer was checked. The steam rate measurement was considered. The officers found that Q could differ with speed both on the road and in tests at Rugby, but there was no proof that such was the case in anything they did. Their analysis of this data was defective and biased the results of their thinking towards the idea that there was a speed effect. This defective analysis is discussed below, because it led to an erroneous method of amending (or intending to amend) the historic data to produce accurate EDBTE and LR (or intended method – it is not clear that such desk-top corrections ever took place). The ideas put forward prove nothing of consequence, and variation in Q could not be detected from the water rates (although the difficulty of detecting small changes in water rates is emphasised). A concomitant problem is not mentioned. It is assumed that the driver could alter the CO as needed to achieve a certain BPP. A few simulations using the Perform program show that the necessary adjustments to CO to maintain a BPP were minute, not physically possible. (The report R13 on testing Duchess 46225 admits, however, that on the Plant, the CO was moved each time to a definite notch and the speed adjusted to give the correct Q, presumably by regulator adjustment, but on the line a definite speed was used for each step and the CO adjusted accordingly; presumably in drawing out the results for Report R13, considerable interpolation was needed to draw the ITE and EDBTE relationships at the usual tens of MPH and thousands of lbs of Q. That well may have been necessary in reporting results for all Test Bulletins). Presumably where the regulator was used to make the adjustments mentioned, the effect on Steam Chest Pressure would have been very small.
More importantly, however, the data recorded specially to show the relationship among Q, BPP and V was wrongly analysed and interpreted. No speed effect on the relationship between Q and BPP was present in the data for a 9F, nor in data with the same items for a Royal Scot. The relationship between Q and BPP was unaffected by speed, as should have been expected from first principles. See below.

Indicating the CRTs

In L116 it is said (as above) that the indicator used on the CRTs gave much the same readings of ITE as did those given by the Rugby indicator. That leads to the question, if the locomotive was indicated on the CRT, why was the BPPabs of any importance, why was the practice continued of trying to replicate the Rugby BPP in the CRT? The only reason which occurs to me is to connect absolutely the Rugby ITE and Q values with those on the road, to ensure that the Q and V for ITE measured at Rugby were exactly the same as those measured on the road, thereby allowing EDBTE to be measured with the same Q, V, CO etc as was the ITE, as is usual in the BR Test Bulletins. Such perfect correspondence, if the reason, is an extreme action - if the road test ITEs and EDBTEs were made at different Qs from those at Rugby, it is always possible to interpolate. Indeed, ITEs obtained on the road must be superior, in view of draft and exhaust effects, to those on a Stationary Plant. In that case, the Rugby results could have been put to one side.

It is remarked in L116 that a comparison was made between Rugby ITE and Derby DBTE by running the engine at equal V and CO, which gave LR without reference to Q. That tells anyone checking what Derby did almost nothing because identical V and CO mean identical Q. If it is thought that Derby needed to explain a V effect, then experiments would have been needed at each speed separately. Indeed there was some of that – see below.

It is obvious, however, that if ITE and EDBTE led to LR results which were obviously wrong (as in Table 1 and by admission, many other tests), then the various ITE results were not compatible, a problem of method and measurement

Test Bulletins Left Uncorrected

The Test Bulletins recording the joint work of Derby and Rugby are listed below. These were invalidated by the problems revealed in L116. Of course L116 contains the following paragraph in the Foreword:

With regard to previously published curves, however (presumably ITE and EDBTE curves in Test Bulletins below) it is considered that the discrepancy is not sufficient to invalidate their use for train timing and similar purposes. No information is given on the size and nature of the discrepancy anywhere in L116 (but see my rough estimates above).

The following Test Bulletins were undermined, those based on Rugby work and Derby CRTs:

Bulletin 2, B1 61353 1950
Bulletin 5, Standard 7 1953
Bulletin 6, Standard 5MT 1952
Bulletin 10, SR 8P 35022
Bulletin 13, Standard 9F 1959, work done up to 1957
(Last steam testing at Rugby 92250, 9F Giesl, no Bulletin
LMR 8P 46225, no Bulletin, but reports R13 and L109 mention the L116 method of adjustment (see below)
[Equipment dismantled 1970, plant demolished 1984]
No attempt to correct these reports is known to have been performed. Despite Fig 16 in L116 attempting to show how the earlier work could be corrected, nothing in L116 shows how a corrected EDBTE could be obtained, even what the error was.

The Proposed Correction and Underlying Research

Some background from L116:

It was concluded that the problem arose from using blast pipe pressure as a steam flow meter without compensating for varying road speed. In the revelation of the odd LR shape problem, it is said early in L116 that the difference between ITE from the LTS and the EDBTE obtained from road tests, which is the Locomotive Resistance (LR), had not been acceptable in shape, that the discrepancy was large and consistent. It was said that it was believed (ie not shown to be the case) that the DBTE resulting from the procedures used was correct in the middle speed range, too low at low speeds and too high at high speeds. It is not stated how this was known, indeed, given the problem, how it was possible to know it. Similarly, in point 9 in the report, it is concluded that the steam rate for a test applied only at the mean speed for the test. This is not sensible if things worked properly. How does the instrumentation know what range of speeds will be tested and how many tests conducted at each speed, ie that the results can be correct for the mean? The mean will vary with the tests conducted.

At Swindon, ITE was measured on both the plant and the road (see Bulletin 1 p 5). Although the Bulletins claim that there were no significant differences in boiler and cylinder performance between the plant and road tests, it is generally considered that the plant tests were undertaken to determine boiler characteristics, and that both ITE and DBTE data used in reports prepared by Swindon were obtained on the road. As they were both subject to the same effect of V on P where P was used as the steam flow meter, they should give reasonable LR.

Discovering the effect of V on BPP at a given steam rate from plant tests to adjust the results of road tests requires correspondence between plant and road in all circumstances. It is doubtful that such correspondence could be achieved. The ability of a given BPP to bring about a given evaporation can be expected to differ on the plant and on the road in ways not considered in the report. There are at least two reasons for this. The first is that the scooping of air into the front damper and under the fire will reduce the need for draft for a given evaporation rate compared with a stationary locomotive on the plant. The same will apply to air drawn from the sides of the ashpan. (If the front damper is closed and underfire air is drawn solely through the rear damper, the draft requirement on the moving locomotive will be increased, to overcome the slight vacuum behind the rear of the ashpan.) The second is that the moving locomotive will create a small vacuum at the chimney top, which will provide a little draft, compared with a stationary locomotive. Both of these effects can be expected to increase with speed. Tests on the plant to establish the effect of speed on evaporation for a given BPP will not detect these two effects. The third possible consideration is that the resistance of the fire cannot be expected to be necessarily identical on the plant and on the road at a given steam rate, on account of firebed depth differing on account of fire management requirements and duration of the run, and different packing down of the burning coal. A given BPP on the road could lead to higher or lower evaporation than on the plant, even if all other factors were made identical.

Surprisingly, it was believed that the incorrect results could be corrected, as a desktop mathematical exercise. To correct something known to be wrong, it is necessary to discover what was wrong and why, and to know the correct answers. None of this applies in this case. Usually, it will not be possible to undo what has been done.

Derby therefore put forward a method for correcting the defective measurements of EDBTE in all the published Bulletins applying to Rugby/Derby tests, or new tests to be done, incorporating these modifications. (Bulletin 13 on the 9Fs was published in 1959, but was based on data gathered before 1957, and Bulletin 20, published in 1960, on the rebuilt Merchant Navy engines) included road test data only, was not tested on the pre-1957 testing system. See the extent of the effect of their correction for the Crosti 9F in Table 1 above. The modification was to develop a process and formula which changes the Q data.

To have any hope of making such a correction, the reason for the error has to be known. As above, it had to be a question of method and measurement. As these factors are likely to differ in effect from test to test, the correction task would seem hopeless. They considered three possible bases – adiabatic heat drop, compensation for change in density, and compensation for speed effect on the BPP/Q relationship. They could not find any thermodynamic reason, which probably meant there was none, and picked, in speed effect, something which did not exist, as I show below. It is true that among the road test data, they had examples of tests where the result differed with the speed, eg by direction. These tests drop out as a basis because they were not comparable with the principle of the testing, constant Q, V and BPP. One wonders if such non constancy by direction in a test was not the reason for the error.

The equations in Fig 16 of L116 do not demonstrate a basis for altering Q, simply playing with the concepts “left over”, not used so far in trying to explain the anomalies. The Derby test officers had observed some peculiar effects of different speeds, which is perhaps why they thought speed was playing a part in explaining the determining the influence of BPP on the Q passing the Blast Pipe. They did not think that through. See my tests below. Note also that where they claimed that the system worked, that a correct LR, or correctly shaped LR, results, there is no case where a correct LR comparator exists. Nor the basis for declaring how a LR would be established from first principles. No prospects for science there.

The officers considered that there must be more to it, however. They considered that the reason for the error was that their assumption held over the whole seven years of testing that Q varied only with BPP was wrong, that the relationship between Q and BPP was affected by V. They therefore sought a relationship among Q, BNP and V. they also believed that the error in procedures and/or measurement were in the EDBTE, which was measured by Derby. But that also depended on ITE registered on the road.

Although L116 was partially accepted and some adjustments made with it, there are memoranda within it from D R Carling, Supervising Engineer of the Rugby plant, and E S Cox, Chairman of the Locomotive Testing Committee. Both have considerable reservations about the report. Both note that no explanation is offered for the supposed effect of V on the relationship between BNP and Q, Cox saying as much as that the variation with V was not established scientifically. Cox believed that the range of experimental data was to a large degree the range of experimental error.

Carling said that on the whole the data examined until then could only be regarded as supporting the method proposed in the report as a workable method for use where necessary, without any pretension to confirming it as a fundamentally correct method.

Neither of these gentlemen called in aid S O Ell or his staff, who were in charge of testing at Swindon. The CRTs conducted at Swindon depended on duplication of the results of boiler and efficiency tests conducted on the Test Plant at Swindon. Ell claimed that the road tests confirmed the plant tests. Ell was surely the person most likely to discover the defect in the Derby practice.

There are more and better reasons for not accepting the correction method. The authors of L116, presumably Rugby officers, were not content with the conclusions and intentions of L116. On p 8, under (2), Joint Analysis of Results, they say "It is desirable that test results should be pooled, so that Indicated and Drawbar Characteristics can be constructed together. Hitherto, the curves have been drawn up entirely independently, and small differences in the methods of construction have added to the difficulties of reconciliation."

In similar vein, they go on "(3) Elimination of Differences in Test Procedure. Testing methods have been developed at Rugby and Derby separately, and the results of tests at both centres are valid for the respective conditions under which the tests were made. It is desirable however, if agreement is to be achieved with joint tests, for local differences to be eliminated as far as practicable . In this connection, it must be mentioned that the mean blast pipe pressure curve established at Rugby cannot be reproduced when a locomotive is subsequently subject to tests on the line. A re-calibration of the orifice meter was therefore necessary , and this work was to be undertaken while the main tests are proceeding. It is considered that anomalies of this nature could be readily eliminated by close co-operation with regard to choice and siting of instruments".

These comments are indicative that the joint tests did not agree for seven years because the procedures were sloppy, and did not lead to automatic reconciliation of results.

Experimental Data on 9F

In L116 the experimental data on Q, V and BPP used in formulating the correction process are presented in Figure 11, ten observations at 15 mph, five at 30 and five at 50 mph. I have transformed these data into Table 2. In Fig 13 of L116, appears another set of BNP against Q for 92050 with differing figures. To increase the number of observations, especially at 30 and 50 mph, the data in Tables 2 and 3 below have been combined into one series, to give 18 observations at 15 mph, ten at 30 and nine at 50 mph, a total of 37. The results are very little different, both in actual answers and goodness of fit. (the comparison was with the 20 observations of Table 2 and the 37 of Tables 2 and 3).

Table 2 Data in Fig 11 of L116 on Blast Pipe Pressure, V in mph, and Q, 9F 92050

BPP gauge lbs/sq in	Q lbs	V mph

1.6	11900	15
1.95	13200	15
2.15	14000	15
3.2	16100	15
3.4	16700	15
4.75	19000	15
4.83	19800	15
5.5	20200	15
6.6	21600	15
7.1	22400	15
2.8	15600	30
4.55	19000	30
6.6	22300	30
7.1	23400	30
8.5	24800	30
2.5	15000	50
3.55	17400	50
4.6	19600	50
5.8	21400	50
7.1	23300	50

Table 3 Data in Fig 13 of L116 on Blast Pipe Pressure, V in mph, and Q. 9F 92050

BPP gauge lbs/sq in	Q lbs	V mph
1.972	13122	15
2.018	13900	15
3.236	16144	15
3.388	16749	15
4.786	18281	15
5.623	20277	15
6.607	21135	15
7.08	22491	15
2.818	15596	30
4.571	19055	30
6.025	21528	30
6.607	22284	30
8.414	24717	30
2.4547	15066	50
3.3884	17378	50
4.5709	21478	50
7.0795	23227	50

In the same Figure 11 of L116 are freehand lines which are meant to represent the relationships among these items, judged to be:
At 15 mph $Q = 9900\ P^{0.415}$
At 30 mph $Q = 10{,}200\ P^{0.415}$

At 50 mph Q = 10,400 $P^{0.415}$

Where P is blast pipe gauge pressure. It is argued in L116 that as these lines are parallel in non-logarithmic form, the index or power can be made the same for each line. The lines in non logarithmic form are not straight, and are therefore cannot be parallel. Nor is the slope of each the same in non-logarithmic form (change in BPP divided by change in Q). (This was a rich claim in any case with only five observations in Fig 11 at each of 30 and 50 mph). They are *in part* the same distance apart, in log form because the centre of the curves of each at that point has been moved a certain distance. A mathematically correct analysis of the data of Both Figs13 and 15 together gives:

At 15 mph, Q = 55$BPPabs^{1.964}$
At 30 mph, Q = 121.5$BPPabs^{1.705}$
At 50 mph, Q = 95$BPPabs^{1.798}$,

Which are mathematically and statistically respectable, whereas the L116 figures are not.

Analyses of 9F Data

There are three important defects in this work. First BPP is measured at atmospheric or gauge pressure, whereas it should be in pressure absolute, as even an apprentice scientist should have known. Second, the three curves in Fig 11 from which Table 2 was drawn above were fitted by freehand, with the initial pressure for each speed picked by eye. More importantly, the data are fitted to lines for the speed at which the tests were made, 15, 30 and 50 mph, and the curves for each speed drawn by eye. That means that the relationship with V is assumed to be that drawn in Fig 11.

Regression of this very same data both with and without its relationship to a speed being assumed finds the effect of V on the relationship between Q and BPP to be in effect nil. Regression also avoids guesswork and has the enormous advantage of giving as a test statistic whether there is any significant (statistically significant) difference in curvature or constant by speed. There is not (see below), which means that eyeing up the gradient and constant introduced a serious bias. Fortunately, it is possible to do this analysis properly, at least in principle.

Third, there are insufficient observations at each of 30 and 50 mph (ten each) to analyse the effects at those speeds properly. It is also desirable to analyse the data in such a way to see whether the implied assumption on the part of the testing officers that the speed effect differed by speed, an assumption for which no reasons are given.

Regressions obtained from these data follow. The physical act of passing a given quantity of steam through a restricted nozzle should have BPPabs on the vertical axis as the result, and Q as the cause, on the horizontal axis. As however, the system is used as a meter, the reverse arrangement of the data is used, ie Q on the vertical, BPP on the horizontal.

The regression results which follow are all in terms of BPPabs, ie in absolute pressure, and speed in RPM.

The following are the results of regressing the useful permutations of BPPabs, Q and RPM, the figures or values in Tables 2 and 3:

1 BPP abs = 0.126$Q^{0.52}$.$RPM^{-0.025}$

Effects and comparisons: A 10% increase in Q, RPM constant, leads to a 5% increase in BPPabs.
A 10% increase in RPM, Q constant, leads to a 0.25% increase in BPPabs (ie a quarter of one per cent) .
If the RPM term and data are eliminated, ie the regression is of Q on BPPabs, the best fit equation of BPPabs on Q scarcely changes. It becomes BPPabs = $0.133Q^{0.51}$
2 Q = $65.BPP^{1.83}.RPM^{0.05}$
Effects and comparisons: at a Q of 16,000lbs, if there is a 10% increase in BPPabs, RPM constant, Q rises 19%.
A 10% increase in RPM, BPPabs constant, leads to a 0.78% increase in Q (ie four-fifths of one per cent)

If the RPM term and data are eliminated, the best fit equation to Q on BPPabs changes only a little from the above with RPM included, to Q = 74. $BPPabs^{1.87}$. A 10% increase in BPP abs at a Q of 20,100 lbs, leads to a 19.5% increase in Q

This equation without the RPM term (ie Q = 61. $BPPabs^{1.9}$) is that used successfully in Swindon testing with BPP as the meter of Q. It was also used by Derby, but not successfully). The equation with the RPM as an extra term, shows how the Q/BPPabs relationship is unaffected by V, ie by RPM) (ie $65.BPP^{1.83}.RPM^{0.05}$).

3 Speeds considered separately as in L116 (as above)

At 15mph, 18 observations, Q = $43BNPabs^{2.06}$
At 30 mph, 10 observations Q = $137BNPabs^{1.66}$
At 50 mph, 9 observations Q = $85BNPabs^{1.33}$
(Compare all speeds together, as in 2 above, Q= $65.BPP^{1.83}$, or Q = 74. $BPPabs^{1.87}$.)
These equations differ vastly from those in Fig 11 in L116.
L116, because they are not based on freehand or by eye curve fitting, and because they employ BPPabs and not BPPgauge, represent the best statistical (scientific) fit to the data. The ratios involved with any change in BPPabs are much smaller than those used in gauge or atmospheric pressure, as in Fig 11 of L116. As before, five observations at each of 30 and 50 mph are totally insufficient for investigation, and ten only on the verge of sufficiency.
The coefficients on RPM are always very small. Q is always a large number in thousands, and RPM always a small number in comparison (50 mph, 280rpm, or less). The effect of V is very small indeed, as in the notes above about the effect of 10% increases in determining variables.
This data from Figs 11 and 13 of L116 does not contain and cannot reveal a speed effect on the relationship between Q and BPP, because there is none. (Statistical tests revealing probabilities available).
There is similar data on a Royal Scot, source now forgotten or lost. The Scot data have been analysed similarly to those of the 9F.
Table 4 Data on Blast Nozzle Pressure, Q and V in mph, Royal Scot

V mph	rpm	BNP, abs pressure, lbs/sq in	Q lbs	ln BNPabs	ln RPM	lnQ
20	83	15.6	11340	2.747277	4.418394	9.336092
35	145.25	15.6	11510	2.747277	4.97801	9.350972
50	207.5	15.6	11590	2.747277	5.334685	9.357898

20	83	16.6	14980	2.809403	4.418394	9.614471
35	145.25	16.6	15260	2.809403	4.97801	9.63299
50	207.5	16.6	15510	2.809403	5.334685	9.64924
20	83	17.6	17640	2.867899	4.418394	9.777924
35	145.25	17.6	18000	2.867899	4.97801	9.798127
50	207.5	17.6	18390	2.867899	5.334685	9.819562
20	83	18.6	19800	2.923162	4.418394	9.893437
35	145.25	18.6	20230	2.923162	4.97801	9.914922
50	207.5	18.6	20750	2.923162	5.334685	9.940302
20	83	19.6	21660	2.97553	4.418394	9.983223
35	145.25	19.6	22160	2.97553	4.97801	10.00604
50	207.5	19.6	22790	2.97553	5.334685	10.03408
20	83	20.6	23310	3.025291	4.418394	10.05664
35	145.25	20.6	23870	3.025291	4.97801	10.08038
50	207.5	20.6	24600	3.025291	5.334685	10.1105
20	83	21.6	24790	3.072693	4.418394	10.1182
35	145.25	21.6	25420	3.072693	4.97801	10.14329
50	207.5	21.6	26240	3.072693	5.334685	10.17504
20	83	22.6	26160	3.11795	4.418394	10.17199
35	145.25	22.6	26830	3.11795	4.97801	10.19728
50	207.5	22.6	27740	3.11795	5.334685	10.23063
20	83	23.6	27440	3.161247	4.418394	10.21976
35	145.25	23.6	28150	3.161247	4.97801	10.2453
50	207.5	23.6	29170	3.161247	5.334685	10.2809
20	83	24.6	28620	3.202746	4.418394	10.26186
35	145.25	24.6	28150	3.202746	4.97801	10.2453
50	207.5	24.6	29170	3.202746	5.334685	10.2809

1 Regressing Q on BPabs and RPM:
$Q = 62.BNPabs^{1.87}.RPM^{0.047}$
If BNPabs rises 10%, RPM constant, Q rises 19.5%; if BNPabs rises 10%, RPM constant, Q rises .044% (ie less than one twentieth of one percent).
2 Regressing BNPabs on RPM and Q
$BNPabs = .16RPM^{-0.234}.Q^{0.50}$
If RPM rises 10%, Q constant, BNPabs falls 2.36%; if Q rises 10%, RPM constant, BNPabs rises 7.4%.
As with the 9F above, the **Royal Scot data shows no relationship between Q and RPM, and that speed has no effect on the relationship between Q and BPPabs,** in complete contrast to the hypothesis of the testing officers.
3 Speeds considered separately as in L116 (as above)

At 15mph, ten observations, $Q = 43BPPabs^{2.06}$
At 30 mph, five observations $Q = 137BPPabs^{1.66}$
At 50 mph, five observations $Q = 85BPP^{1.33}$

These equations differ considerably from those in Fig 11 in L116, because they are not freehand curve fitting, and because they employ BPPabs and not BPPgauge, and because they represent the best statistical (scientific) fit to the data. The ratios involved with any change in BPPabs are much smaller than those used in gauge or atmospheric pressure, as in Fig 11 of L116.

Use of gauge pressure instead of the correct absolute pressure will have considerably distorted any relationships including BPP, including the idea that V is needed in determining a relationship between Q and BPP.

The coefficients on V are always very small. Q is always a large number in thousands, and V always a small number (50 mph, 280rpm, or less). The effect of V is very small indeed, as in the notes above about the effect of 10% increases in determining variables above. This Royal Scot data does not contain and cannot reveal a speed effect on the relationship between Q and BPP.

Analysed via Perform

There is no effect of V on the relationship between Q and BPPabs in the data, data gathered in the testing of 92050 and Royal Scot. Further, despite the claims in L116, the three lines, one for each speed, in L116 are not straight, nor parallel (not that those charactristics matter provided a good fit is obtainable). Fig 15 does not connect C to V; while one can be graphed against the other, a constant remains a constant. Something else must have been in mind. That can be observed in test applications of the Perform program, as could be done by considering indicator diagrams. What can be learnt from Table 5 is that at low RPM, the exhaust from each stroke is separate, but as rpm rises, the exhausts merge, to give a continuous flow of the Q. This can be observed through trials of Perform.

Table 5

Perform Exercise to Show Effect of V on Relationship between Q and BNP abs, Standard 9F

V	Q	CO	BNP Gauge*	Inlet Steam Temper-ature, ℃	Release Pressure	ITE Perform	ITE, Table 10 of Bulletin 13	Ratio ITEs, Perform to Bulletin 13
20	12000	19.2	1.9	316	40	13,700	14,400	0.95
40	12000	13	1.71	316	52	8040	7200	1.12
60	12000	11	1.7	316	52	5420	5000	1.08
20	18000	33.9	4.75	366	91	22,800	22,000	1.036
40	18000	21.7	4.45	366	52	12,900	12,000	1.08
60	18000	18.2	4.36	366	39	8,850	8,100	1.09
20	24000	44.9	9.2	377	110	27,600	28,000	0.99
40	24000	29.1	8.67	377	78	16,400	16,000	1.03
60	24000	24.5	8.45	377	52	11,300	11,000	1.03
20	30000	57.1	16.3	393	143	31,800	32,600	0.98
40	30000	36.6	15.1	393	84	19,200	19,000	1.01
60	30000	30.8	14.8	393	72	13,300	13,000	1.02

*Perform works in pressure absolute and automatically converts gauge pressure to absolute. Absolute pressure is simply gauge pressure plus 14.6 lbs/sq in.

In each set of 20, 40 and 60 mph at a certain Q, BPP gauge is close, but falls a little from 20 to 60 mph. In each set of three, BPP gauge is highest at 20 mph, because at 20 mph, the BPP discharges are more distinct than at higher speeds, but decline from 20 to 60 mph. The Perform ITE is close to the Bulletin 13 figure. There is no evidence here for a speed effect on BPP at each speed at each Q.

This table can be rearranged to have sets of three for Q and for CO.

Speed in Normal Test Results

In the Test Bulletins, ITE and DBTE are shown in Figures or Graphs against Speed, with the following shown across the Figures: Q, CO, fuel and efficiency. These performance maps clearly show that ITE varies at a given Q with speed, that as V increases, the ITE curve declines with speed.

That has to be. As speed increases less steam is available per stroke, and ITE from a given Q falls. The issue in L116 is different. The issue in L116 is whether during a given test, changing V affects the relationship between Q and BPP. Fig 11 in L116 is drawn to imply that it does. The data collected to test that point for the 9F and Royal Scot, and the simulations with Perform show that V does not make the slightest difference to the relationship between Q and BPP, that it is BPP which affects Q, unaffected by V.

The officers were not clear about this distinction. It is said on p 5 of L116 that it was first observed with B1 61353 that during the course of a day's test running from Carlisle to Skipton and return, the Q produced by a particular BPP during the outward run could not be accurately repeated on the return. The only difference of any significance between the two test runs was that the overall average speed was lower on the return, owing to the nature of the test route. They refer to average speed. The whole aim of the MTUs was to keep speed constant for a given Q. The aim would have been to select the speed to be run for a given Q on the ruling gradient of 1 in 100, and to add to the resistance to maintain the speed where the gradient eased. The Q and V were to be maintained for the whole test section, downhill as well as up. The average speed was of no significance. Nothing is said about how much downhill running was converted to 1 in 100 (or other desired gradient) uphill, but it would have been useful for testing for significant periods at the higher speeds. Average speed is not of interest for either CRTs or the effect of V on the relationship between BPP and Q.

There is also confusion on pp 5 and 6 of L116, and in Figures 7, 8, 9 and 10 of L116. There is mention of constant BPP, mean Q for a test applying only to power developed at the mean speed for the test, adjustments being needed at all other speeds. There is obviously mistaken thinking and measurement here, because the data are those which appear in the EDBTE diagram in Test Bulletins, Q and EDBTE on the vertical axis, V on the horizontal axis, with a series of lines showing the relationships among those items (usually with CO and efficiency superposed). There is nothing to do with the effect of speed on the relationship between Q and BPP here. Nothing is discovered through the idea that effects differ at high and low speed tests, so called.

Experimental Data on Duchess 46225
The Last Attempt to Get it Right, Rugby and Derby?

This was the testing of Duchess 46225 in 1958, which involved use of the steam flow meter to ensure constancy of Q on the road, and the MTUs. There was an extensive time gap between the Rugby tests of this engine and those on the road. Further the valve heads were set back for the road tests to even the tractive effort produced from the front and rear ends of all cylinders; that led in turn to a given CO giving a higher ITE at any speed, ie plant and road ITEs could be expected to have differed for a given Q. On the other hand, the road tests were separately indicated at all cylinder ends (the same applied to other of the joint tests, but Rugby results were preferred for indicated results in those cases. L116 had just been published, identifying the defect, and proposing a solution. Efforts were made to confirm the discrepancy between ITE on the plant and on the road, but the results were not conclusive. For all that, report R13 says that the L116 method of adjusting CRT results was used and brought agreement between the two types of test. In addition, LR was measured directly on line as the difference between ITE and EDBTE (but very poorly presented – in specific terms with exact weight indeterminant, wind effects unknown, the statement "average service conditions" undefined; even broad values or conditions still leave an unusually low rate of increase with speed. Although water consumed (Q) was measured incrementally on a time basis, the BPP was used to measure the instantaneous Q. For all the care taken, plant and road ITEs differed by speed, as in Table 6 above.

Neither R13 and L109 give any detail on the application of the L116 method of reconciliation. No reasons are adduced for the ability of the method to reconcile the results from the two types of test; there was none.

No progress was therefore made in comprehending the problem of wrong EDBTE values encountered by the Derby testing people, explaining it, and finding a solution, in these last tests. No correct cure.

L116 admits the error and the period over which it prevailed. It is singularly deficient in not saying what went wrong and why. There is the idea that the error was the result of failing to take into account the effect of speed on the relationship between Q and BPP, but the above analyses show that idea to be fanciful. In particular, allocating observations into speed bands vastly exaggerates the effect of V in the results of the analyses. For all that, it is obvious what was going wrong. The method was not connecting ITE at Rugby with ITE on the road for, so far as we readers sixty years later can tell, a given BPP advised to the driver in a CRT. Even if the ITE on the road for a given Q and V was equal to the Rugby ITE, the EDBTE for that ITE (hence Q and V) was not measured or calculated properly.

ITE on the road and on the Rugby plant ought to differ for reasons already given, to do with draft on the fire and exhaust effects on the road compared with the plant, and the road figures ought to be preferred. That does not really answer the question of what happened in Derby controlled CRTs. L116 does not give the road ITEs, except in a very indirect way for the Crostis. (The fig 11 data in L116 is wrongly presented and analysed, as discussed above). The exception is in yet another internal report L109, in Fig 20. This shows ITE recorded by Rugby and Derby for various Qs from 16,000 to 38,000 for a Duchess at speeds from 20 to 80 mph. At a Q of 28,000 (one of many Qs available), Derby ITE differs from Rugby ITE as follows:

Table 6 ITE Recorded by Rugby and Derby at Q of 28,000 lbs/hour Duchess 46225, 1956

mph	Rugby ITE	Derby ITE	Rugby ITE/Derby ITE
25	25,500	24,000	1.0625
30	22,400	21,700	1.032
40	17,700	17,400	1.017
50	14,600	14,600	1
60	12,200	12,500	0.976
70	10,400	11,000	0.945
80	9,000	9,800	0.918

Source, Table 20, Internal Report L109. The road tests (Derby figures) were conducted March to May 1956.

Here remerges the pattern of Table 1. The Derby figure is the lower from 20 mph to 50, and the higher from 50 to 80 mph, with results equal at 50 mph (39 mph for the Crosti 9F). Note above that the indicators were compared. So were the Qs (ie water consumption) and not found to be the source of error or explanation. Even if there was error in measurement of Q, it would be expected to be a constant quantity or proportion, not one operating in one direction below 50 mph and the other above 50 mph, and to different extents. Nor would it be expected that the ITEs would be equal at 50 mph. There is no measurement of EDBTE in this data, but if EDBTE were properly measured relative to Derby ITE, it would follow a similar pattern of ratios.

This data does not appear in R13, reporting the same tests of the same engine. But R13 says:

When the two sets of test results were first compared there appeared to be an even larger discrepancy between them as regards power output than there was between similar tests on the plant and on the line in the case of the (Class 9 locomotives). The extent of the disagreement was shown in Fig 20 of L109 (and in part in Table 6 just above).

Application of the methods (in L116) has, it is claimed, however, brought agreement of the two sets of tests within the normal limits of experimental error, having regards to the circumstances of the tests mentioned above (ie the time gap). This does not apply, however, because the correcting equation is wrong in principle.

The Duchess data on the Rugby plant and on the road are definitely not comparable. Between the tests at each place, the valves were set back to increase the work done at the rear end of each cylinder. Nevertheless, the pattern and extent of the ratio of Rugby to Derby ITEs, as in Table 6, could still not be explained. That is of course if any consideration had been given to why it could exist.

However, Report L109 states:

An attempt was made to determine whether the same blast pipe pressure produced different rates of evaporation under constant and variable conditions of speed respectively. The constant speed tests were carried out during the first two weeks, and difficulties encountered during the early stages of the tests (Effect not given) … prevented them being strictly comparable with the remainder of the tests. The results were therefore not conclusive. Despite which:

As regards the degree of reconciliation with the results obtained during the Stationary Plant tests, …..as on previous occasions, however, there is some discrepancy between the ITE characteristics established on the Stationary Testing Plant and the road. Results were of the type appearing in Table 6 above.

It then goes on ……"Tests will be carried out in the near future at Rugby to investigate this discrepancy." So only after testing had ended was the error to be investigated, and then only on the test plant.

So no progress was made in understanding the difference between road and plant ITEs from a given Q, even at the very end of steam testing.

Unscientific Presentation of the Results of the Derby tests and the Supposed Correction Procedure

As the commission of the error was so long lived, its effect was so unusual and gross, and the correction procedure was of such doubtful validity, a lot more explanation should have been given than is present in L116. The following would be expected:

1 Showing the Error – about 30 examples of what were meant to be corresponding Rugby and Derby results, the Q, BPP, V of the test, ITE, EDBTE and any Vs which might have affected the BPP/Q relationship. In particular, additional characteristics of the Derby ITE and EDBTE results, especially such as Derby and Rugby ITEs which are the same at some central speed but which are different at other speeds, and to increasing or decreasing extents from some central value.

2 Application of the Intended Correction, in particular the application of Figure 16 of L116. What adjustments are made to the Derby Q for road ITE and DBTE tests. Then, for a given recorded erroneous road ITE and EDBTE, the source of the corrected ITE and EDBTE (what is their source without running special tests; were the corrected values interpolated from other

data, and if so, what? Are there examples of whether during a given test, changing V affects the relationship between Q and BPP. Even more basically, it cannot be expected that variations in Q on the basis of the correcting equation can be correct . how is it supposed to produce what it is said to do. The ITE and EDBTE developed on the road should be derived from accurate measurement, not an invalid formula.

3 Results of the Correction Made – the different Q, and the associated road ITE and EDBTE; where did they come from, how do they fit into a continuity of ITE and EDBTE, ie the results of the corrected Q and associated ITE and EDBTE, for both Rugby and Derby.

4 The LR of the loco for which these adjustments were made and what was the comparison locomotive, and how its LR was obtained. That and any easier and more accurate tests, such as road tests run at a constant speed and CO for ITE, EDBTE and LR.

The Correction Equation

This is of the form $Q = CP^n$. Its derivation is not explained, either what it is intended to do, nor its origin. There is ready comparison with the equations derived above from the research data for the 9F. The conclusions reached, however, are very different. P is BPP, which is probably in gauge pressure, whereas it should be in pressure absolute. C varies with V, according to Figure 15, from 99 at 15 mph to 104 at 50 mph, or by a ratio of 1 to 1.05. That is the ratio of the constants in Figure 11, remarked upon above as a bias towards a speed effect. In my regressions, across all speeds together, the value of this constant is 57 with a speed term present, or 61 with no speed term present (as above).

In the L116 correction equation, the index on BPP is 0.415 in all circumstances. By the regression of the test data on which it is based, the index on BPPabs is 1.9, whether a term for RPM is included or not, a vastly greater influence of BPPabs than the index on BPP in the freehand L116 equations.

The correcting equation is therefore Q= (99 to 104, depending on speed)$BPP^{0.415}$. As the regressions of the same data show there is no dependence on speed, a conclusion confirmed by the Perform analysis, and no explanations or instructions are given in L116 (despite Fig 16) on the circumstances in which the correction equation is to be used and how, it should not be used to correct any data. And it cannot correct the old Derby data. In L116 not only is the correcting equation based on wrong thinking, it is based on wrong data and relationships.

The correct equation relating Q to BPPabs is $Q = 61BPPabs^{1.9}$, at all speeds and BPPabss. That is based on the test data collected for 9F, and applies to that class. See the analyses and results of the data above. Subject to the reliability of that data, it gives correct Q for any BPPabs for a 9F.

These two equations (99 to 104, 61 etc) are not correcting equations, but relationships between Q and BPPabs. The 99 to 104 equation is wrong, for reasons already given, and the 61 equation is the best fit to the data collected to research the V effect on the relationship between Q and BPP. L116 gives no rules for declaring that a Q is incorrect, although an LR might be judged to be the wrong shape. Even if a Q can be said to be incorrect, where does the correct BPP to obtain a correct Q come from, and from that the correct ITE and EDBTE. As Derby had made so many mistaken estimates of road ITE and EDBTE, it is not satisfactory to suggest that it will have a large notebook of observations for each engine tested, certainly not correct ones, because it had no way of saying which if any were correct. Nor should any further tests a Rugby be expected to solve the problem

Conclusions

The conclusions are not favourable to the Derby team. First, the results being anomalous over the whole testing period, it follows that the Derby team did not know how to achieve satisfactory road ITE and EDBTE results for a given Q despite years of practice. They wasted time in developing a supposed speed effect on the relationship between Q and BPPabs and V. The same applies to the supposed correction equation and procedure.

Different and more scientific expertise (including statistical) should have been called in early in the testing programme (before the end of the first year say) rather than tolerate anomalous results for years on end, ie better technical expertise on the generation and detection of correct data on the road of ITE and EDBTE, the function of the Derby Testing Section.

This paper first considered the large number of wrong results, admitted in internal report L116. It then considered how incorrect results could have arisen, and the modest research conducted to allow correct the incorrect results to be corrected, research which was extremely poorly applied. The officers concerned considered that their results were wrong because they had not taken into account the effect of speed on the use of the Blast Pipe Pressure on the metering of steam. In that they were mistaken, for there was no such speed effect. The correcting mechanism and equation they devised did not fit the data available, which led to wrong conclusions. They believed that they could conduct desktop corrections of results, but in that they were mistaken also, and no corrections of results proved possible. Nor did they perfect the testing and measurement, and to the end the Derby measurements of ITE proved defective, including that of a Duchess. Although Derby thought it had a system which could correct LR, it never explained where the comaparator locomotive came from. Checks were made of the apparatus and procedure, but the Derby errors were never corrected. This failure by Derby is surprising because testing procedure with similar intentions took place at Swindon and seemed to operate satisfactorily – it was Derby which did not succeed in measuring properly, and which devised correcting mechanism which was not a logical explanation for the mismeasurement which occurred.

The data available has been analysed much more soundly here than was done for L116.

Derby did not run its side of the joint Rugby – Derby testing soundly.

Some conclusions are drawn in the text on the peculiarities of some of the testing.

The conclusions of L116 should be forgotten, such as they are. That includes the supposed LR of a 9F.