13 June 1934.

Dear Whatley Carington,

I have looked through the material on the Distribution of Reaction Time, and find it really very reassuring. The extremely long times are not very numerous, though in the case of Leonard and Feda there certainly are a few that would be better for being scaled down. As regards practical procedure, if you think reducing the times to percentages of the medians is worth the labour then I think your whole problem would be solved by taking 10 classes, below 25%, 25%-50%, and so on up to above 250% as class No. 10, and using the numbers 1-10 as your variate. This will simplify the arithmetic and the classing of the reaction times observed will only involve a short table based on the median for each series.

As regards the distribution of ratio to median, those of Garrett and Uvani are not perceptively different, χ^1 = 10.985 for 9 degrees of freedom, while Feda and Leonard also are not significantly different, χ^1 = 14.014. In both cases the chief differences are in the longer times, which appear more frequently in Garrett than in Uvani, but less frequently in Leonard than in Feda. A curious contrast, though, perhaps,

there is nothing more than chance in it.

On your second question of gap-filling, I really should rather look at the data, i.e. apply some exact tests before suggesting any of the simpler possibilities. For ten sittings your words will fall into classes, some being available from all sittings, Class 1, some having a particular reading missing, e.g. third session with Garrett, and so on. Now if you could give me a table showing the number of words of each class and the total scores of each class on the occasions for which they are available, I could form a fair opinion as to how much additional information can be got out of the uncompleted words, as well as on how it could be got out. I state the first question first because it may be that the additional information requires disproportionate additional labour, and should not be done so long as you can spend your time better.

The last point is really very subtle. When experimenters ask me how high odds they ought to get in order to be satisfied I am always inclined to tell them to set the standard of significance rather low, say at 5%, knowing what they are doing, i.e. asking to be decieved, if the supposed effect is really absent just once in 20 trials. When they get a significant result, therefore, they have not certainty, but something worth following up. I should say that a scientific effect was scientifically demonstrable if we know how to conduct an experiment so

that we really fail to get a significant result. The whole utility of this definition being that only a small number of repetitions are needed to distinguish between things that happen once in 20 trials and things which really fail to happen. Consequently, the experimenter whose only evidence up to date is a single significant experiment ,i.e. a single significant discrepancy from some previous expectation has as his next job to design an experiment, which, if his interpretation is correct, will really fail to show up the effect he is after. If his first experience is only just significant his design may have to be somewhat better, i.e. either more extensive or improved in other ways by the elimination of what he has since come to recognise as disturbing factors.

On pure chance I think he ought to expect to encounter a certain number of inexplicable occurrences, e.g. when I first bred mice which were expected to give me 1 in 4 albinos, I remember one mating which produced only 1 in 31; absurdly below expectation. To test the mother I mated her with an albino buck, thus increasing the expectation to a half - and her first test litter consisted of 9 albinos, with no coloured mice. Had it gone the other way, I should certainly have been confirmed in my suspicion that she was abnormal and was producing too few albino eggs. One of the troubles of psychical research seems to be that whereas in a mouse-room it would have been

unreasonable to put this down to any of the traditional sorts of bogey, in a séance designed for the purpose, it would certainly be the proper explanation of an equally improbable event. I say proper, of course, not necessarily true. The unavoidable necessity seems to be reproducability at will, and the fact that one may be dealing with other wills is no real obstacle to this, for I could write 500 letters to different people, asking for replies, and at their will obtain quite a consistent series of responses; or to take the opposite side of the parallel: some friends of mine have been engaged for nearly 15 years now testing the effect of high tension electrical discharges on the growth of barley and have frequently obtained significant effects, none of which are regularly reproducable. Consequently, they have proved nothing, except that there is a mystery or bogey in it.

Yours sincerely,