7 March 1932.

Dr. Hubert D. Goodale,
Mount Hope Farm,
Williamstown,
Massachusetts,
U.S.A.

Dear Dr. Goodale:

Thanks for your letter and cheque; I am sorry you were troubled twice with the latter.

I see by your letter that we are getting right down to the heart of the problem. In your third paragraph you suggest that if we compared

$$H + \tfrac{1}{10}D \qquad (a)$$

with

$$2H - D \qquad (b)$$

the relative variability of (a) will be less than that of (b). I must say at once that I think this most unlikely, and that in suggesting it you must be thinking of the absolute variabilities. The absolute variability <u>will</u> probably favour (a) against (b) as I have written it; but I might equally have written

$$H - \tfrac{1}{2}D \qquad (b')$$

or

$$\frac{H}{10} - \frac{D}{20} \qquad (b'')$$

and so cut down the absolute variability as low as I liked, without in the least affecting the value of the formula. This only shows that absolute variability is the wrong basis for comparison. The case is different for relative variability, for on changing from (b) to (b′) not only is the variance of different estimates of the same bull divided by 4, but the variance of the mean estimates of different bulls is also divided by 4, and so the equivalence of formulae b, b′, b″, etc. is shown by giving the same relative variability.

If, however, (a) did in fact give the lowest relative variability, I should not hesitate to say it was in fact the best formula. It would mean, I suppose, that within the offspring of the same bull, there was zero correlation between dam and heifer. But in any case, is not the formula which discriminates most clearly between different bulls bound to be the best?

If you could send along the actual data, with sums of squares and products of heifers and dams, and totals merely for different bulls, I think I could convince you.

Yours sincerely,