A MATHEMATICAL THEORY OF OBJECTIVITY AND ITS CONSEQUENCES FOR MODEL CONSTRUCTION

G. Rasch

Institute of Statistics, University of Copenhagen (Denmark)

European Meeting on Statistics, Econometrics and Management Science,
Amsterdam 2-7 september 1968

A MATHEMATICAL THEORY OF OBJECTIVITY AND ITS CONSEQUENCES FOR MODEL CONSTRUCTION

by G. Rasch (Denmark)

Institute of Statistics, University of Copenhagen

1. A SUFFICIENT CONDITION FOR SEPARABILITY OF TWO SETS OF PARAMETERS.

As an introduction to the ideas developed under the heading "specific objectivity" we shall consider a simple model suggested in (1960) [3] for analyzing dichotomous items of a psychological test.

A number of persons (no. $\nu = 1,2,\ldots,n$) give one out of two possible answers, denoted 1 and 0, to each of a set of questions (no. $i=1,2,\ldots,k$). With respect to such questions each person (ν) is assumed to be fully characterized by a scalar positive parameter, ξ_ν, and similarly each question (i) is assumed to be fully characterized by a scalar positive parameter, ε_i. Furthermore all the answers

$$(1.1) \qquad a_{\nu i} = \begin{cases} 1 \\ 0 \end{cases}$$

are taken to be independent stochastic variables with the probabilities

$$(1.2) \qquad p\{a_{\nu i} = 1\} = \frac{\xi_\nu \varepsilon_i}{1+\xi_\nu \varepsilon_i}, \quad p\{a_{\nu i} = 0\} = \frac{1}{1+\xi_\nu \varepsilon_i}$$

Consider first the case of one arbitrary person with parameter ξ answering two arbitrary items with parameters ε_1 and ε_2.

The possible outcomes

(1.3)

		i = 2: 1	i = 2: 0
i=1	1	(1,1)	(1,0)
	0	(0,1)	(0,0)

form a universe D^2 with the probabilities

$$
(1.4) \qquad i = 1 \quad
\begin{array}{c|cc}
 & \multicolumn{2}{c}{i = 2} \\
 & 1 & 0 \\
\hline
1 & \dfrac{\xi^2 \varepsilon_1 \varepsilon_2}{(1+\xi\varepsilon_1)(1+\xi\varepsilon_2)} , & \dfrac{\xi\varepsilon_1}{(1+\xi\varepsilon_1)(1+\xi\varepsilon_2)} \\[2ex]
0 & \dfrac{\xi\varepsilon_2}{(1+\xi\varepsilon_1)(1+\xi\varepsilon_2)} , & \dfrac{1}{(1+\xi\varepsilon_1)(1+\xi\varepsilon_2)}
\end{array}
$$

and accordingly the probability of the combination (1,0) in the set

$$(1.5) \qquad V = (1,0) \cup (0,1)$$

is

$$(1.6) \qquad p\{(1,0)|V\} = \frac{\varepsilon_1}{\varepsilon_1+\varepsilon_2} ,$$

as the terms containing ξ obviously cancel.

An immediate consequence is that from the answers to the two items of any collection of n persons we may obtain an estimate of the ratio $\varepsilon_1/\varepsilon_2$ that is uninfluenced by the unknown values of the parameters $\xi_1,\ldots,\xi_n$. In fact, if c out of the n persons have just one 1-answer to the two items then the probability of $a_{\nu 1} = 1$ for b of them is given by the binomial distribution

$$(1.7) \qquad p\{b|c\} = \binom{c}{b} \left(\frac{\varepsilon_1}{\varepsilon_1+\varepsilon_2}\right)^{b} \left(\frac{\varepsilon_2}{\varepsilon_1+\varepsilon_2}\right)^{c-b}$$

according to which

$$(1.8) \qquad \frac{\varepsilon_1}{\varepsilon_1+\varepsilon_2} \approx \frac{b}{c} ,$$

irrespective of the values of $\xi_1,\ldots,\xi_n$.

Some colleagues have argued that c itself is a stochastic variable, the distribution of which is a function of $\xi_1,\ldots,\xi_n$ and that b/c therefore indirectly does depend on the unknown parameters. This objection, I think, is met by the observation that whichever collections of persons are used, in so far as the model (1.2) holds we shall, according to (1.7), always get estimates (1.8) which statistically are in accordance with each other. Significant deviations serve as signing where and how the model breaks down.

To the results (1.6) and (1.7) which were presented in [3], chapter X,3 two points may be added: 1) Interchanging persons and items leads to the possibility of estimating the ratio of any two person parameters irrespective of the item parameters, but when items are few, as is usual in practice, the method becomes unpractical. 2) When estimates like (1.8) are available for different pairs of items the question of how to amalgamate them arises. This question does not seem to have an easy answer and therefore the simultaneous treatment of more items than two requires more refined techniques which have been outlined in [3], [5] and [6], being based upon a result which generalizes (1.6), viz. that for any given person (ν) the conditional probability

$$(1.9)\qquad \begin{cases} p\{a_{\nu 1},\ldots,a_{\nu k} \mid \sum\limits_{i=1}^{k} a_{\nu i} = r\} = \dfrac{\varepsilon_1^{a_{\nu i}} \ldots \varepsilon_k^{a_{\nu k}}}{\gamma_r(\varepsilon_1,\ldots,\varepsilon_k)} \quad , \\[2ex] \gamma_r(\varepsilon_1,\ldots,\varepsilon_k) = \sum\limits_{1 \leq i_1 < \ldots < i_r \leq k} \varepsilon_{i_1}\ldots\varepsilon_{i_r} \end{cases}$$

is independent of the person parameter ξ_ν.

Similarly

$$(1.10)\qquad p\{a_{1i},\ldots,a_{ni} \mid \sum_{\nu=1}^{n} a_{\nu i} = s\} = \frac{\xi_1^{a_{1i}}\ldots\xi_n^{a_{ni}}}{\gamma_s(\xi_1,\ldots,\xi_n)}$$

is independent of the item parameter ε_i.

The main result of this section we may formulate as

Theorem 1.

On the assumption that the answers 1 and 0 of different persons to a set of items are independent dichotomous stochastic variables following the distributions (1.2) it is possible to separate the person parameters ξ_ν *and the item parameters* ε_i *such that the person parameters are eliminated while the item parameters are being estimated, and vice versa.*

2. A NECESSARY CONDITION FOR SEPARABILITY.

In view of theorem 1 the question arises whether other models also enjoy the separability or the model (1.2) is unique in this respect.

The answer is given in

Theorem 2.

On the assumption that the answers of different persons to a set of items are independent, dichotomous stochastic variables and that the probabilities of the two possible answers 1 and 0 of a person (ν) *to an item* (i) *depend only on two scalar positive parameters, characterizing the person* (ξ_ν) *and the item* (ε_i) *respectively:*

$$(2.1) \qquad p\{1|\nu,i\} = f(\xi_\nu,\varepsilon_i), p\{0|\nu,i\} = 1-f(\xi_\nu,\varepsilon_i) \ ,$$

then

$$(2.2) \qquad f(\xi,\varepsilon) = \frac{\xi\varepsilon}{1+\xi\varepsilon} \ ,$$

but for trivial transformations, is a necessary condition for the parameters always to be separable as mentioned in theorem 1.

Consider any person with a parameter ξ being exposed to two arbitrary items with parameters ε_1 and ε_2. The set of possible outcomes is D^2 as defined by (1.3), the probabilities of which are

$$\text{(2.3)} \quad \begin{aligned} p\{(1,1)\,|D^2,\xi,\varepsilon_1,\varepsilon_2\} &= f(\xi,\varepsilon_1)f(\xi,\varepsilon_2) \\ p\{(1,0)\,|D^2,\xi,\varepsilon_1,\varepsilon_2\} &= f(\xi,\varepsilon_1)(1-f(\xi,\varepsilon_2)) \\ p\{(0,1)\,|D^2,\xi,\varepsilon_1,\varepsilon_2\} &= (1-f(\xi,\varepsilon_1))f(\xi,\varepsilon_2) \\ p\{(0,0)\,|D^2,\xi,\varepsilon_1,\varepsilon_2\} &= (1-f(\xi,\varepsilon_1))\ (1-f(\xi,\varepsilon_2)) \end{aligned}$$

A being a subset of D^2 our first question is whether $P\{A|D^2\}$ for any choice of the function f could be independent of ξ. Clearly A would have to comprise at least two elements of D^2 and the same must hold for its complement in D^2. Thus, if at all possible, A should be a pair. However, the four pairs

$$\text{(2.4)} \quad \begin{aligned} &(0,0),(0,1) \\ &(0,0),(1,0) \\ &(0,1),(1,1) \\ &(1,0),(1,1) \end{aligned}$$

are excluded unless $f(\xi,\varepsilon)$ is independent of ξ which is a trivial case. Thus we are left with two possibilities:

$$\text{(2.5)} \qquad P\{(0,0)\cup(1,1)\,|D^2\}=(1-f(\xi,\varepsilon_1))(1-f(\xi,\varepsilon_2))+f(\xi,\varepsilon_1)f(\xi,\varepsilon_2)$$

and

$$\text{(2.6)} \qquad P\{(0,1)\cup(1,0)\,|D^2\}=(1-f(\xi,\varepsilon_1))\ f\,(\xi,\varepsilon_2)+f(\xi,\varepsilon_1)(1-f(\xi,\varepsilon_2)\ ,$$

both of which for

$$\varepsilon_1 = \varepsilon_2 = \varepsilon \tag{2.7}$$

depend on ξ except in the trivial case, the conclusion being that no such non-trivial function $f(\xi, \varepsilon)$ exists that any $P\{A|D^2\}$ could be independent of ξ.

Turning now to conditional probabilities it is obvious that V must consist of either a pair or a triple.

As a representative of triples we may take

$$V: \{(0,0),(0,1),(1,0)\} \tag{2.8}$$

in which the probability of at least one of the elements should have the desired property. Consider for instance

$$p\{(0,0)|V\} = \frac{(1-f(\xi,\varepsilon_1))(1-f(\xi,\varepsilon_2))}{1-f(\xi,\varepsilon_1)\ f(\xi,\varepsilon_2)} \tag{2.9}$$

which in the case (2.7) reduces to

$$\frac{(1-f(\xi,\varepsilon))^2}{1-f^2(\xi,\varepsilon)} = \frac{1-f(\xi,\varepsilon)}{1+f(\xi,\varepsilon)} \tag{2.10}$$

And similarly all other cases with V as a triple is reduced in absurdity.

As regards pairs (2.4) of course do not count as possible V's and so we are left with two cases:

$$V:\{(0,0),(1,1)\} \tag{2.11}$$

for which

$$p\{(0,0)|V\} = \frac{(1-f(\xi,\varepsilon_1))(1-f(\xi,\varepsilon_2))}{(1-f(\xi,\varepsilon_1))(1-f(\xi,\varepsilon_2))+f(\xi,\varepsilon_1)f(\xi,\varepsilon_2)} \tag{2.12}$$

reduces to

$$(2.13) \qquad \frac{1}{1+\left(\frac{f(\xi,\varepsilon)}{1-f(\xi,\varepsilon)}\right)^2}$$

when the ε's are equal, and

$$(2.14) \qquad V:\{(0,1),(1,0)\}$$

with

$$(2.15) \qquad p\{(0,1)|V\} = \frac{(1-f(\xi,\varepsilon_1))f(\xi,\varepsilon_2)}{(1-f(\xi,\varepsilon_1))f(\xi,\varepsilon_2)+f(\xi,\varepsilon_1)(1-f(\xi,\varepsilon_2))}$$

which for coinciding ε's reduces to $\frac{1}{2}$. Accordingly our problem has been boiled down to the question: For which function $f(\xi,\varepsilon)$ is (2.15) independent of ξ for any unequal ε_1 and ε_2? Clearly, the ratio

$$(2.16) \qquad \frac{f(\xi,\varepsilon_1)}{1-f(\xi,\varepsilon_1)} : \frac{f(\xi,\varepsilon_2)}{1-f(\xi,\varepsilon_2)}$$

should be independent of ξ, e.g. equal to its value for $\xi = 1$; accordingly

$$(2.17) \qquad \frac{f(\xi,\varepsilon_1)}{1-f(\xi,\varepsilon_1)} = \left(\frac{f(1,\varepsilon_1)}{1-f(1,\varepsilon_1)} : \frac{f(1,\varepsilon_2)}{1-f(1,\varepsilon_2)}\right) \cdot \frac{f(\xi,\varepsilon_2)}{1-f(\xi,\varepsilon_2)}$$

must hold for arbitrary ε_1, ε_2 and ξ. In particular $\varepsilon_2 = 1$ yields the relation

$$(2.18) \qquad \frac{f(\xi,\varepsilon_1)}{1-f(\xi,\varepsilon_1)} = \frac{f(1,\varepsilon_1)}{1-f(1,\varepsilon_1)} \cdot \frac{f(\xi,1)}{1-f(\xi,1)} : \frac{f(1,1)}{1-f(1,1)}$$

Take now

$$(2.19) \qquad \xi' = \frac{f(\xi,1)}{1-f(\xi,1)}$$

and

$$(2.20) \qquad \varepsilon_1' = \frac{f(1,\varepsilon_1)}{1-f(1,\varepsilon_1)} : \frac{f(1,1)}{1-f(1,1)}$$

as new parameters, then we get

$$(2.21) \qquad f(\xi,\varepsilon_1) = \frac{\xi'\varepsilon_1'}{1+\xi'\varepsilon_1'}$$

from which it follows that with a suitable choice of parameters (2.2) is the only solution of our problem.

3. A GENERALIZED SEPARABILITY THEOREM.

In order to suggest a possible generalization of theorem 1 to more response categories than two we shall first consider a homogeneous version of model (1.2)

$$(3.1) \qquad \begin{aligned} p\{x^{(1)}|\nu,i\} &= \frac{\xi_\nu^{(1)}\varepsilon_i^{(1)}}{\gamma_{\nu i}} \\ p\{x^{(2)}|\nu,i\} &= \frac{\xi_\nu^{(2)}\varepsilon_i^{(2)}}{\gamma_{\nu i}} \end{aligned}$$

where

$$(3.2) \qquad \gamma_{\nu i} = \xi_\nu^{(1)}\ \varepsilon_i^{(1)} + \xi_\nu^{(2)}\ \varepsilon_i^{(2)} \quad .$$

Obviously (3.1) is connected with (1.2) by the relation

$$\xi_\nu = \frac{\xi_\nu^{(1)}}{\xi_\nu^{(2)}} \quad ; \quad \varepsilon_i = \frac{\varepsilon_i^{(1)}}{\varepsilon_i^{(2)}} \quad , \tag{3.3}$$

but the symmetry between the two probabilities in (3.1) suggests a generalization to more response categories for which an analogue to theorem 1 holds:

Theorem 3.
Consider a situation where each out of n persons answer k questions to each of which m response categories

$$U : \{x^{(1)},\ldots,x^{(e)},\ldots,x^{(m)}\} \tag{3.4}$$

are available, all nk answers being stochastically independent. Assume that each person (ν) as well as each question (i) is characterized by an homogeneous vector, respectively

$$\xi_\nu = (\xi_\nu^{(1)},\ldots,\xi_\nu^{(e)},\ldots,\xi_\nu^{(m)}) \tag{3.5}$$

and

$$\varepsilon_i = (\varepsilon_i^{(1)},\ldots,\varepsilon_i^{(e)},\ldots,\varepsilon_i^{(m)}) \tag{3.6}$$

such that the probabilities of the possible answers are given by

$$p\{x^{(e)}|\nu,i\} = \frac{\xi_\nu^{(e)}\varepsilon_i^{(e)}}{\gamma_{\nu i}} \quad , \tag{3.7}$$

$$\gamma_{\nu i} = \sum_{e=1}^{m} \xi_\nu^{(e)}\varepsilon_i^{(e)} \quad .$$

Counting for each person (ν) *the number,* $a_{\nu o}^{(e)}$, *of questions to which his answer was* $x^{(e)}$ *and similarly for each question* (i) *the number,* $a_{oi}^{(e)}$, *of persons that gave this answer, then the conditional probability of the whole set of* $a_{oi}^{(e)}$*'s* ($i=1,\ldots,k$, $e=1,\ldots,m$), *given the whole set of* $a_{\nu o}^{(e)}$*'s* ($\nu=1,\ldots,n$, $e=1,\ldots,m$) *depends on the question parameters* (3.5). *And, symmetrically, the probability of all* $a_{\nu o}^{(e)}$*'s, given the set of* $a_{oi}^{(e)}$*'s depends on the* ξ_ν*'s, but not on the* ε_i*'s, cf.* [4].

As regards the number of persons and questions the formulation of theorem 3 is more inclusive than that of theorem 1, covering also the generalization indicated by (1.10).

For the proof a special notation in vector-matrix algebra is expeditious: Let

$$x = (x_1, \ldots, x_m) \tag{3.8}$$

be a real vector and the elements of

$$a = (a_1, \ldots, a_m) \tag{3.9}$$

be integers ≥ 0. We define x raised to the power a as the scalar

$$x^a = x_1^{a_1} \ldots x_m^{a_m} \tag{3.10}$$

and notice the rule

$$x^{a+b} = x^a \cdot x^b . \tag{3.11}$$

Denoting now the answer $x^{(e)}$ of person ν to question i by the selection vector

$$a_{\nu i} = (0, \ldots, \overset{e}{1}, \ldots, 0) \tag{3.12}$$

we may condense the m equations (3.7) into one formula

$$(3.13)\qquad \begin{cases} p\{a_{\nu i}\} = \dfrac{\xi_\nu^{a_{\nu i}} \cdot \varepsilon_i^{a_{\nu i}}}{\gamma_{\nu i}} \,, \\ \gamma_{\nu i} = \xi_\nu \varepsilon_i^{*} \,. \end{cases}$$

According to the stochastical independence we have for a given person

$$(3.14)\qquad p\{a_{\nu 1},\ldots,a_{\nu k}\} = \frac{\xi_\nu^{a_{\nu o}} \cdot \prod_{(i)} \varepsilon_i^{a_{\nu i}}}{\prod_{(i)} \gamma_{\nu i}} \,,$$

where

$$(3.15)\qquad a_{\nu o} = \sum_i a_{\nu i} = (a_{\nu o}^{(1)},\ldots,a_{\nu o}^{(m)}) \,.$$

By summing (3.14) over all sets of selection vectors with the same vector sum $a_{\nu o}$ we get the marginal probability of $a_{\nu o}$:

$$(3.16)\qquad p\{a_{\nu o}\} = \frac{\xi_\nu^{a_{\nu o}} \cdot \gamma((\varepsilon_i)|a_{\nu o})}{\prod_{(i)} \gamma_{\nu i}}$$

where

$$(3.17)\qquad \gamma((\varepsilon_i)|a_{\nu o}) = \sum_{\Sigma_{(i)} a_{\nu i} = a_{\nu o}} \varepsilon_1^{a_{\nu 1}} \cdots \varepsilon_m^{a_{\nu m}}$$

is a polynomial in the ε's.

From (3.14) and (3.16) we obtain the conditional probability of $a_{\nu 1},\ldots,a_{\nu k}$ for given marginal vector $a_{\nu o}$:

(3.18) $$p\{a_{\nu 1},\dots,a_{\nu k}|a_{\nu o}\} = \frac{\prod_{(i)} \varepsilon_i^{a_{\nu i}}}{\gamma((\varepsilon_i)|a_{\nu o})}$$

which is independent of the person parameter ξ_ν.

Utilizing this result for all of the persons we find, due to the stochastical independence, the joint probability of the whole set of selection vectors, given all the marginal vectors of the persons

(3.19) $$p\{((a_{\nu i}))|(a_{\nu o})\} = \frac{\prod_{(i)} \varepsilon_i^{a_{oi}}}{\gamma((\varepsilon_i)|(a_{\nu o}))}$$

where

(3.20) $$a_{oi} = \sum_{(\nu)} a_{\nu i} = (a_{oi}^{(1)},\dots,a_{oi}^{(m)})$$

and

(3.21) $$\gamma((\varepsilon_i)|(a_{\nu o})) = \prod_{(\nu)}\gamma((\varepsilon_i)|a_{\nu o}) \quad ,$$

and from (3.19) we get

(3.22) $$p\{(a_{oi})|(a_{\nu o})\} = \begin{bmatrix}(a_{\nu o})\\(a_{oi})\end{bmatrix} \cdot \frac{\prod_{(i)} \varepsilon_i^{a_{oi}}}{\gamma((\varepsilon_i)|(a_{\nu o}))}$$

where the bracket indicates a combinatorial coefficient, viz. the number of matrices

(3.23) $$((a_\nu)) \text{ , } \nu = 1,\dots,n,\ i = 1,\dots k \text{ ,}$$

with selection vectors as elements, which have the two sets of marginal vectors $(a_{\nu o})$ and (a_{oi}).

Symmetrically, the formula

(3.24) $$p\{(a_{\nu o})|(a_{oi})\} = \begin{bmatrix}(a_{\nu o})\\(a_{oi})\end{bmatrix} \cdot \frac{\prod_{(\nu)} \xi_\nu^{a_{\nu o}}}{\gamma((\xi_\nu)|(a_{oi}))}$$

also holds.

Thus the proof of theorem 3, *the separability theorem for the model* (3.7) has been completed.

4. PARTIAL AND COMPLETE SEPARABILITY.

An inversion of theorem 3 requires that a further property of the model is realized.

In practice it is quite often felt that the categorization of the observable qualities is somewhat arbitrary, or that there are too many categories to be manageable, or that some of the categories are rarely observed in a given set of data. In such cases it is costumary to undertake a revision of the categorization, quite often by pooling two or more categories into one. This,of course, is quite legitimate, but in connection with the model (3.7) it should be noticed that the probability of a subset V of U being

$$(4.1) \qquad P\{V|\nu,i\} = \frac{\sum_{x^{(e)}\in V} \xi_\nu^{(e)}\varepsilon_i^{(e)}}{\gamma_{\nu i}} \quad ,$$

its numerator does not usually enjoy the multiplicativity of the numerator in (3.7), on which the separability was founded. Thus the separability usually gets lost by pooling categories, if it were attainable for the original categorization.

However, the separability is retained by a different kind of modifying the categorization, viz. that of *concentrating upon the responses falling within a subset V of U* , neglecting observations falling into its complementary set $\bar{V}$ in U.

First it follows from (3.17) and (4.1) that

$$(4.2) \qquad p\{x^{(e)}|V,\nu,i\} = \frac{\xi_\nu^{(e)}\ \varepsilon_i^{(e)}}{\sum_{x^{(h)}\in V} \xi_\nu^{(h)}\varepsilon_i^{(h)}}$$

i.e. the model within V is of the same type as (3.7). Thus, in cases where all nk answers fall within V theorem 3 may be applied to demonstrate the separability of the vectors

$$\left.\begin{array}{ll} (4.3) \qquad \xi'_\nu = (\xi_\nu^{(e)}) & \\ \text{and} & \\ (4.4) \qquad \varepsilon'_i = (\varepsilon_i^{(e)}) & \end{array}\right\} x^{(e)} \in V.$$

However, also for arbitrary response matrices the separability holds. In fact, in the proof of theorem 3 we may just interpret $a_{\nu i}$ as a selection vector in V if $x^{(e)} \in V$ and otherwise let $a_{\nu i} = (o,\dots,o)$ symbolize a "missing plot". The whole argument from (3.13) through (3.24) may then be repeated with obvious modifications, leading to the separability of (4.3) and (4.4).

This result we shall express in a more general setting, leaving aside the references to (3.7) and (4.2).

Let persons and questions be characterized by parameters ξ_ν and ε_i, in so far as the observations are elements of U. For simplicity we shall assume that the ξ's and the ε's are vectors of the same dimension $r \leq m-1$. The elements of the response matrices $((a_{\nu i}))$ are assumed to be independent stochastic variables, which for each (ν,i) has a probability distribution

$$(4.5) \qquad p\{x^{(e)} | \nu,i\} = f_e(\xi_\nu, \varepsilon_i)$$

that depends solely upon the parameters ξ_ν and ε_i. We shall say that *the person parameters and the question parameters can be separated from each other if two non-trivial probability statements, direct or conditional, about the matrices* $((a_{\nu i}))$ *hold, one of which depends on the* ξ_ν*'s, but not on the* ε_i*'s, while the other one depends on the* ε_i*'s, but not on the* ξ_ν*'s.*

Consider now what may happen when an $a_{\nu i}$ is only recorded if the corresponding $x^{(e)} \in V \subset U$ and then as a selection vector referring to V. Within this reference frame the probabilities

$$(4.6) \qquad p\{x^{(e)} | V, \nu, i\} = \frac{f_e(\xi_\nu, \varepsilon_i)}{P\{V | U, \xi_\nu, \varepsilon_i\}}$$

may be expressible in terms of two different sets of parameters, ξ'_ν and ε'_i , of a dimension r' not exceeding r, and $\leq m' - 1$, V consisting of m' elements. ξ'_ν, being characteristic of person no. ν in a more limited sense than ξ_ν, is presumed to be a function of ξ_ν. And analogously for ε'_i. Whether or not the separability holds for ξ_ν and ε_i in relation to U, it may hold for ξ'_ν and ε'_i in relation to V. If it does we shall say that *the model (4.5) allows for a partial separability of parameters for persons and questions. If the partial separability holds for any $V \subset U$ we shall say that the model (4.5) allows for a complete separability of the two sets of parameters.*

The above result may now be expressed as:

Theorem 4.

The model (3.7) allows for complete separability of the two kinds of parameters.

5. A NECESSARY AND SUFFICIENT CONDITION FOR COMPLETE SEPARABILITY OF PARAMETERS FOR PERSONS AND QUESTIONS IN CASE OF MAXIMAL DIMENSION.

Turning now to the inversion of theorem 4 we first notice that if in (4.5) we assume that the person parameter ξ_ν could be fully determined from the m probabilities $p\{x^{(e)} | \nu, i\}$ provided the item parameters ε_i were known, then the number of elements in the vector ξ_ν could not exceed m-1 since

$$(5.1) \qquad \sum_{e=1}^{m} p\{x^{(e)} | \nu, i\} = 1.$$

And the symmetric condition requires that neither the dimension of ε_i

exceeds m-1. In this sense m-1 is considered to be the maximal dimension of both ξ_ν and ε_i.

Apparently the parameters in theorem 3 are m- dimensional, but from (3.7) it is clear that ξ_ν and ε_i are expressed in homogeneous coordinates, a representation we shall make use of whenever convenient.

We may now formulate the inversion of theorem 4.

Theorem 5.

On the assumption that the answers of different persons to a set of items are independent stochastic variables for which the probability distributions over m possible categories $\{x^{(1)},\dots,x^{(e)},\dots,x^{(m)}\}$ depend on two sets of homogeneous m-dimensional parameters ξ_ν and ε_i, referring to respectively persons and items, cf. (4.5), then the validity of the model (3.7) is a necessary and sufficient condition for complete separability of the two kinds of parameters.

From the complete separability it follows that partial separability holds if for V in (4.6) we take any of the pairs

$$(5.2) \qquad V_{em}: \{x^{(e)},x^{(m)}\}\ , \quad e = 1,\dots,m-1\ ,$$

to each of which theorem 2 may be applied. The parameters in (2.2) are then to be understood as person and item parameter corresponding to the pair considered, i.e. as scalar-functions of ξ and ε, say

$$(5.3) \qquad \xi^{(e)} = \varphi_e(\xi), \varepsilon^{(e)} = \psi_e(\varepsilon).$$

Now

$$(5.4) \qquad p\{x^{(e)}|V_{em}\} = \frac{p\{x^{(e)}|U\}}{p\{x^{(e)}|U\}+p\{x^{(m)}|U\}}$$

and therefore

$$(5.5) \qquad \frac{p\{x^{(e)}|U\}}{p\{x^{(m)}|U\}} = \xi^{(e)}\varepsilon^{(e)} .$$

Since the vectors ξ and

$$(5.6) \qquad (\xi^{(1)},\ldots,\xi^{(m-1)},1)$$

both have the dimension m and the m-1 elements of ξ are presumed to be functionally independent - otherwise ξ could have been replaced by a parameter of lower dimension - ξ may, due to (5.3), be expressed in terms of (5.6), i.e. this vector might have been taken as the person parameter. Applying the same argument to

$$(5.7) \qquad (\varepsilon^{(1)},\ldots,\varepsilon^{(m-1)},1)$$

we see that both sets of parameters might have been chosen in such a way that (3.7) holds.

This proves the necessity while theorem 3 ascertains the sufficiency.

6. CONDITIONS FOR COMPLETE SEPARABILITY OF EQUIDIMENSIONAL PARAMETERS FOR PERSONS AND ITEMS.

What happens when the dimensions of ξ and ε are not maximal? So far, I do not possess a complete answer to that question, but I do have an answer of a generality that corresponds to the epistemological viewpoint to be developed in the last section.

For a preliminary motivation I may just point out a further property of the model under consideration which becomes conspicuous by conversion of (3.7) into an exponential form. Setting

$$(6.1) \qquad \log \xi_\nu^{(e)} = \theta_\nu^{(e)} , \quad \log \varepsilon_i^{(e)} = \sigma_i^{(e)}$$

we get

(6.2) $$p\{x^{(e)}|\nu,i\} = \frac{e^{\theta_\nu^{(e)}+\sigma_i^{(e)}}}{\gamma_{\nu i}} \quad , \quad e = 1,\dots,m$$

where now

(6.3) $$\gamma_{\nu i} = \sum_{e=1}^{m} e^{\theta_\nu^{(e)}+\sigma_i^{(e)}} \quad .$$

With the notations

(6.4) $$\theta_\nu = (\theta_\nu^{(1)},\dots,\theta_\nu^{(m)}), \quad \sigma_i = (\sigma_i^{(1)},\dots,\sigma_i^{(m)})$$

the version (3.13) of (3.7) takes on the form

(6.5) $$p\{a_{\nu i}\} = \frac{1}{\gamma_{\nu i}} \cdot e^{a_{\nu i}(\theta_\nu+\sigma_i)^*} ,$$

when $\gamma_{\nu i}$ of course may be considered as a function of $\theta_\nu+\sigma_i$.

Thus for any combination of ν and i the distribution of the answers over U is governed by a parameter

(6.6) $$\zeta_{\nu i} = \theta_\nu+\sigma_i .$$

Removing an additive constant, from θ_ν and σ_i, e.g. $\theta_\nu^{(m)}$ and $\sigma_i^{(m)}$, and expanding the idea somewhat we may say that the answer distribution is characterized by a parameter of the same dimension as the person and item-parameters, of which it is a function. The special feature of (6.5) then is, that the dimension of the three parameters is maximal and that the parameters can so be chosen that the general function

(6.7) $$\zeta_{\nu i} = \mu(\theta_\nu,\sigma_i)$$

reduces to the sum (6.6).

Consider now the more general case where the dimensions of ζ,

θ and σ are still equal $(=r)$, but possibly $< m-1$. In that case (4.5) is replaced by

$$p\{x^{(e)}|\nu,i\} = f_e(\zeta_{\nu i}). \tag{6.8}$$

Let us investigate the conditions for complete separability of the parameters for persons and for items, requiring, however, for partial separability somewhat more then above (sect. 4), viz. that (4.6) can be expressed in terms of an answer parameter $\zeta'_{\nu i}$ which in relation to V has the same properties as $\zeta_{\nu i}$ according to (6.7) has in relation to U, i.e.

$$\zeta'_{\nu i} = \mu'(\theta'_\nu,\sigma'_i) , \tag{6.9}$$

where ζ', θ' and σ' are equidimensional parameters characterizing respectively answers, persons and items, μ' being a fixed vector function.

In order to derive a necessary condition we shall apply theorem 2 to the partial separability for the pairs (5.2) and conclude that for each pair we may have

$$\zeta'_{\nu i} = \theta'_\nu + \sigma'_i \tag{6.10}$$

where θ'_ν and σ'_i are scalar functions of respectively θ_ν and σ_i. Thus for each $e = 1,\dots,m-1,m$

$$\frac{p\{x^{(e)}|U\}}{p\{x^{(m)}|U\}} = e^{\phi_e(\theta_\nu)+\psi_e(\sigma_i)} \text{, say,} \tag{6.11}$$

interpreting $\phi_m(\theta)$ and $\psi_m(\sigma)$ as 0.

However, due to (6.8) and (6.7) we have when dropping suffixes ν and i

$$\frac{f_e(\mu(\theta,\sigma))}{f_m(\mu(\theta,\sigma))} = e^{\phi_e(\theta)+\psi_e(\sigma)} \tag{6.12}$$

or, say,

$$(6.13) \qquad g_e(\mu(\theta,\sigma)) = \log \frac{f_e(\mu(\theta,\sigma))}{f_m(\mu(\theta,\sigma))} = \phi_e(\theta)+\psi_e(\sigma)$$

Considering first the r first of these equations we may - with proper enumeration - take

$$(6.14) \qquad (\phi_1(\theta),\ldots,\phi_r(\theta)) = (\Theta^{(1)},\ldots,\Theta^{(r)}) = \Theta$$

and

$$(6.15) \qquad (\psi_1(\sigma),\ldots,\psi_r(\sigma)) = (\Sigma^{(1)},\ldots,\Sigma^{(r)}) = \Sigma$$

as new parameters instead of θ and σ. μ then becomes a function of Θ and Σ, but since the vector

$$(6.16) \qquad (g_1,\ldots,g_r) = \Theta + \Sigma,$$

μ turns out to be some vector function of $\Theta + \Sigma$, say

$$(6.17) \qquad \mu(\theta,\sigma) = M(\Theta + \Sigma).$$

Next we insert this result in the rest of the quations (6.13), at the same time, however, utilizing that θ and σ according to (6.14) and (6.15) may be expressed in terms of Θ and Σ and that the same therefore holds for $\phi_e(\theta)$ and $\psi_e(\sigma)$, $e = r+1,\ldots,m-1$. Thus we may write

$$(6.18) \qquad g_e(M(\Theta + \Sigma)) = \Phi_e(\Theta) + \Psi_e(\Sigma), \text{ say.}$$

If in these equations we in turn differentiate with regard to Θ^* and Σ^* we get

$$(6.19) \qquad \frac{dM(\Theta+\Sigma)}{d(\Theta+\Sigma)^*} \cdot \frac{dg_e(M)}{dM^*} = \frac{d\Phi_e(\Theta)}{d\Theta^*} = \frac{d\Psi_e(\Sigma)}{d\Sigma^*}$$

from which it follows that

$$(6.20) \qquad \phi_e(\theta) + \psi_e(\sigma) = \Phi_e(\Theta) + \Psi_e(\Sigma) = \alpha_e + (\Theta+\Sigma)\beta_e^*$$

where α_e is a scalar and β_e a vectorial constant, an equation which on proper interpretation of α_e and β_e also holds for $e = 1,\ldots,r$ and $e = m$.

Going back to (6.12) we now have

$$(6.21) \qquad f_e(\mu(\theta,\sigma)) = f_m(\mu(\theta,\sigma))e^{\alpha_e+\beta_e(\Theta+\Sigma)^*}$$

and summation over $e = 1,\ldots,m$ determines f_m as

$$(6.22) \qquad f_m(\mu(\theta,\sigma)) = \frac{1}{\sum_{e=1}^{m} e^{\alpha_e+\beta_e(\Theta+\Sigma)^*}},$$

Realizing that we might have chosen, to begin with, Θ and Σ as parameters and therefore in the end replacing them by θ and σ, we obtain for the probability (6.8)

$$(6.23) \qquad p\{x^{(e)}|\nu,i\} = \frac{1}{\gamma(\theta_\nu+\sigma_i)} \cdot e^{\alpha_e+\beta_e(\theta_\nu+\sigma_i)^*}$$

where

$$(6.24) \qquad \gamma(\theta) = \Sigma\, e^{\alpha_e+\beta_e\theta^*}$$

the similarity to and distinction from the Darmois-Koopman class of distributions will be recognized.

In so far as α_e and β_e are considered as known constants the model

(6.23) allows for separability of the θ_ν's and the σ_i's.

In fact, using the selection vector (3.12) the m equations (6.23) may be condensed into

$$(6.25)\qquad p\{a_{\nu i}\} = \frac{1}{\gamma(\theta_\nu+\sigma_i)} \cdot e^{a_{\nu i}(\alpha^* + \beta^*(\theta_\nu+\sigma_i)^*)}$$

with the notations

$$(6.26)\qquad \alpha^* = \begin{pmatrix} \alpha_1 \\ \text{-----} \\ \alpha_m \end{pmatrix}, \qquad \beta^* = \begin{pmatrix} \beta_{11},\ldots,\beta_{1r} \\ \text{----------} \\ \beta_{m1},\ldots,\beta_{mr} \end{pmatrix}.$$

From the stochastic independence of the $a_{\nu i}$'s it follows that the probability of the whole response matrix$((a_{\nu i}))$ is

$$(6.27)\qquad p\{((a_{\nu i}))\}$$

$$= \frac{1}{\prod\limits_{(\nu)} \prod\limits_{(i)} \gamma(\theta_\nu+\sigma_i)} \cdot e^{a_{oo}\alpha^* + \sum\limits_{(\nu)} a_{\nu o}\beta^*\theta_\nu^* + \sum\limits_{(i)} a_{oi}\beta^*\sigma_i^*}$$

where $a_{\nu o}$ and a_{oi} are the marginals (3.15) and (3.20) while

$$(6.28)\qquad a_{oo} = \sum_{(\nu)} a_{\nu o} = \sum_{(i)} a_{oi} .$$

Accordingly the joint probability of the marginals becomes

$$(6.29)\qquad p\{(a_{\nu o}),\ (a_{oi})\}$$

$$= \frac{\begin{bmatrix} (a_{\nu o}) \\ (a_{oi}) \end{bmatrix}}{\prod \prod \gamma(\theta_\nu+\sigma_i)} \cdot e^{a_{oo}\alpha^* + \sum\limits_{(\nu)} a_{\nu o}\beta^*\theta_\nu^* + \sum\limits_{(i)} a_{oi}\beta^*\sigma_i^*} ,$$

the bracket just as in sect. 3 indicating the number of matrices of selection vectors with the two sets of marginals $(a_{\nu o})$ and (a_{oi}). Summing over those combinations of the a_{oi}'s that are compatible with a fixed set $(a_{\nu o})$ we get the marginal distribution

$$p\{(a_{\nu o})\} = \frac{e^{a_{oo}\alpha^{*} + \sum\limits_{(\nu)} a_{\nu o}\beta^{*}\theta^{*}_{\nu}}}{\Pi\ \Pi\ \gamma(\theta_{\nu}+\sigma_{i})} \cdot \gamma((\sigma^{*}_{i})|(a_{\nu o})) \tag{6.30}$$

where the second factor on the right only depends on the σ_i's:

$$\gamma((\sigma_i)|(a_{\nu o})) = \sum_{(a_{oi})} e^{\Sigma a_{oi}\beta^{*}\sigma^{*}_{i}} . \tag{6.31}$$

Finally, on dividing this expression into (6.29) we obtain the probability of the set (a_{oi}) as conditioned by the set $(a_{\nu o})$:

$$p\{(a_{oi})|(a_{\nu o})\} = \frac{\begin{bmatrix}(a_{\nu o})\\(a_{oi})\end{bmatrix}}{\gamma((\sigma_i)|(a_{\nu o}))} \cdot e^{\sum\limits_{(i)} a_{oi}\beta^{*}\sigma^{*}_{i}} , \tag{6.32}$$

which leads to estimating the item parameters irrespective of the person parameters.

Symmetrically we have of course,

$$p\{(a_{\nu o})|(a_{oi})\} = \frac{\begin{bmatrix}(a_{\nu o})\\(a_{oi})\end{bmatrix}}{\gamma((\theta_\nu)|(a_{oi}))} \cdot e^{\sum\limits_{(\nu)} a_{\nu o}\beta^{*}\theta^{*}_{\nu}} \tag{6.33}$$

which allows for the estimation of the θ_ν's irrespective of the σ_i's.

Thus the two sets of parameters may be separated [4].

Clearly this argument runs parallel to that of sect. 3 and the similarity may even be carried a step further. Introducing in (6.23) the m-dimensional parameters ξ_ν and ε_i defined by

$$\log \xi_\nu^{(e)} = \alpha_e + \beta_e \theta_\nu^* , \quad e = 1,\dots,m \tag{6.34}$$

and

$$\log \varepsilon_i^{(e)} = \beta_e \sigma_i^* , \quad e = 1,\dots,m \tag{6.35}$$

it takes on just the form of (3.7), and (6.32) is transformed to (3.22). Thus we may consider (3.7) as the fundamental form of the model which we in practice may have to try out as a first step in an analysis. Being satisfied that the data anyhow may be represented by (3.7), i.e. by taking r = m-1, and then estimating the ξ_ν's and the σ_i's, the subsidiary question arises: Is it possible to reduce the parameters to parameters of a lower dimension and still retain the complete separability?

The answer is: If that is possible, then $\log \xi_\nu^{(e)}$ and $\log \varepsilon_i^{(e)}$ must be linear functions of the new parameters θ_ν and σ_i.

It remains to be seen that the separability is complete. Consider in the model (3.7) a set $V \subset U$ of m' elements $x^{(e)}$ and the corresponding partial vectors ξ'_ν and ε'_i of ξ_ν and ε_i. According to sect. 4 (ξ'_ν) and (ε'_i) are separable in the submodel (4.2). However, for these partial vectors of (ξ_ν) and (ε_i) those equations (6.34) and (6.35) for which $x^{(e)} \in V$ must hold. Now the corresponding vectors β_e form a coefficient matrix β' of the order (m',r), but the rank r' of this matrix may be less than r. Accordingly β' may be split into the product of two matrices of the orders (m',r') and (r',r), both having the rank r':

$$\underset{(m',r)}{\beta'} = \underset{(m',r')}{\gamma} \cdot \underset{(r',r)}{\delta} . \tag{6.36}$$

Thus for each of the relevant vectors we have

$$\underset{(1,r)}{\beta_e} = \underset{(1,r')}{\gamma_e} \; \underset{(r',r)}{\delta} \tag{6.37}$$

and therefore

$$\left.\begin{aligned} \log \xi_\nu^{(e)} &= \alpha_e + \gamma_e \cdot \delta\theta_\nu^* \\ \log \varepsilon_i^{(e)} &= \gamma_e \cdot \delta\sigma_i^* \end{aligned}\right\} \quad \text{for } x^{(e)} \in V . \tag{6.38}$$

It follows that

$$\theta_\nu'^* = \delta\theta_\nu^* \quad \text{and} \quad \sigma_i'^* = \delta\sigma_i^* \tag{6.39}$$

may be taken as those functions of θ_ν and σ_i which can be separated on the basis of observations limited to V.

The results of this section may be summed up in a generalization of theorem 5:

Theorem 6:

Assuming that:

1. *the answers of different persons to a set of different items are independent stochastic variables;*
2. *for each combination (ν,i) of person and item the probability distribution over the set U of m possible answers $\{x^{(1)},\dots,x^{(m)}\}$ is defined as m functions of an r-dimensional parameter $\zeta_{\nu i}$ $(r \leq m-1)$;*
3. *the parameter $\zeta_{\nu i}$ is a vector function of two parameters θ_ν and σ_i, both vectors of dimension r, characterizing respectively person no. ν and item no. i;*

   then the necessary and sufficient condition for complete separability of the two sets of parameters, (θ_ν) and (σ_i) is that the mentioned probability distribution is of the form (6.23).

7. SEPARABILITY, MEASUREMENT, AND OBJECTIVITY.

The structure of the model (6.23) is quite remarkable. First we notice that the psychological investment of the preceding sections is purely incidental, being due to the historical fact that the type of problems it deals with first became manifest in psychology and that so far most of the applications have been restricted to this and related fields. But from a formal point of view "persons" and "items (or questions)" may be replaced by any set of "objects", respectively of "agents", any "contact" or "interaction" of an object and an agent resulting in a "response" belonging also to a set which may be finite or infinite. The aim of producing the contacts is to achieve comparisons between the objects (and/or the agents) by means of their interactions with the agents (or the objects). With a view to this each object, each agent as well as each contact is assumed to be fully characterized by a vector parameter, denoted respectively by the letters θ, σ and ζ. Already this assumption guarantees the uniqueness of ζ as a function of θ and σ. Adding to this the probabilistic assumption (6.8) - together with the stochastic independence of the responses - and the equidimensionality [x]) of θ, σ and ζ, then we have everything that is needed for the proof of theorem 6.

Accordingly this theorem stands a chance of being applicable in literally any field of science where the observations are qualitative and, with a simple modification, also when the observations are quantitative.

Of course, the theorem does not tell that responses to a certain set of interactions do follow a probabilistic law like (6.23), but it does tell that *whether observations come from Physics, Psychology, Social Sciences or Humanities, and whether they be quantitative or qualitative, gives no a priori reason for believing in or abolishing methods founded upon strong probabilistic models like (6.23) and its consequences.*

---

x) The discussion of this assumption I have to defer to some other occasion.

In Psychometrics measurement is often characterized as "the art of assigning figures to qualities" or the like. And discussing the foundations of measurement rarely goes beyond the axiomatics of categorization, of the natural numbers and of positive as well as all real numbers. With a few exceptions, such as the Beaufurt scale of wind force and the scale of hardness, measuring in Physics is defined in relation to an established physical law. By way of an example I may refer to Cl. Maxwell's analysis of mass and force [3] and [2], chapt. VII, according to which the introduction of these two concepts and of the measurements attached to them is in principle based upon the empirical fact that the acceleration suffered by a rigid body when translated by a mechanical device is proportional to the product of two factors, one referring to the body - the reciprocal of its "mass" - the other one referring to the instrument - the "force" applied. Thus the two concepts are derived simultaneously from the interactions of bodies and instruments.

The model (6.23) opens up similar possibilities in fields where the observations are qualitative, but may be mapped upon probability distributions. Clearly, if we could determine the probabilities $p\{x^{(e)}|\nu,i\}$ then from

$$(7.1) \qquad \log p\{x^{(e)}|\nu,i\} = -\log \gamma_{\nu i} + \alpha_e + \beta_e(\theta_\nu+\sigma_i)$$

for a number of objects and agents a simultaneous determination of the θ_ν's and the σ_i's would be accessible provided the α_e's and the β_e's were known in advance. *Thus in principle a measurement of object parameters and agent parameters in the same sense as in Physics has been established* as soon as the validity of the model has been ascertained [4].

For this reason I have characterized (6.23) as *a model for measuring.*

In practice the probabilities cannot be determined, but the parameter estimations indicated in sect. 6 as a substitute.

The coefficients α_e and β_e may sometimes be derived from theoretical considerations. As a case in point I may mention the multiplicative

Poisson Law represented in [3], chapt. II and VIII,

$$(7.2) \qquad p\{a|\nu,i\} = e^{-\lambda_{\nu i}} \cdot \frac{\lambda_{\nu i}^{a}}{a!} \quad ; \quad \lambda_{\nu i} = \xi_\nu \varepsilon_i$$

where a may be the number of misreadings committed by pupil no. ν in text no. i.

If no apriori knowledge about the coefficients is available they may be inferred from the data, but that presents a new class of problems which I have to leave at present.

A further conclusion from the model is laid down in the formulae (6.32) and (6.33), i.e. that the estimation of the agent parameters may be directed in such a way that it is unaffected by the lacking knowledge of the object parameters, and the reverse.

The kind of argument that lead from (6.29) to (6.32) and (6.33) may even be applied once more, thus leading to the elimination of all parameters, but two in both cases:

$$(7.3) \qquad p\{a_{oh}, a_{oj} | (a_{\nu o}), (a_{oi})'\} = \begin{bmatrix} (a_{\nu o}) \\ (a_{oi}) \end{bmatrix} \cdot \frac{\varepsilon_h^{a_{oh}} \cdot \varepsilon_j^{a_{oj}}}{\gamma(\varepsilon_h, \varepsilon_j | (a_{\nu o}), (a_{oi})')}$$

and

$$(7.4) \qquad p\{a_{\lambda o}, a_{\mu o} | (a_{oi}), (a_{\nu o})'\} = \begin{bmatrix} (a_{\nu o}) \\ (a_{oi}) \end{bmatrix} \cdot \frac{\xi_\lambda^{a_{\lambda o}} \cdot \xi_\mu^{a_{\mu o}}}{\gamma(\xi_\lambda, \xi_\mu | (a_{oi}), (a_{\nu o})')}$$

where $(a_{oi})'$ indicates the set of all a_{oi}'s except a_{oh} and a_{oj} and analogously for $(a_{\nu o})'$. Thus it is possible to separate any two ε's as well as any two ξ's from all other parameters in the system under consideration.

At this stage I wish to remind you of a word which is very well known and very much used - in a lot of different meanings, the word *objectivity*. On this occasion I shall not enter upon a historicalphilosophical discussion of this term but just point out that objectivity seems to be an almost ubiquitous request as regards "scientific statements". A precise formalization of objectivity will be given elsewhere. At present I shall draw the attention to (7.3) as a comparative statement about the two agents A_h and A_j, based upon the whole matrix of stochastic variables $((a_{\nu i}))$, but mathematically independent of all other parameters than just those to be compared - independent of all the ξ's and of all the ε's, but ε_h and ε_j, and therefore independent of the corresponding θ's and σ's.

The model (6.23) of course only holds for a certain class $\mathcal{O}$ of objects O_ν contacting a certain class $\mathcal{A}$ of agents A_i, the set $\mathcal{R}$ of responses being fixed. Theorem 6 tells that *the invariance* just pointed out *holds for any set of elements* $O_1,\dots,O_n \in \mathcal{O}$ *and any set of elements* $A_1,\dots,A_k \in \mathcal{A}$. Therefore we shall qualify the statement (7.3) about A_h and A_j, based upon $((a_{\nu i}))$, as *specifically objective* [1], [5], [6], [7], [8] - the qualification specific being chosen in order to distinguish between this particular kind of objectivity and other senses of that word.

Clearly (7.4) is also specifically objective as a statement about O_λ and O_μ, based upon $((a_{\nu i}))$.

When, in the title, I speak about consequences the mathematical theory of objectivity may have upon the choice of models I am referring to the demand for objectivity in the conclusions. If the objectivity is understood to be specific and if this specificity is presumed to be complete and the parameters have the same dimension then the modeltype (6.23) is in fact the only possibility for independent $a_{\nu i}$'s.

Thus, *if a set of empirical data cannot be described by that model then complete specifically objective statements cannot be derived from them.*

Firstly, the failing of specific objectivity means that the conclusion about, say, any set of person parameters will depend on which other persons are also compared. As a parody we might think of the comparison of the volumes of a glass and a bottle as being influenced by the heights of some of the books on a shelve.

Secondly, the conclusions about the persons would depend on just which items were chosen for the comparison, a situation to which a parallel would be, that the relative height of two persons would depend on whether the measuring stick was calibrated in inches or in centimeters.

Avoiding such irrelevant dependencies is just my reason for recommending the use of the models for measuring whenever they may be utilized. Which, by the way, is not at all always - but that raises another class of problems.

Literature

[1] Gerhard H. Fischer, Stochastische Testmodelle. Psychologische Testtheorie, 3. Kap. Ed. Gerhard H. Fischer. Huber. Bern und Stuttgart. 1968.

[2] J. Cl. Maxwell, Matter and Motion. London 1876. Reprinted in Dover Series.

[3] G. Rasch, Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danmarks Paedagogiske Inst., 1960.

[4] G. Rasch, On General Laws and the Meaning of Measurement in Psychology. Proc. Fourth Berk. Symp. Math. Stat. and Prob. IV 321-34. Berkeley: Univ. of Calif. Press., 1961.

[5] G. Rasch, An Individualistic Approach to Item Analysis. In: Reading in Math. Social Science. 89-107. Chicago: Science Research Associates, Inc., 1966.

[6] G. Rasch, An Item Analysis Which Takes Individual Differences into Account. Br. J. of Math. and Stat. Psych. 19, 49-57, 1966.

[7] G. Rasch, An Informal Report of Objectivity in Comparisons. In "Psychological measurement theory". Proceedings of the NUFFIC international summer session in science at "Het

Oude Hof", Den Haag, Juli 14-28, 1966. Red. L.J.Th. van der Kamp and C.A.J.Vlek, Leyden 1967.

[8] Jon Stene, Einführung in Rasch's Theorie der psychologischen Messung. Psychologische Testtheorie, 9. Kap. Ed. Gerhard H. Fischer. Huber. Bern und Stuttgart. 1968.