

The relation between standard deviation and mean strength of concrete test cubes*

by A. M. Neville, M.C., M.Sc.(Eng.), Ph.D., A.M.I.C.E., A.M.N.Z.I.E.

Contribution by T. N. W. Akroyd,
LL.B., M.Sc.(Tech.), A.M.I.C.E., A.M.I.Struct.E.
(Constructional Services Ltd)

In his paper Dr Neville raises the question of the application of the standard deviation or the coefficient of variation to concrete control. I am not sure that he has realized fully the implication of all the statements given in his paper.

He states that, from his tests and from an examination of site test data, for the same degree of control the standard deviation is proportional to the mean strength. One could, of course, be rude and say this is balderdash because it appears to contradict our everyday experience, but before jumping to such a conclusion let us consider the following example.

A concrete is required having a minimum strength of 3,000 lb/in² at 28 days; and a water/cement ratio of 0.71; assume the standard deviation will be 600 lb/in² which, according to Himsworth⁽¹⁾, is equivalent to good site control. If the standard deviation method of designing the mix is used, the average strength becomes $3,000 + 2.33 \times 600 = 4,400$ lb/in² (water/cement ratio = 0.57).

Now suppose I carry out a trial mix in the laboratory and test the cubes at 7 days. According to Road Note No. 4⁽²⁾, I can expect a strength of 2,800 lb/in² but the 7-day figure for minimum strength corresponding to 3,000 lb/in² at 28 days is 1,900 lb/in² (water/cement ratio = 0.71). The standard deviation at 7 days is then 360 lb/in².

Suppose, however, that the cubes are not tested for twelve months. The strength corresponding to a water/cement ratio of 0.57 is 7,100 lb/in² and this, compared with the guaranteed minimum strength of 5,300 lb/in² (water/cement ratio = 0.71) which corresponds to the 28-day minimum of 3,000 lb/in², means that the standard deviation is 775 lb/in². Therefore, says Dr Neville, if out of a batch of concrete I make cubes and test them at different ages, I can expect the results to give me different values of the standard deviation. I agree with him. From experience I have found that for any one mix on the site the standard deviation is lower at 7 than at 28 days.

Now if it is true, as I have always been led to believe, that in the application of statistics to the quality of concrete the standard deviation could be used as a measure of the degree of quality control, it

is absurd to suggest that out of one batch of concrete the control which I exercise is "better than excellent", "good", or "fair", according to whether I test the resulting cubes at 7 or 28 days or one year, and yet Dr Neville goes on to state (p. 83) that "it would seem, therefore, that it is not the degree of homogeneity of the concrete at the time of making the specimens that determines the scatter of the strength, but merely the value of the strength at the time of testing". If, as is most likely, the standard deviation continues to be used as a measure of the site control, wherein lies the anomaly? In Dr Neville's experiments we are dealing with one standard of control so that his results can be summarized quite simply as follows: on any one site with any one type of concrete the scatter of results increases with increase of 28-day strength. This immediately becomes recognizable as site experience. Furthermore, the standard deviation as a measure of control between different sites is usually applied to the 28-day strength only and this is usually recognized in the practice of basing the design of the mix on the 28-day strength.

Consider, however, the question of designing concrete for one site to give minimum strengths of 1,000, 3,000, and 6,000 lb/in² at 28 days. Table I shows the results that I would normally expect.

TABLE I

Required minimum (lb/in ²)	Practical average for which concrete would be designed (lb/in ²)	Standard deviation of previous column (lb/in ²)	Coefficient of variation (%)
1,000	1,500	215	14.3
3,000	4,000	430	10.7
6,000	7,500	645	8.6

The second column shows the average strength for which the concrete would, in practice, be designed, provided the control was good; from our experience, for site work concrete required to have a minimum strength of 1,000 lb/in² would be designed to have an average strength of 1,500 lb/in² and so on.

According to Himsworth's Table⁽¹⁾, if the standard deviation means anything, then from the third column of Table I it is possible to say that control ranges from "better than excellent" to "medium" (fair to good), but according to Stanton Walker, if the coefficient of variation means anything, control ranges from "good" to "better than excellent" in the reverse direction.

*Pages 75-84 of Magazine No. 32.

Our normal practice is therefore not to assume a constant standard deviation or a standard coefficient of variation. Below 3,000 to 4,000 lb/in² the method based on the coefficient of variation has more to recommend it than that based on the standard deviation; above this range the latter method is of greater value. In practice we use neither.

I am not convinced, however, either by Dr Neville's tests or by his interpretation of the site test data, that there is any justification whatsoever for his last paragraph, especially as this would be against my experience. I would not expect to obtain a coefficient of variation for "run of the mill" 3,000 lb/in² concrete on a site and apply that coefficient to the design of 6,000 lb/in² concrete for the same site.

Notwithstanding the logic displayed in the statistical analysis of his results, I am surprised that Dr Neville has summed the matter up by saying that a standard coefficient of variation can be used for concrete where there is a constant degree of control, especially so when he uses almost the same example as that used by Himsworth⁽⁴⁾ who pointed out that, if a coefficient of variation derived from 3,000 lb/in² concrete is applied to 6,000 lb/in² concrete, the difference between the minimum and the mean strength for the 6,000 lb/in² concrete is twice what it is for the 3,000 lb/in² concrete. This is a result which Himsworth suggested is impracticable and if this means that one would not use it in practice I agree.

Incidentally, it is a pity Dr Neville did not carry out the present series of tests before he investigated and presented his paper on the effect of cube size on mean strength and standard deviation, for in that paper he assumed to be true what he has now "proved" to be untrue.

Contribution by H. C. Erntroy, M.Sc.(Eng.),
A.M.I.C.E.
(Cement and Concrete Association)

Dr Neville has produced some additional laboratory results and has re-analysed reported data to prove his contention (p. 83) that "the standard deviation [of site-made cubes] is proportional to the mean strength, for the *same degree of control* of manufacture of concrete", so that "the present practice of mix design on the basis of a constant coefficient of variation . . . appears to be satisfactory". I do not intend to dispute Dr Neville's conclusion (p. 77) from his laboratory tests that for strengths above 1,500 lb/in² the coefficient of variation is "sensibly constant", although results based on 12 specimens only are naturally subject to some inaccuracies, as indicated by the widely differing values in Table 1 for nominally identical pairs of batches (e.g. with aggregate/cement ratios of 6.0 and 4.0).

While I agree with Dr Neville that it does not seem that either the age at test or the type of cement affects the relation between standard deviation and mean

strength, I must regard the subsequent conclusion as somewhat doubtful since the specimens for the new "concrete cube test" for cement testing are made individually, as are the standard vibrated mortar cubes, and not in the sets of 12 on which Dr Neville apparently bases his comparison.

His interpretation of the Road Research Laboratory data obtained on 58 jobs is claimed (p. 82) to offer "very strong evidence of proportionality between the standard deviation and the mean strength of site-produced concrete for a *constant* degree of control". A critical study of Figure 9, however, would hardly support this assertion. The range of mean strength for "fair" and "poor" control is limited to about 2,000 to 5,000 lb/in² and neither of the regression lines as drawn passes through the origin; this is particularly true of the results for "good" control which cover a somewhat wider range of mean strength. (It may also be noted that the variation for "poor" control is lower than that for "fair" control.)

The conclusions drawn from Figures 10 and 11 would also have been modified had Dr Neville recalled his own remark (p. 81), made when considering the data provided by Himsworth and Murdock, that varying degrees of control in one set of data introduce "an additional vital factor in the relation between the standard deviation and mean strength" and must therefore be treated separately. If the three different symbols in Figure 10 are considered independently it will be noted that there is little justification for drawing only one regression line through them.

Any study involving comparative variabilities of site concrete is complicated by the fact that a large number of "individuals" is required to establish an accurate measure of the standard deviation, and Dr Neville's requirement (p. 83) that "a full confirmation . . . can be obtained only from test data on concretes of widely differing strengths made on the same site, with the same control and under the same supervision" is not likely to be satisfied frequently. However, by accumulating a large amount of data obtained under known degrees of control from a wide range of sources, a good estimate of the real relation can be obtained. To this end, the Cement and Concrete Association has conducted a survey^(5,6) since 1953 of the variation of works test cubes on about 300 construction sites where a study of the degree of control has been made. An analysis of these results indicates that neither the standard deviation nor the coefficient of variation is linearly related to the mean strength over the entire range but that the standard deviation is approximately proportional to the mean strength for mean strengths of up to about 3,000 lb/in² and approximately constant for mean strengths of about 5,000 lb/in² and over. As Dr Neville has pointed out, from a mix designer's point of view the relation between the mean and minimum strengths is required, and the C & C A work has shown that, although these

strengths are not related in a simple way, the water/cement ratios required to give these strengths bear a constant proportion for any one standard of control; this leads to a modification of the Road Note No. 4 method of estimating the required mean to give a specified minimum strength. The mean strength given by this modification to satisfy a minimum strength requirement of 6,000 lb/in², quoted by Dr Neville (p. 76), is of the same order as the 7,160 lb/in² obtained by the "standard deviation method". It is of interest to note, however, that for concrete of low mean strength this modified method gives a result approaching that obtained by the "coefficient of variation method".

It seems all the more surprising that Dr Neville maintains (p. 83) that the method of Road Note No. 4 "appears to be satisfactory" when, as revealed in his own example (p. 76), it calls for a mean strength of 9,800 lb/in² to satisfy a minimum strength of 6,000 lb/in². Since he admits (p. 76) that it is "generally assumed that the distribution of concrete strength follows the normal distribution . . . although only approximate[ly]", this would lead to a maximum strength in the region of 13,500 lb/in². These values of mean and maximum strength are hardly ever attained in practice with normal concreting techniques, although the specified minimum strength of 6,000 lb/in² is often satisfied.

Dr Neville states (p. 83) that the method based on the coefficient of variation is inapplicable to "very low strengths (less than, say, 1,500 to 1,200 lb/in²) Fortunately, these extremely low strengths are not of practical importance". He has, however, not established that it is appropriate for strengths of about 4,000 lb/in² or over. It would thus seem that this method has a severely limited range of application and could well be superseded by the C & C A method.

Contribution by W. J. Larnach, M.Sc.,
A.M.I.C.E.
(University of Bristol)

Dr Neville has performed a service in examining and rationalizing site test data on concrete strength. With this reinterpretation, and the support of his experimental results, it seems that the basis given in Road Note No. 4⁽²⁾ for assessing the design strength, when the required "minimum" is known, is sound. Hitherto, one has been rather confused by the divergent views inherent in Road Note No. 4 and in the work of Himsworth and Murdock (mentioned by Dr Neville) and of Wright⁽⁷⁾.

With regard to Table 2 of the paper, it would be interesting to know the age of the mortars at test, and the number of specimens per batch. In an investigation such as that described in the paper, when a large range of variables is to be investigated, the number of replicate specimens must be kept small. The degree of uncertainty, however, in an estimate of the standard deviation of a population from the results of a small

sample could be significant. Tests on larger samples in a few instances might therefore prove worth while as a check on the trends observed in the paper.

Such tests on large samples seem to have been rarely reported. Williams⁽⁸⁾ presents the results for two series each of 54 mortar cubes, but their mean strengths differed only by some 10%. Some years ago, I carried out similar tests, and the results are summarized in Table II.

TABLE II

Series	Number of cubes	Age at test (days)	Mean strength (lb/in ²)	Range (lb/in ²)	Standard deviation (lb/in ²)
1	54	3	2,000	700	200
2	54	3	3,150	1,060	230
3	48	7	4,660	1,100	250

The specimens were standard mortar cubes of 1:3 composition, with a water/cement ratio of 0.4. Series 2 and 3 were compacted with a B.S. mortar cube vibrator, and Series 1 with a Westool Stewart table vibrator. Each series was manufactured from a single batch of ordinary Portland cement which was not used for any other series.

On the evidence of these three "large sample" series, it appears that the same trend of increasing standard deviation with increasing mean strength occurs. Clearly, however, it would be interesting and informative to have further results from samples of this size, and having a wider range of mean strength.

Contribution by F. A. Sharman, B.Sc.(Eng.),
A.C.G.I., M.I.C.E.
(Sir William Halcrow & Partners)

It was high time that somebody attempted, as Dr Neville has done, to answer experimentally the question of whether there is any inherent change in standard deviation with strength. The results given challenge some of the existing data and form a valuable addition to information on this subject, but I doubt whether they truly support all the conclusions Dr Neville has drawn.

I was surprised to find that, in the experiments described in this paper, standard deviation has been computed from the results of only 12 tests—or 6 to 12 tests for "between-batch" variation. It is generally supposed that at least 30 results are needed to establish the variation parameters with any reliability. With 12 results only, there is a 5% chance of any of the standard deviations—which look so accurately determined in Table I—being in fact more than 41% out, and it is therefore only to be expected that the corresponding coefficients of variation should be rather wildly scattered, as they are indeed in Figure 2. What is not to be expected is that we should be told (p. 77) that "the significance of an approximately linear relation" for

the points plotted in Figure 2 is “considerably better than 0.1%”. This sort of thing brings statistics into disrepute.

But, however the results are interpreted, they are in clear contradiction to those of the ASTM series⁽⁹⁾, which are quoted by Himsworth⁽¹⁰⁾ and form a vital link in the reasoning of his paper. Dr Neville implies that Himsworth relied entirely on site data, and ignores the ASTM results altogether. These American results may of course be wrong or misleading in some undisclosed particular, but at least they were based on sets of 100 tests, and form a convincing consistent series. They clearly point to constant standard deviation. I hope that Dr Neville will clear up this curious conflict of evidence for us.

What matters to the designer and specification-writer, however, is the type of variation to be expected on site. Dr Neville’s argument, that better control is generally exercised on stronger concrete, thus masking in some of the previously published data any increase in standard deviation with increased strength, is perfectly valid. But it hardly leads to the conclusion that this observed fact of life should be ignored, and that mixes ought to be designed on a constant coefficient basis.

There is in fact no single rule for mix design. Sometimes it must be done before a contractor is chosen, and in ignorance of the plant and general supervisory system which will be used. Sometimes the control possibilities of the machinery for producing the concrete are fully known: sometimes they can be dictated within limits. The importance of cement economy varies, as does the tolerance of the design to a few sub-standard batches. As at Woodhead New Tunnel⁽¹¹⁾, the distribution of results may not be symmetrical about the mean, and it becomes necessary to study the pattern of distribution. In all this there is room for far more diversity than there is in the disputed constancy of deviation or coefficient. But it would certainly be nice to know exactly what happens in the laboratory.

Contribution by P. J. F. Wright, B.Sc.
(Road Research Laboratory)

Dr Neville has made a very useful examination of the variations that occur in concrete cube strengths. In general his discussion assumes that the variation is a single quantity which might be expected to bear a very simple relation to the average cube strength. It seems more likely, however, that the variation is compounded from several sources and, whereas some of the component variations may be constant, others may increase with the average strength. It is useful to consider the various sources in turn.

Most of Dr Neville’s laboratory investigation is concerned with what may be termed testing error. This includes such variations as arise from differences in compaction between one specimen and another,

slight departures from planeness in the moulded surfaces, and variations in the rate of loading or other characteristics of the testing technique. It might be expected that the errors arising in this manner would be proportional to the strength, as Dr Neville’s work indicates. This has also been the general experience at the Road Research Laboratory although, in general, the coefficient of variation is found to decrease slightly as the strength increases. This effect may be observed in the figures for variation within batches quoted in Table 1b of my paper dealing with variations in the strength of Portland cement.⁽⁷⁾

For cubes made on the site the largest errors probably arise in batching. The batching errors affect the strength largely through variations in water/cement ratio, whether these are due to variations in moisture content of the aggregate, to uncontrolled variations in the amount of water delivered by the measuring tank or metering system, or to changes in the amount of water being used to counteract some other variation such as a change in grading. The changes in water content will be similar whether the concrete is of high or low average strength and, if the cement is batched without appreciable error, the percentage variations in water/cement ratio will also be independent of the average strength. Table III shows the effect of a given percentage change in water/cement ratio on the crushing strength and indicates that this type of variation would tend to give a relatively constant error in crushing strength, irrespective of the average strength of the concrete.

A third source of variation arises from sampling errors, and in the experience of the Road Research Laboratory these errors can be very considerable. Their nature, however, is similar to that of batching errors in that the water/cement ratio of the sample may differ from that in the batch as a whole and it would therefore be expected that sampling errors would also lead to a constant component in the variation of cube strength.

A fourth source of variation in field strengths, which is not normally allowed to arise in laboratory work, is that of variation in the characteristics of the cement. Table 1c of my paper previously referred to⁽⁷⁾ shows that the standard deviation arising from this cause decreases approximately in proportion to the average

TABLE III: Approximate change in crushing strength to be expected from a 10% change in water/cement ratio, based on data in Road Note No. 4.

Water/cement ratio	Average crushing strength (lb/in ²)	Change in crushing strength caused by 10% change in water/cement ratio	
		(lb/in ²)	(% of average)
0.4	6,900	650	9½
0.6	4,100	650	16
0.8	2,400	500	21

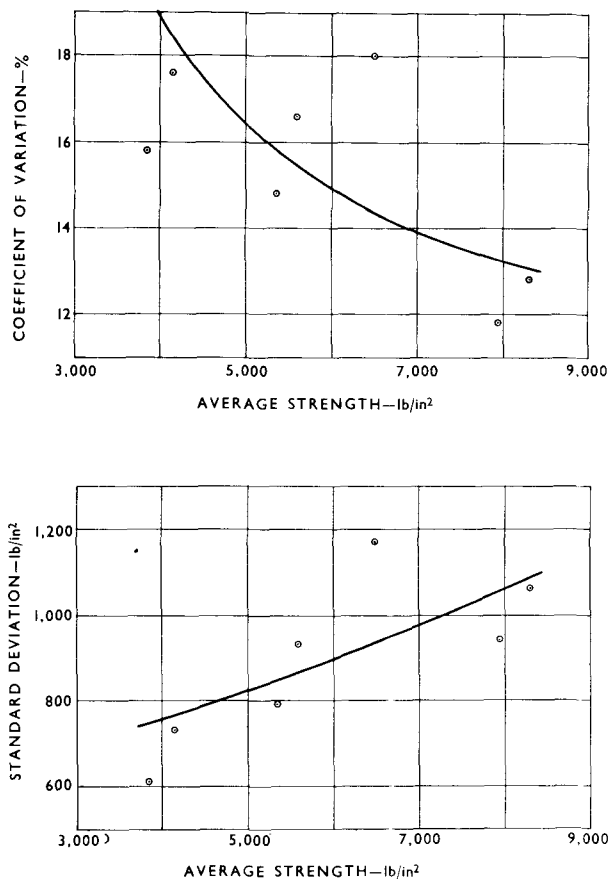


Figure 1: Results of crushing strength tests at Llangyfelach.

strength with different mixes, but is practically unchanged as the strength increases with increasing age.

On the assumption that the over-all variation will consist of several components, some being proportional to the average and some being constant, a linear relation would be expected between the variance of the cube strengths and the square of the average strength. Dr Neville makes a plea in his conclusion for data referring to mixes of different average strengths made in the field with identical degrees of control. Such data are available from a full-scale experiment carried out by the Road Research Laboratory at Llangyfelach in Glamorgan, although the number of results is rather limited. On this site seven different mixes were used for the various experimental sections of the road, but the control and attention given to each of them was exactly the same. About 30 specimens prepared from each of these mixes were tested at the age of 28 days and the coefficients of variation and standard deviations of these are plotted against the average strengths in Figure 1. The specimens were $4 \times 4 \times 20$ in. beams which were tested both for flexural strength and for crushing strength by the equivalent cube technique, but it is unlikely that the use of beams in place of cubes would affect the present discussion. These data strongly support the suggestion that the coefficient of variation decreases with increasing

average strength. The curves shown correspond to the regression line between variance and square of average strength and, although there is considerable scatter among these results, the regression coefficient is significant at the 1% level of probability. The relation represents a constant standard deviation of 620 lb/in² combined with a variable component equal to 10.7% of the average crushing strength.

Reply by the author

I am grateful to Messrs Akroyd, Erntroy, Larnach, Sharman, and Wright, for their thoughtful contributions which clearly show the importance of the topic under consideration. It seems to me, however, that Mr Akroyd has rather missed my argument about the lack of influence of age *per se* on the standard deviation.

Mr Wright is correct in stating that the variation in strength is compounded from several sources but, as he himself points out, I investigated "laboratory errors" only.

I subscribe fully to his comments on the effect of the error in the water/cement ratio. May I add that his data are virtually the same as those given by me in 1954⁽¹²⁾, which of course is not surprising since we both based our calculations on Road Note No. 4. In the present investigation this source of error would have been a very minor factor.

Mr Wright's study of different sources of variation would go a long way towards explaining the difference between my experimental results and those of Mr Erntroy, or between laboratory and site behaviour in general. I feel that although laboratory investigations on the variability of concrete are clearly of only limited interest they may be of help in analysing site data. Mr Wright's Figure 1—a particularly valuable example of site data—suggests that the increasing part of the standard deviation, quoted by him as some 10.7% of the mean strength, is probably due to the same causes which were responsible for the increase in the standard deviation with strength in my tests.

The regression coefficient for this graph of Mr Wright's is significant at the 1% level of probability, and if Mr Sharman compares the scatter of results in this Figure with the scatter in my Table 1 he will see better how my analysis of variance has produced the result of a 0.1% significance level for the linear relation suggested.

May I draw Mr Erntroy's attention to the fact that my regression lines in Figure 9 are not the only ones that do not pass through the origin: neither does Mr Wright's regression line in Figure 1.

There is no doubt that the size of the sample affects the accuracy of the resulting data, but the trends in the variation of the standard deviation can be clearly seen even with samples of 12 cubes, and the significance of the observed variation can be tested by the usual

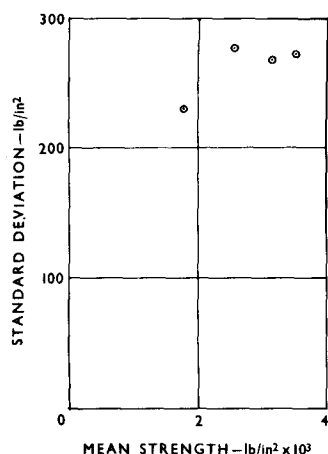


Figure II: The relation between the standard deviation and the mean strength for the data of Crum and Leavitt⁽⁹⁾.

statistical methods for the appropriate number of degrees of freedom.

This brings us to the subject of the four large samples of Crum and Leavitt⁽⁹⁾. Their tests were made some 30 years ago and, although this does not decrease their validity, it is difficult to know the precise conditions of their tests. A plot of the standard deviation against mean strength (Figure II) does not lead to any obvious interpretation. Furthermore, Fisher's variance ratio test of the variances at the strengths of 2,566 and 1,784 lb/in² shows that they differ at the 1% significance level so that we cannot interpret the data of Crum and Leavitt to mean that the standard deviation in their tests was constant. We should also remember that the range of strengths tested was between 1,784 and 3,504 lb/in² only and three out of four tests were within a range of 938 lb/in² whereas I reported tests on concretes with cube strengths between 500 and 13,000 lb/in².

While on the subject of large samples may I say that I was most interested to read of Mr Larnach's tests on samples of 48 to 54 cubes each: the results of his test are in very close agreement with those reported in my paper.

In answer to his question the cubes of Table 2 were cast in batches of 12 and were tested at 28 days.

It is also particularly pleasing to hear that Mr Erntroy does not dispute my conclusions regarding laboratory tests, for his experimental resources are far greater than could ever be hoped for in a university laboratory. Regarding his comments on Figures 9, 10, and 11, I think I ought to repeat that classification into various degrees of control is perforce arbitrary when applied years later to sites in different parts of

the world where different contractors have worked during some 30 years. It is not even possible to have the subjective judgement of a man who had seen all the jobs in progress.

I am not sure that Mr Erntroy's numerical example is one that would occur on site because concrete with a minimum strength of 6,000 lb/in² would not normally be manufactured with the degree of control implied by the figures quoted on p. 76. This point was made in the paper and is, of course, borne out by contracting practice, as mentioned for instance by Mr Sharman.

The practical design recommendations of Mr Erntroy seem both interesting and simple and I am looking forward to hearing details of them.

REFERENCES

1. HIMSWORTH, F. R. The application of statistics to concrete quality. *Proceedings of a Symposium on Mix Design and Quality Control of Concrete*, May 1954. London, Cement and Concrete Association, 1955. p. 470.
2. ROAD RESEARCH LABORATORY. *Design of concrete mixes*. 2nd edition. London, H.M.S.O., 1950. pp. 16. Road Research Road Note No. 4.
3. WALKER, S. *Application of theory of probability to design of concrete for strength specifications*. Paper presented at 14th Annual Meeting of National Ready Mixed Concrete Association at New York, 27th January 1944. N.R.M.C.A., Washington, 1955.
4. HIMSWORTH, F. R. op. cit. pp. 469-470.
5. ERNTROY, H. C. *The variation of works test cubes*. London, Cement and Concrete Association, January 1959. pp. 23+63. Technical Report TRA/287.
6. ERNTROY, H. C. *The effects of variation in cement and of testing error on the variability of concrete*. London, Cement and Concrete Association, June 1959. pp. 12+9. Technical Report TRA/302.
7. WRIGHT, P. J. F. Variations in the strength of Portland cement. *Magazine of Concrete Research*. Vol. 10, No. 30. November 1958. pp. 123-132.
8. WILLIAMS, T. E. H. Vibrated concrete and mortar cubes: distribution of compressive strength. *Magazine of Concrete Research*. Vol. 3, No. 9. March 1952. pp. 107-112.
9. AMERICAN SOCIETY FOR TESTING MATERIALS. *Report on the significance of tests of concrete and concrete aggregates*. 2nd edition. Philadelphia, American Society for Testing Materials, 1943. pp. 176. Special Technical Publication 22-A.
10. HIMSWORTH, F. R. The variability of concrete and its effect on mix design. *Proceedings of the Institution of Civil Engineers*. Part I. Vol. 3, No. 2. March 1954. pp. 163-200.
11. SHARMAN, F. A. Concrete quality control at Woodhead New Tunnel. *Proceedings of the Institution of Civil Engineers*. Part I. Vol. 3, No. 5. September 1954. pp. 564-578.
12. NEVILLE, A. M. Contribution to discussion on paper by F. R. Himsworth: The variability of concrete and its effect on mix design. *Proceedings of the Institution of Civil Engineers*. Part I. Vol. 3. November 1954. pp. 742-743.