Danmarks
Pædagogiske
Institut

Hermodsgade 28
2200 København N
Tlf. (01) 810140

Professor Benjamin Wright
The University of Chicago
The Department of Education
5835 Kimbark Ave., Chicago
Illinois 60637
USA

Den 19.10.79
Journal nr.

Dear Ben

Referring to your last letter I must disappoint you. What you indicate about my dawning realization of how a proper analysis of testdata should be carried out is a phantasy of your own which has no connection whatsoever with realities!

I did do something in 1947 that shook the psychologists and educationalists in Scandinavia - so much that one of the leading Danish newspapers (Berlingske Tidende) brought an extensive interview with me (a full page). My main point was obvious, even though I had no better data than the raw scores per person (some 1200 conscripts) in a very mixed sort of intelligence test (an "omnibus test") selected from a large Swedish study of intellectual achievements in various directions of children at different ages.

Benefiting from that the psychologist E. Rubin & E. Tranekjær Rasmussen, then heading the just established (in 1944) course in Educational Psychology, undertook the construction of a new Danish Intelligence Test, which should be ready when the negotiations about establishing a section of psychology within our Defense had been finished. As a teacher in statistics at the said course I acted as a statistical consultant for the group, constructing and trying out the new intelligence test.

But mind you: By then I had no previous experience whatsoever in test psychology and knew about "standardization" only what I could read in existing textbooks. An a narrow time limit.

So I had to start from sample rawscores from a test that on purpose was composed of items covering as large a variety of intellectual performances as at all feasible - a real omnibus test!

Under these circumstances I could do no better than to throw such light on the raw score distribution as the data offered. And they did offer some external data for each person, viz.:

a) Living place while growing up: Capital, towns, rural districts.
b) Father's occupation (ranging from manual workers to bankers and professors).
c) Own school education (4 groups) from 4 forms or more.

What I found was to begin with what is found everywhere:

1) Growing up in the capital leads to higher raw scores than growing up in other towns, which again was better than growing up in rural areas;
2) Better social conditions when growing up gave better results than more poorly conditions;
3) Better educations lead to higher intelligence scores than poor education.

Wellknown and with obvious comments! But when the criteria were intersected then the picture changes completely. In short:

<u>For any given education the distribution of the raw scores is completely independent of both the social and the geographical criteria!</u>

A result I by then was not able to trace in the literature then available to me.

How well known is it today?

So <u>that</u> was what happened to me in 1947.

After that I pursued my teaching in statistics to educationalists and to groups of mathematicians and actuarians - in the spirit of Laplace, Kapteyn and R.A. Fisher till the end of the fifties, when I began to master my own ideas.

But parallel to that I worked as a consultant to the group of military psychologists (since 1952) and the Danish Institute for Educational Research (since 1955). Apart, of course, from free lance work in almost every field of experimental and observational research.

But if you wish to know how I came across the famous Multiplicative Model for Dichotomic Responses (MMDR), I can tell it

quite precisely, and it had in fact nothing to do with item analysis, but was a pure mathematical byproduct to the first discovered multiplicative model, viz. the one linked to the Poisson distribution!

You know about the problem I was faced with on my return from India ab. February 1st in 1952: having to follow up the development in reading ability in a number of individuals, tried out through some years with different reading tests (based on records of both misreadings and reading speed).

A purely selfish hope for mathematical simplicity conducted my choice of statistical tool to the Poisson distribution, whose parameter then should characterize both the reader and the text. Mathematically it was obvious that a very nice theory would come out of it, if that parameter happened to be a product of one parameter for the person and one for the test.

My goodness - it did work in practice!!

But having tried it out over some months I became again mathematically infatuated:

How the deuce could such a peculiar sort of Poisson distribution come into existence?

Poisson (ab. 1838) himself stumbled upon his version of it as a limiting case of the binomial distribution with large n and a parameter $\to 0$, inversely proportional to n. Much later, I think, the proof was extended to cover the sum of many small independent random variables, provided the mean value of the sum has a finite limit as the number of terms $\to \infty$, who first did that I really don't know.

However, if a machinery like that should lead to the MPM for the number of reading mistakes, then we must go back to the single words of a text, each of which is rather easy, but all of them - or at least a majority of them - totalling to something not at all negligible and even characteristic of the text as a whole.

Well, this can be specified in an infinity of ways, but it would seem an obvious presumption that each word has its own probability of being misread by a given person. And so I was driven right down to specifying probabilities for dichotomies with parameters referring to two distinct sources: the person and the word - a point of view that had never occurred to me before. And furthermore, when the words were lumped together to make a "text", then - according to my empirical studies - the

probability of so and so many words being misread ought to be a parameter of two factors, one pertaining to the person, the other to as much of the text as was read.

This, I shall think, could hardly be achieved unless the parameters $\lambda_{\nu i}$ (ν= person, i = word) of the person's probability of misreading word No. i of the text, written on the always permissible form

$$\frac{\lambda_{\nu i}}{1+\lambda_{\nu i}}, \quad \lambda_{\nu i} \geq 0$$

was assumed to be the product of a person factor and a word factor, i.e.

$$p\{+|\nu,i\} = \frac{\xi_\nu \varepsilon_i}{1+\xi_\nu \varepsilon_i} \; ; p\{-|\nu,i\} = \frac{1}{1+\xi_\nu \varepsilon_i}$$

There you are: my discovery was a somewhat intuitive achievement, but wholly within my own mathematical playground - with no relation to any actual item analysis problem!

Of course, afterwards I immediately realized that I might have stumbled upon a tool for formalizing, thus handling such problems, and as soon as my substantial report on the reading retarded children had been delivered - 7 months after my return - I made a temporary analysis of accessible data on Raven's matrix test, which I found on the whole beautifully represented by the new model. And then I returned to the 1947-data and realized that they did not at all agree with the model! But on sorting the - by intention - extremely inhomogenous test according to subject matter things looked much better within each "subtest" - though not as good as for the Raven test.

This work - carried out as a consultant to the newly established group of military psychologists - I could present to the head, Poul Borking, of that group, who immediately saw the significance of my discovery, called in a fellow psychologist, Børge Prien - whom I already then knew very well - and assigned the following task to him: constructing four subtests covering quite different subject matters, i.e. requiring different fields of intellectual

activities, each subtest fulfilling the demands of the new model as well as at all possible, having it ready for use - i.e. a table for transforming raw scores to model-measurements for each subtest. All of it ready for being printed exactly 6 months later! Because it had to be used for intelligence testing of the several thousands of conscripts - with a view to selection and distribution of the selected to different units within the defense - starting November 1st. Børge got a staff of young psychologists and students, with me as statistical adviser through the wealth of analyses required. - I dare say we worked fast and fine in 1953! And this was before the era of electronic computers - paper and pencil all of it - and of course a small electric computer!

Well, for computing and evaluation I could by then not offer anything better than a primitive paper and pencil procedure - the outcomes of which you can read about in Prob. Mod., the material of which largely stems from 1952 and 1953. - Not until 1958/59 when I had to write the book, was a mathematical theory for the item analysis worked out.

Well, friend, this is the real story about how the Multiplicative Model for Dichotomic Responses came into existence - so please, drop your nonsense about 1947. And consider how erratically a human mind may work!

The information offered here you may use at your own discretion (but no misrepresentation!) at any opportunity you think fit.

And now to the next letter to you, hopefully reaching you in another couple of days.

Yours as ever,

Georg

Dear Ben

We got this letter handwritten from Georg who is not able to leave his island of Læsø, and thought it appropriate to take the time needed to type it - thus delaying it ab. 1½ month. We hope the increase in readibility compensates.

Yours