PSYCHOMETRIKA—VOL. 33, NO. 2
JUNE, 1968

THE THEORY OF ABSTRACT PARTIALS: AN INTRODUCTION

WILLIAM W. ROZEBOOM

UNIVERSITY OF ALBERTA

It is customary in multivariate analysis to search out relational structure by partitioning measures of total variation or overall relatedness into components variously attributable to different sources. Although the multivariate systems most familiar today—notably, linear correlational analysis, analysis of variance, and Information theory—base their decompositions upon very different substantive quantities, they all embody the very same abstract mathematical structure which, moreover, is capable of generating potentially fruitful patterns of data analysis in many other substantive applications as well.

Possibly the single most important achievement of modern scientific methodology has been the development of analytic systems for disclosure and dissection of relational patterning within complex multivariate data, notably, linear correlational analysis, Fisherian analysis of variance, and, most recently, Information theory (Uncertainty analysis). Although previous writers have pointed out parallels and near-equivalences between certain aspects of these systems (notably, Garner & McGill, 1956; Ross, 1962), it has not as yet become generally recognized that in the main, they are all but different interpretations of the very same abstract mathematical structure, which, moreover, is also capable of potentially fruitful application to areas into which it has not so far penetrated. In Sections 1 and 2 of this paper I shall exhibit this structure, here called the *theory of abstract partials*, in disembodied mathematical purity. Subsequently (Section 3) we shall see how it is variously instanced by Information theory, analysis of variance, the analysis of conditional probabilities, and the system of partial statistics comprised by linear correlational analysis (whence the title "abstract partials").

Since the abstract theory is not altogether easy to follow in places, the reader is urged to refer to its substantive instances in Section 3 whenever it is at all helpful to do so. Moreover, those portions of the abstract system which are most interesting mathematically have only dubious practical importance beyond adding depth to the concept of "interaction," and the reader may skim or omit Section 2 without major handicap to his grasp of the remainder.

1. *The Fundamental System*

A system of abstract partials arises whenever we have defined a measure function G with the two following properties: (*a*) There exists a domain $\mathbf{d}$ of entities such that for any element $d^{(n)}$ of any n-fold product set $\mathbf{d}^n$ (n = 1, 2, $\cdots$) on $\mathbf{d}$—i.e., $d^{(n)}$ is any ordered n-tuple of not-necessarily-distinct elements in $\mathbf{d}$—G maps $d^{(n)}$ into a number. (*b*) The value of G is unaffected by permutation of the $\mathbf{d}$-elements in its argument—i.e., if $d_i^{(n)}$ and $d_j^{(n)}$ contain each different element the same number of times, then $G(d_i^{(n)}) = G(d_j^{(n)})$. (Both of these conditions can be relaxed further, but with one exception to be described later, additional generality would needlessly complicate the present discussion.) I shall call a function G which has these properties a *generating function* over domain $\mathbf{d}$.

If it were not for possible duplications among the elements in its argument, a generating function could be characterized as a measure over subsets of its domain. I shall adopt this manner of speech anyway, with the understanding that a given "subset"—call it a "quasi-set" or *q-set* for short to distinguish it from an ordinary set—may include multiple occurrances of the same element. q-sets from domain $\mathbf{d}$ will here be designated by upper-case letters X, Y, Z, etc., and individual elements of $\mathbf{d}$ by lower-case letters x, y, etc. A concatenation of symbols for q-sets and $\mathbf{d}$-elements—e.g., XY, $xYzW$, $x_1x_2Z_1Z_2$—will designate the q-set comprising all elements indicated in the concatenation, each taken as many times as it is mentioned. Thus if X is the q-set $x_1 \cdots x_m$, Y is the q-set $y_1 \cdots y_n$ and z is a single element, XYz is the q-set (i.e., permutable ordered set) $x_1 \cdots x_m y_1 \cdots y_n z$. According to this convention, an element of $\mathbf{d}$ may be thought of as a q-set of unit length.

Using the notation just introduced, the permutability condition stipulated to hold for a generating function G may be written

$$G(WxyZ) = G(WyxZ), \tag{1}$$

in which W, Z, or both may be empty. Iteration of [1] yields the required indifference of G to the order of elements in its argument.

In most statistical models of the abstract-partials system, domain $\mathbf{d}$ is a set of jointly distributed scientific variables (i.e., "variates") while G is some scalar-valued multivariate statistic. However, there is nothing inherently statistical about abstract partials. For example, if $\mathbf{d}$ is a set of numbers, one generating function over $\mathbf{d}$ is the numerical product of any $n(n \geq 1)$ not-necessarily-distinct numbers in $\mathbf{d}$. For a more interesting example, let each element $x_i \in \mathbf{d}$ be a particular denomination of postage stamp as these are classified by philatelists. Then $G(X)$ might be the average market value of a set $X = x_1 \cdots x_n$ of n stamps whose denominations are $x_1, \cdots, x_n$, respectively. Either of these examples should make clear the sense in which duplications of $\mathbf{d}$-elements are admissible in G's argument. Thus in the stamp instance, $G(x_1x_1x_1x_2x_2)$ would be the market value of a set of five stamps, three of which are of denomination x_1 while the other two are of denomination x_2. Incidently, a system of market values for commodity bundles, as in the postage-stamp example, embodies the full abstract-partials structure in a fashion which can not only be comprehended with a minimum of effort, thus affording an excellent didactic model of the abstract theory, but may also have value for serious research in economics and the psychology of choice.

The quantity $G(X)$ assigned by generating function G to an argument X will be called the *G-value* of q-set X. In the postage-stamp model, the G-value of stamp collection X (more precisely, of a collection containing stamps with denomination-configuration X) is its market price.

So far, nothing has been said about a generating function's value for an argument containing *no* elements of its domain. Under most primary interpretations of G, the quantity $G(X)$ remains undefined when X is empty and can be assigned an arbitrary value. Since proofs and statements of results are greatly expedited if the null q-set has a G-value of zero, we let "ϕ" designate the null set and stipulate that

[2] $$G(\phi) = 0.$$

Note that for any q-set X, $X\phi = \phi X = X$.

Once a generating function G over q-sets from a domain **d** has been identified, we are in position to extract an outrageously abundant array of relations among the elements of **d** with respect to G. We begin with the conditional G-value, $G(X \mid Z)$, of *q*-set X given q-set Z, defined as

[3] $$G(X \mid Z) = G(XZ) - G(Z).$$

(Since the conditional G-function takes two q-set arguments which enter asymmetrically, it differs importantly in mathematical form from generating function G, and it is not strictly proper to use the same function-symbol "G" for both. That more is gained than lost by this impropriety, however, will soon be evident.) The conditional G-value of X given Z is in effect the G-value of X as a supplement to Z, i.e., the increment in G-value which occurs when Z is augmented by X. (When $G(X)$ is the market value of an assortment X of stamps, $G(X \mid Z)$ is how much extra it would cost to buy assortment X on the same occasion that assortment Z is purchased.) While the adjective "conditional" is here applied to $G(X \mid Z)$ in a generic sense which will recur, it is more insightful in most interpretations to think of $G(X \mid Z)$ as the "residual" G-value of X when Z is given. For reasons to appear shortly, $G(X \mid Z)$ will also be described as a "partial G-value of order r," where r is the number of elements in Z. From [2] and [3] we have

[4] $$G(X \mid \phi) = G(X),$$

so unconditional G-values are partial G-values of zero order.

The *configural savings*, $C(X)$, of q-set X with respect to measure G is now defined as the algebraic amount by which the G-value of X fails to be a simple sum of the G-values of its elements. More generally, the conditional (partial) configural savings of q-set $x_1 \cdots x_n$, given q-set Z, with respect to G is defined as

[5] $$C(x_1 \cdots x_n \mid Z) = \sum_{i=1}^{n} G(x_i \mid Z) - G(x_1 \cdots x_n \mid Z),$$

the zero-order case of which is

[6] $$\begin{aligned} C(x_1 \cdots x_n) &= C(x_1 \cdots x_n \mid \phi) \\ &= \sum_{i=1}^{n} G(x_i) - G(x_1 \cdots x_n). \end{aligned}$$

When the G-value of a q-set X is analyzed into constituent contributions, the configural savings of X (with respect to G) may be construed as the amount of G which is subtracted from the aggregate contributions to $G(X)$ of the individual elements in X by the patterning which emerges when these are G-evaluated jointly. The quantity $C(X)$ may also be thought of as the "negative gestalt" value of configuration X, since in a sense it represents the degree to which, regarding G, a whole is less than the sum of its parts. In the postage-stamp model, $C(X)$ would be an especially interesting measure were serious research to be done on the sources of philatelist values, for in this and similar economic situations, the monetary worth of a bundle of commodities tends to be the sum of the values of its constituents, and the configural savings for various assemblages of stamps could be expected to result rather cleanly from a small number of identifiable patterning effects, such as a tendency for duplications to depreciate a collection's value below what the same aggregate of individual values invested in all different denominations would be worth (i.e., positive configural savings), or for a collection's value to become enhanced as it nears completion with respect to one or more of the goals that philatelists prize (i.e., negative configural savings). In multivariate analysis, the configural-savings measure is of very recent origin, having first appeared as "total correlation" [Watanabé, 1960] or "total constraint" [Garner, 1962] in Information theory, and remaining so far unrecognized as such in the older multivariate systems although its analysis-of-variance interpretation, namely, total interaction, is also a familiar concept.

Next we have the *G-contingency*, $R(X; Y)$, of q-set X upon q-set Y. This is the algebraic amount by which the G-value of X exceeds its conditional G-value given Y. More generally, the conditional (partial) G-contingency of X upon Y, given Z, is defined as

[7] $$R(X; Y \mid Z) = G(X \mid Z) - G(X \mid YZ),$$

of which $R(X; Y)$ is the zero-order case

[8] $$\begin{aligned} R(X; Y) &= R(X; Y \mid \phi) \\ &= G(X) - G(X \mid Y). \end{aligned}$$

In virtually all interpretations, G-contingency is a measure of relationship, and has probably the greatest importance of all the abstract partials. In the postage-stamp model, the G-contingency of one assortment of stamps upon another is the amount by which the price of the first is reduced (or enhanced if R is negative) by being sold as a supplement to Y in a package deal.

Finally, in many—though not all—interpretations of the abstract-partials system, a q-set $x_1 \cdots x_n$ containing $n \geq 2$ elements of **d** cannot be distinguished by measure G from a certain single entity $^{|}\overline{x_1 \cdots x_n}^{|}$ which is either an element of **d** to begin with or can be added to **d** without changing the latter's logical character. For example, in addition to treating a collection X of stamps as an ensemble of items whose market value is constituted in some perhaps complex fashion out of the values of its components, we can also conceive of this collection as a single commodity $^{|}\overline{X}^{|}$ which can be bought and sold as a unit and whose market value is a datum which can be processed in the same way as data about the values of individual stamps. The single element $^{|}\overline{X}^{|}$ coordinated with q-set X is governed by the axioms

[9F] $$^{|}\overline{X}^{|} \varepsilon \mathbf{d},$$

[10F] $$G(^{|}\overline{X}^{|}Y) = G(XY),$$

and will here be called the *fusion* of X. Those theorems below which pertain to fusion elements will be labeled "F." For interpretations in which the "fusion" concept is not defined, the F-equations are not false but simply meaningless.

As immediate, or almost immediate, consequences of the above axioms and definitions we have the following basic theorems, in which X, Y, etc. are any q-sets from **d** containing zero or more elements and x, y, $^{|}\overline{X}^{|}$, etc. are **d**-elements.

[11] $$G(X \mid YZ) = G(XY \mid Z) - G(Y \mid Z)$$

[12] $$G(X\phi \mid Z) = G(X \mid \phi Z) = G(X \mid Z)$$

[13] $$G(\phi \mid Z) = 0$$

[14] $$C(x \mid Z) = 0$$

[15] $$R(X; \phi \mid Z) = 0$$

[16] $$\begin{aligned} R(X; Y \mid Z) &= G(X \mid Z) + G(Y \mid Z) - G(XY \mid Z) \\ &= C(XY \mid Z) - C(X \mid Z) - C(Y \mid Z) \end{aligned}$$

[17] $$R(X; Y \mid Z) = R(Y; X \mid Z)$$

[18] $$R(x; y \mid Z) = C(xy \mid Z)$$

$$[19] \quad R(X_1X_2\,;\,Y_1Y_2 \mid Z) + R(X_1\,;X_2 \mid Z) + R(Y_1\,;\,Y_2 \mid Z) = R(X_1Y_1\,;\,X_2Y_2 \mid Z) + R(X_1\,;\,Y_1 \mid Z) + R(X_2\,;\,Y_2 \mid Z)$$

$$[20] \quad R(X; Y \mid WZ) = R(X; YW \mid Z) - R(X; W \mid Z)$$

$$[21] \quad C(X \mid YZ) = C(XY \mid Z) + (n - 1)C(Y \mid Z) - \sum_{i=1}^{n} C(x_iY \mid Z)$$

$$= C(XY \mid Z) - C(Y \mid Z) - \sum_{i=1}^{n} R(x_i\,;\,Y \mid Z)$$

$$= C(X \mid Z) + R(X; Y \mid Z) - \sum_{i=1}^{n} R(x_i\,;\,Y \mid Z) \qquad (X = x_1 \cdots x_n)$$

$$[22F] \quad C({}^{|}\overline{X}{}^{|}Y \mid Z) = C(XY \mid Z) - C(X \mid Z)$$

$$[23F] \quad R(X; Y \mid Z) = C({}^{|}\overline{X}{}^{||}\overline{Y}{}^{|} \mid Z)$$

$$[24F] \quad G({}^{|}\overline{{}^{|}\overline{W}{}^{|}X}{}^{|}Y) = G({}^{|}\overline{WX}{}^{|}Y) = G(WXY)$$

$$[25F] \quad G({}^{|}\overline{X}{}^{|}Y \mid {}^{|}\overline{W}{}^{|}Z) = G(XY \mid WZ),\ C(X \mid {}^{|}\overline{Y}{}^{|}Z) = C(X \mid YZ),$$
$$R({}^{|}\overline{X}{}^{|}Y_1\,;\,Y_2 \mid \overline{W}Z) = R(XY_1\,;\,Y_2 \mid WZ).$$

While we shall not discuss these equations in detail, a few remarks on their more salient features are in order. [11], [20], and [21] show how higher-order partials are built up from lower-order partials of the same kind. In particular, for each measure G, C, or R, respectively, all partials of that kind higher than order r can be derived wholly from rth order partials of that kind. (The reverse is not true, however—given all partials of orders higher than r, we cannot reclaim those of order r or lower.) The connection between G-contingencies and configural savings is spelled out by [16], [18], and [23], which show that the G-contingency of one q-set upon another is in effect a special case of configural savings, and that $R(X; Y)$ is also the amount by which the total configural savings in q-set XY exceeds a simple sum of the configural savings in its two parts X and Y considered separately. An obvious but important consequence of these relations, made explicit in [17], is that $R(X; Y \mid Z)$ is symmetric in X and Y, a fact not at all apparent in R's definition. A more general G-contingency invariance under interchange of elements in R's arguments is given in [19].

It also follows from the preceding results that

$$[26] \quad G(X \mid Z) = G(X \mid YZ) + R(X; Y \mid Z)$$

$$[27] \qquad R(Y; X_1 \cdots X_n \mid Z) = R(Y; X_1 \mid Z) + R(Y; X_2 \mid X_1 Z) + \cdots + R(Y; X_n \mid X_1 \cdots X_{n-1} Z)$$

$$[28] \qquad C(XY \mid Z) = C(X \mid Z) + C(Y \mid Z) + R(X; Y \mid Z)$$

$$[29] \qquad C(XY \mid Z) = C(Y \mid Z) + C(X \mid YZ) + \sum_{i=1}^{n} R(x_i ; Y \mid Z) \qquad (X = x_1 \cdots x_n)$$

$$[30] \qquad C(x_1 \cdots x_n \mid Z) = R(x_2 ; x_1 \mid Z) + R(x_3 ; x_1 x_2 \mid Z) + \cdots + R(x_n ; x_1 \cdots x_{n-1} \mid Z) \cdot$$

Equations [26], [28], and [29] are merely transpositions of [7], [16], and [21], respectively, [27] is an iterated transposition of [20], and [30] is an iteration of [28] for q-sets of unit length. These may be called the "partitioning theorems," since they analyze a given G-value or derivative measure as a sum of component quantities attributable to different sources. The basic partition is [26], the nature of which shows forth most visibly in its special case

$$G(y) = R(y; x_1 \cdots x_n) + G(y \mid x_1 \cdots x_n).$$

This says in essence that the total G-value of an element y equals a component attributable to the effect on y of other elements $x_1 , \cdots , x_n$, plus a residual G-value which persists for y even when $x_1 , \cdots , x_n$ are given—a notion highly familiar to any student of multivariate analysis. The component $R(y; x_1 \cdots x_n)$ of $G(y)$ jointly attributed to $x_1 , \cdots , x_n$ may be further analyzed by [27] into a sum of effects allocated more specifically to the various individual "predictor" elements. Apart from its special case [18], the configural-savings measure does not appear as an additive component in any partition of G, and hence cannot properly be regarded as a G-partial. When of interest for its own sake, however, C is shown by [28]–[30] to be likewise susceptible to intriguing partitions. Further, the asymmetries in [27] and [30] can be eliminated in favor of the more elaborate partitions given in [59] and [61] below.

For interpretations in which the concept of a q-set's "fusion" is meaningful, [28] and [30] are special cases, respectively, of

$$[31F] \qquad C(X_1 X_2 \cdots X_n \mid Z) = C({}^{|}\overline{X_1}{}^{|}\,{}^{|}\overline{X_2}{}^{|} \cdots {}^{|}\overline{X_n}{}^{|} \mid Z) + \sum_{i=1}^{n} C(X_i \mid Z)$$

$$[32F] \qquad C(X_1 X_2 \cdots X_n \mid Z) = C(X_1 \mid Z) + C({}^{|}\overline{X_1}{}^{|} X_2 \mid Z) + C({}^{|}\overline{X_1 X_2}{}^{|} X_3 \mid Z) + \cdots + C({}^{|}\overline{X_1 \cdots X_{n-1}}{}^{|} X_n \mid Z),$$

both of which follow from [22]. Actually, [31] and [32] are but two of an enormous number of different ways in which a configural-savings value can be iteratively partitioned by [22]. In particular, iteration of [31] yields a branching-type analysis of total configural value—e.g.,

$$
\begin{aligned}
C(X_1X_2X_3X_4) &= C(|\overline{X_1X_2}||\overline{X_3X_4}|) + C(X_1X_2) + C(X_3X_4) \\
&= C(|\overline{X_1X_2}||\overline{X_3X_4}|) + C(|\overline{X_1}||\overline{X_2}|) + C(|\overline{X_3}||\overline{X_4}|) \\
&\quad + C(X_1) + C(X_2) + C(X_3) + C(X_4),
\end{aligned}
$$

etc.—whose invariance under alternative patterns of branching has been described by Watanabé [1961] as "the fundamental theorem of ITCA [Information-theoretical correlation analysis]."

2. *The Ramified System*

Given a generating function G over a domain **d**, it is possible to derive from G additional measures G' over q-sets from **d** which also satisfy the formal requirements for a generating function and which hence bud off subsidiary systems of abstract partials based on G', including still another round of felial generating functions G'' and so on ad infinitum. The opportunities for distinctive variations in these ramifications, if not endless, at least bulge beyond the confines of any systematic treatment I am able to give them—which is one reason why the title of this paper bears the qualification, "an introduction." How much of the ramified system of abstract partials will prove to have useful application is very much an open question, and there is probably little reason to explore its more wonderous complexities other than pure mathematical curiosity. To illustrate what can be done by ramification, this section will present a generalized development of the hierarchy of "interaction" terms whose charming symmetries have previously been described in their Information-theoretical interpretation by McGill [1954] and Garner [1962]. Since we shall now be making reference to a variety of generating functions, and their derivative measures, over the same domain, all abstract partials in the same system will be identified with a common subscript—i.e., C_α is configural savings with respect to generating function G_α, R_α is G_α-contingency, and similarly for other measures based on G_α.

Consider the class of all linear functions of various orders of G_α-values, configural savings with respect to G_α, and G_α-contingencies for various q-set arguments involving a given q-set X. Since all partial G_αs, C_αs, and R_αs of all orders are themselves linear combinations of zero-order G_αs, this is the class of all functions of form

$$[33] \qquad L_{\alpha,\mathbf{a}\mathrm{Y}}(X) \stackrel{\mathrm{def}}{=} \sum_{i=0}^{m} a_i G_\alpha(XY_i) + \sum_{i=m+1}^{n-1} a_i G_\alpha(Y_i) + a_n ,$$

where $a_0, \cdots, a_m, \cdots, a_n$ are numerical constants and $X, Y_0, \cdots, Y_{n-1}$ are not-necessarily-distinct q-sets from G_α's domain. (The roman subscript "aY" will be clarified shortly.) The possibility that $Y_i = \phi$ is not excluded, and to make systematic provision for the occurrence of $G_\alpha(X)$ (i.e., $G_\alpha(X\phi)$) in [33] it is convenient to stipulate that Y_0 is the null set. This is, of course, no restriction on the generality of [33] since the term $G_\alpha(XY_0)$ can always be eliminated by setting $a_0 = 0$.

If "X" in [33] is now regarded as an argument place-holder for which any q-set from **d** may be substituted, $L_{\alpha,\mathrm{aY}}$ becomes a function over q-sets from **d** which could be taken for a generating function were it not that $L_{\alpha,\mathrm{aY}}$ generally fails to satisfy axiom [2]. This defect is easily corrected, however, by the adjustment

$$[34]\qquad \begin{aligned} G_{\alpha,\mathrm{aY}}(X) &\stackrel{\text{def}}{=} L_{\alpha,\mathrm{aY}}(X) - L_{\alpha,\mathrm{aY}}(\phi) \\ &= \sum_{i=0}^{m} a_i[G_\alpha(XY_i) - G_\alpha(Y_i)] \\ &= \sum_{i=0}^{m} a_i G_\alpha(X \mid Y_i)\cdot \end{aligned}$$

A generating function defined from G_α by an equation of form [34] may be called a *stationary linear development* of G_α. Any stationary linear development of a generating function G_α is unambiguously identified by the ordered subscript "α, aY", in which "a" and "Y" abbreviate the parameter-vectors $\langle a_0, \cdots, a_m\rangle$ and $\langle Y_1, \cdots, Y_m\rangle$, respectively. Each $G_{\alpha,\mathrm{aY}}$ then generates its own subsidiary system of abstract partials. In particular,

$$[35]\qquad \begin{aligned} G_{\alpha,\mathrm{aY}}(X \mid Z) &= G_{\alpha,\mathrm{aY}}(XZ) - G_{\alpha,\mathrm{aY}}(Z) \\ &= \sum_{i=0}^{m} a_i G_\alpha(X \mid Y_i Z) \end{aligned}$$

$$[36]\qquad \begin{aligned} C_{\alpha,\mathrm{aY}}(X \mid Z) &= \sum_{i=1}^{n} G_{\alpha,\mathrm{aY}}(x_i \mid Z) - G_{\alpha,\mathrm{aY}}(X \mid Z) \\ &= \sum_{i=0}^{m} a_i C_\alpha(X \mid Y_i Z) \qquad (X = x_1 \cdots x_n) \end{aligned}$$

$$[37]\qquad \begin{aligned} R_{\alpha,\mathrm{aY}}(X_1 ; X_2 \mid Z) &= G_{\alpha,\mathrm{aY}}(X_1 \mid Z) - G_{\alpha,\mathrm{aY}}(X_1 \mid X_2 Z) \\ &= \sum_{i=0}^{m} a_i R_\alpha(X_1 ; X_2 \mid Y_i Z)\cdot \end{aligned}$$

(The stationary linear developments of generating function G_α constitute the class of all generating functions which can be defined from G_α by linear operations with a fixed set of numerical constants and q-sets from G_α's domain as parameters. However, these by no means exhaust the possibilities for deriving subsidiary generating functions. For example, a nonstationary linear development of G_α would be a linear function such as C_α whose defining form is not fixed but varies with the number of elements in its argument.

Further, nonlinear developments of G_α after the fashion of [34] are also available in unlimited supply. In particular, the measure

$$G_{\alpha,f}(X) \stackrel{\text{def}}{=} f[G_\alpha(X)] - f[0]$$

for any numerical transformation f is also a generating function over G_α's domain.)

Now let

$$[38k] \qquad G_{\alpha,\mathrm{a}_k\mathrm{Y}_k}(X) \stackrel{\text{def}}{=} \sum_{i=0}^{n_k} a_{ki} G_\alpha(X \mid Y_{ki}) \qquad (k = 1, 2, \cdots)$$

be a not-necessarily-finite sequence of stationary linear developments of G_α in which

$$[39k] \qquad Y_{k0} \stackrel{\text{def}}{=} \phi. \qquad (k = 1, 2, \cdots)$$

Since α is actually a parameter in [38], the series can be expanded by replacing G_α at different positions k by other generating functions derived from G_α. In particular, functions $G_{\alpha,\mathrm{a}_k\mathrm{Y}_k}$ $(k = 1, 2, \cdots)$ can be nested by the recursive definition

$$[40k] \qquad (k = 1, 2, \cdots) \qquad \begin{aligned} G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_k\mathrm{Y}_k}(X) &\stackrel{\text{def}}{=} G_{(\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_{k-1}\mathrm{Y}_{k-1}),\mathrm{a}_k\mathrm{Y}_k}(X) \\ &= \sum_{i=0}^{n_k} a_{ki} G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_{k-1}\mathrm{Y}_{k-1}}(X \mid Y_{ki}), \end{aligned}$$

which gives rise to a hierarchical sequence of derivative generating functions based on parent function G_α. The number k of simple functions $G_{\alpha,a_k\mathrm{Y}_k}$ which are nested in $G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_k\mathrm{Y}_k}$ will be called the latter's *level* and represented in the notation by a parenthetical superscript whenever it is useful to do so. By induction on k, it is easily seen that

$$[41k] \qquad (k = 1, 2, \cdots) \qquad \begin{aligned} &G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_k\mathrm{Y}_k}(X \mid Z) \\ &\quad = \sum_{h=0}^{n_1} \sum_{i=0}^{n_2} \cdots \sum_{j=0}^{n_k} a_{1h} a_{2i} \cdots a_{kj} G_\alpha(X \mid Y_{1h} Y_{2i} \cdots Y_{kj} Z), \end{aligned}$$

which shows both that the class of stationary linear developments of G_α is closed under nestings of form [40] and also, more interestingly, that a hierarchically nested generating function is unaffected by permutation of the simple functions from which it is derived. That is, the parameter-couples $\mathrm{a}_1\mathrm{Y}_1, \cdots, \mathrm{a}_k\mathrm{Y}_k$ in $G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_k\mathrm{Y}_k}$ occur symmetrically. If the simple functions from which a sequence of hierarchically nested generating function is formed differ only in their q-set parameters, the sequence may be called a *homogeneous hierarchy* on G_α—i.e., $G_{\alpha,\mathrm{a}_1\mathrm{Y}_1,\cdots,\mathrm{a}_k\mathrm{Y}_k}$ $(k = 1, 2, \cdots)$ are a homogeneous hierarchy iff $\mathrm{a}_k = \mathrm{a}$ $(k = 1, 2, \cdots)$.

We now reflect that the q-set parameters $\mathrm{Y}_1, \cdots, \mathrm{Y}_k$ in a nested generating function $G_{\alpha,a_1Y_1,\cdots,a_kY_k}$ may just as well be construed as argument place-holders for which various ordered sets of q-sets from G_α's domain can be substituted. In particular, the function $H^{(k)}_{\alpha,\mathrm{a}}$ of nk + 2 q-set arguments may be defined

[42] $$H^{(k)}_{\alpha,\mathrm{a}}(\mathrm{Y}_1 ; \cdots ; \mathrm{Y}_k ; X \mid Z) \overset{\text{def}}{=} G_{\alpha,\mathrm{aY}_1,\cdots,\mathrm{aY}_k}(X \mid Z),$$

where each $\mathrm{Y}_i = \langle Y_{i1}, \cdots, Y_{in}\rangle$ is an n-component vector of q-sets from G_α's domain, $\mathrm{a} = \langle a_0, \cdots, a_n\rangle$ is an $(n + 1)$-component vector of numerical constants, n and k are any two nonnegative integers, and the absence of differentiating subscripts on the numerical parameter-vectors in the right-hand side of [42] signifies that $G_{\alpha,\mathrm{aY}_1,\cdots,\mathrm{aY}_k}$ $(k = 1, 2, \cdots)$ is a homogeneous hierarchy. Since interchanging parameter-couples aY_i and aY_j in a homogeneous hierarchy is the same as interchanging just Y_i and Y_j, $H^{(k)}_{\alpha,\mathrm{a}}(\mathrm{Y}_1 ; \cdots ; \mathrm{Y}_k ; X \mid Z)$ is invariant under any permutation of the q-set vectors $\mathrm{Y}_1, \cdots, \mathrm{Y}_k$, though not necessarily so under permutation of the q-sets *within* a given Y_i.

Finally, let ι be the ordered pair of numerical constants

[43] $$\iota \overset{\text{def}}{=} \langle -1, 1\rangle.$$

Then with $k = 1$, $a_{10} = -1$, and $a_{11} = 1$ in [41], we have

[44] $$\begin{aligned} -G_{\alpha,\iota}Y(X \mid Z) &= -[-G_\alpha(X \mid Z) + G_\alpha(X \mid YZ)] \\ &= R_\alpha(X; Y \mid Z), \end{aligned}$$

which shows that the G_α-contingency of X upon Y may be regarded as the negated value for X of G_α's first-level development with parameters ι and Y. More generally, the negated homogeneous hierarchy developed from G_α by numerical parameter ι is the series of "interaction" terms whose substantive embodiments in analysis-of-variance and, more recently, Uncertainty analysis are familiar multivariate concepts. Specifically, the kth level *interaction* among $k + 1$ q-sets $X_1, \cdots, X_{k+1}$ with respect to generating function G_α, conditional upon another q-set Z, is defined

[45] $$\begin{aligned} I^{(k)}_\alpha(X_1 ; \cdots ; X_{k+1} \mid Z) &\overset{\text{def}}{=} -H^{(k)}_{\alpha,\iota}(X_1 ; \cdots ; X_k ; X_{k+1} \mid Z) \\ &= -G_{\alpha,\iota X_1,\cdots,\iota X_k}(X_{k+1} \mid Z), \end{aligned}$$

the unconditional interactions of course being

[46] $$\begin{aligned} I^{(k)}_\alpha(X_1 ; \cdots ; X_{k+1}) &\overset{\text{def}}{=} I^{(k)}_\alpha(X_1 ; \cdots ; X_{k+1} \mid \phi) \\ &= -G_{\alpha,\iota X_1,\cdots,\iota X_k}(X_{k+1}). \end{aligned}$$

The interaction hierarchy has many interesting properties, though explicit statement of these tends to become formidable notationally. To begin,

[47] $$I^{(k)}_\alpha(X_1 ; \cdots ; X_{k+1} \mid Z) = R_{\alpha,\iota X_1,\cdots,\iota X_{k-1}}(X_k ; X_{k+1} \mid Z),$$

which follows from [45] by substituting "$\alpha, \iota X_1, \cdots, \iota X_{k-1}$" for open parameter "$\alpha$" in [46]. We have already seen on more general grounds (cf. comment about permutations following [42]) that $I_\alpha^{(k)}(X_1 ; \cdots ; X_k ; X_{k+1} | Z)$ is indifferent to the order of $X_1, \cdots, X_k$. Since by [47] and [17], X_k and X_{k+1} can also be interchanged without affecting the value of the interaction, we may write

$$[48i] \qquad I_\alpha^{(k)}(\cdots ; X_i ; X_{i+1} ; \cdots | Z)$$
$$(i = 1, \cdots, k) \qquad = I_\alpha^{(k)}(\cdots ; X_{i+1} ; X_i ; \cdots | Z),$$

which by iteration says that $I_\alpha^{(k)}(X_1 ; \cdots ; X_{k+1} | Z)$ is invariant under all permutations of its primary arguments (i.e., excluding Z). Together, [47] and [48] show that $I_\alpha^{(k)}$ is the contingency between any two of its primary arguments with respect to a certain nested generating function derived from the parent G_α and the remainder of its primary arguments. In particular,

$$[49] \qquad I_\alpha^{(1)}(X_1 ; X_2 | Z) = R_\alpha(X_1 ; X_2 | Z),$$

which states that the G_α-contingency between two q-sets is also their first-level interaction with respect to G_α. Zero-level interactions likewise merit special mention, since they are simply negated G_α-values—i.e.,

$$[50] \qquad I_\alpha^{(0)}(X | Z) = -G_\alpha(X | Z).$$

It is further instructive to note from [47], [7], and [48] that

$$[51] \qquad I_\alpha^{(k)}(X_1 ; \cdots ; X_{k+1} | Z)$$
$$= I_\alpha^{(k-1)}(X_1 ; \cdots ; X_k | X_{k+1}Z) - I_\alpha^{(k-1)}(X_1 ; \cdots ; X_k | Z)$$

and from [45] that

$$[52h] \qquad I_\alpha^{(k)}(X_1 ; \cdots ; X_{k+1} | Z)$$
$$(h = 0, \cdots, k) \qquad = I_{\alpha, \iota X_1, \cdots, \iota X_h}^{(k-h)}(X_{h+1} ; \cdots ; X_{k+1} | Z).$$

Both [51] and [52] are of course unaffected by any permutations of $X_1, \cdots, X_{k+1}$.

An additional consequence of the symmetry of $I_\alpha^{(k)}$ in its primary arguments is that the unconditional interactions among the elements of the system's domain determine still another generating function over the latter, i.e.,

$$[53] \qquad G_{I(\alpha)}(x_1 \cdots x_n) \stackrel{\text{def}}{=} I_\alpha^{(n-1)}(x_1 ; \cdots ; x_n)$$

where, since $I_\alpha^{(n-1)}$ is undefined when $n = 0$, we are free to stipulate that $G_{I(\alpha)}(\phi) = 0$. The element-interaction function $G_{I(\alpha)}$ stands in a rather special relation to its parent generating function G_α, for while $G_{I(\alpha)}$ is a nonstationary linear development of G_α and hence opens still another dimension of the ramified theory erected upon G_α, it will be shown later that there is an important sense in which G_α and $G_{I(\alpha)}$ are transformationally equivalent.*

For the remaining interaction theorems we need an operator which combines selection and summation. Let F be a function whose argument is any ordered r-tuple of entities of some kind K (here q-sets from **d**), and let "$\xi_1, \cdots, \xi_r$" ambiguously designatean "r-selection" from a more inclusive ordered n-tuple ($r \leq n$) of K-entities, namely, an ordered r-tuple of indexically different terms taken from $X_1, \cdots, X_n$ and arranged in the order of their occurrence in the latter. (Two terms X_i and X_j from n-tuple $X_1, \cdots, X_n$ are "indexically different" iff $i \neq j$. This does not preclude the possibility that $X_i = X_j$.) That is, $\xi_1, \cdots, \xi_r$ is formed from $X_1, \cdots, X_n$ by deleting $n - r$ indexically different terms from the latter. Two r-selections from $X_1, \cdots, X_n$ differ if and only if there is a term in one which is indexically different from every term in the other. From an n-tuple $X_1, \cdots, X_n$, a total of

$$\binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

different r-selections can be formed. (For example, the three different 2-selections from 3-tuple X_1, X_2, X_3 are X_1X_2, X_1X_3, and X_2X_3.) The quantity

$$\mathsf{S}^r_{X_1,\cdots,X_n}F(\xi_1, \cdots, \xi_r)$$

may now be defined to be the sum of values of function F for all the $\binom{n}{r}$ different r-selections $\xi_1, \cdots, \xi_r$ from $X_1, \cdots, X_n$. A pivotal property of the S-operator is that

[54] $$\mathsf{S}^r_{X_1,\cdots,X_{n+1}}F(\xi_1, \cdots, \xi_r) = \mathsf{S}^r_{X_1,\cdots,X_n}F(\xi_1, \cdots, \xi_r) + \mathsf{S}^{r-1}_{X_1,\cdots,X_n}F(\xi_1, \cdots, \xi_{r-1}, X_{n+1})$$

for all integers r such that $1 < r \leq n$. [54] also holds for $r = 1$ and $r > n$ if

[55] $$\mathsf{S}^r_{X_1,\cdots,X_n}F(\xi_1, \cdots, \xi_r) = \begin{cases} F & \text{if } r = 0 \\ 0 & \text{if } r > n \end{cases},$$

where for $r = 0$, $F(\xi_1, \cdots, \xi_r)$ is a function of no arguments, i.e., is a constant F. It is possible to argue that [55] follows from the S-operator's verbal definition, but since these limiting cases are conceptually fuzzy, [55] may simply be regarded as a formal definition of S when $r = 0$ or $r > n$.

It may now be seen that for any $k \geq 0$,

*It would be pleasant to supplement [53] with the theorem that $G_{I(\alpha)}(x_1 \ldots x_n|Z) = I_\alpha^{(n-1)}(x_1; \ldots; x_n|Z)$ for all Z, but unhappily this is not the case.

[56] $$I_\alpha^{(k)}(X_1 \,;\, \cdots \,;\, X_{k+1} \mid Z)$$

$$= -\sum_{r=0}^{k} \mathsf{S}^r_{X_1,\ldots,X_k}\{(-1)^{k-r} G_\alpha(X_{k+1} \mid \xi_1 \cdots \xi_r Z)\}$$

$$= \sum_{r=0}^{k} (-1)^{k-r+1}\{\mathsf{S}^r_{X_1,\ldots,X_k} G_\alpha(\xi_1 \cdots \xi_r X_{k+1} \mid Z)$$

$$- \mathsf{S}^r_{X_1,\ldots,X_k} G_\alpha(\xi_1 \cdots \xi_r \mid Z)\}$$

$$= \sum_{r=1}^{k+1} (-1)^{k-r}\{\mathsf{S}^{r-1}_{X_1,\ldots,X_k} G_\alpha(\xi_1 \cdots \xi_{r-1} X_{k+1} \mid Z)$$

$$+ \mathsf{S}^r_{X_1,\ldots,X_k} G_\alpha(\xi_1 \cdots \xi_r \mid Z)\}$$

$$= \sum_{r=1}^{k+1} (-1)^{k-r} \mathsf{S}^r_{X_1,\ldots,X_{k+1}} G_\alpha(\xi_1 \cdots \xi_r \mid Z).$$

(The first line of [56] follows from [45] and [41]; the second applies [11]; the third is a reorganization of the second with the help of [55] and [13]; and the last line follows by [54].) Equation [56] makes explicit how interactions of all levels with respect to generating function G_α are composed of G_α-values for all possible combinations of the interaction's arguments. Moreover, the hierarchy of interactions based on G_α symmetrically partitions G_α-values. Specifically,

[57] $$-G_\alpha(X_1 \cdots X_n \mid Z) = \sum_{r=1}^{n} \mathsf{S}^r_{X_1,\ldots,X_n} I_\alpha^{(r-1)}(\xi_1 \,;\, \cdots \,;\, \xi_r \mid Z),$$

which says that if a q-set X is analyzed as a concatenation of n sub-sequences $X_1 \cdots X_n$, the negated G_α-value of X, given Z, equals the sum of all different G_α-interactions, given Z, at all levels (including zero) among these sub-sequences. [57] is most easily proved by induction on n. When $n = 1$, [57] reduces to [50], which establishes the base of the induction, while the induction step is

$$-G_\alpha(X_1 \cdots X_{n+1} \mid Z)$$

$$= -G_\alpha(X_1 \cdots X_n \mid X_{n+1} Z) - G_\alpha(X_{n+1} \mid Z)$$

$$= -G_\alpha(X_{n+1} \mid Z) + \sum_{r=1}^{n} \mathsf{S}^r_{X_1,\ldots,X_n} I_\alpha^{(r-1)}(\xi_1 \,;\, \cdots \,;\, \xi_r \mid X_{n+1} Z)$$

$$= I_\alpha^{(0)}(X_{n+1} \mid Z) + \sum_{r=1}^{n} \mathsf{S}^r_{X_1,\ldots,X_n}[I_\alpha^{(r-1)}(\xi_1 \,;\, \cdots \,;\, \xi_r \mid Z)$$

$$+ I_\alpha^{(r)}(\xi_1 \,;\, \cdots \,;\, \xi_r \,;\, X_{n+1} \mid Z)]$$

$$= \sum_{r=1}^{n+1} [\mathsf{S}^r_{X_1,\ldots,X_n} I_\alpha^{(r-1)}(\xi_1 \,;\, \cdots \,;\, \xi_r \mid Z)$$

$$+ \, \mathsf{S}^{r-1}_{X_1,\cdots,X_n} I^{(r-1)}_\alpha(\xi_1 \, ; \, \cdots \, ; \xi_{r-1} \, ; X_{n+1} \mid Z)]$$

$$= \sum_{r=1}^{n+1} \mathsf{S}^r_{X_1,\cdots,X_{n+1}} I^{(r-1)}_\alpha(\xi_1 \, ; \, \cdots \, ; \xi_r \mid Z) \qquad \text{QED.}$$

(The first line is authorized by [11]; the second line applies the induction hypothesis; the third follows from [50] and [51]; the fourth is a rearrangement of the third with the help of [55]; and the last follows by [54].)

A useful generalization of [57] which follows from it by way of [52] and [50] is

$$[58] \qquad I^{(k)}_\alpha(Y_1 \, ; \, \cdots \, ; \, Y_k \, ; X_1 \cdots X_n \mid Z) = \sum_{r=1}^{n} \mathsf{S}^r_{X_1,\cdots,X_n} I^{(k+r-1)}_\alpha(Y_1; \, \cdots \, ; \, Y_k \, ; \xi_1 \, ; \, \cdots \, ; \xi_r \mid Z).$$

This shows how an interaction at any level can be further expanded if its primary arguments are not all of unit length. An important special case of [58] is

$$[59] \qquad R_\alpha(Y; X_1 \cdots X_n \mid Z) = \sum_{r=1}^{n} \mathsf{S}^r_{X_1,\cdots,X_n} I^{(r)}_\alpha(Y; \xi_1 \, ; \, \cdots \, ; \xi_r \mid Z) = \sum_{i=1}^{n} R_\alpha(Y; X_i \mid Z) + \sum_{r=2}^{n} \mathsf{S}^r_{X_1,\cdots,X_n} I^{(r)}_\alpha(Y; \xi_1 \, ; \, \cdots \, ; \xi_r \mid Z)$$

(from [58] by [49]), which says that given Z, the G_α-contingency of q-set Y upon another q-set analyzed as a concatenation of n sub-sequences $X_1 \cdots X_n$ equals the sum of Y's G_α-contingencies with each X_i taken separately, plus all the G_α-interactions of all higher orders involving Y and the different r-selections from $X_1, \cdots, X_n$.

For the limiting case wherein $Z = \phi$ and each sub-sequence X_i is of unit length, [57] reduces to

$$[60] \qquad -G_\alpha(x_1 \cdots x_n) = \sum_{r=1}^{n} \mathsf{S}^r_{x_1,\cdots,x_n} I^{(r-1)}_\alpha(\xi_1 \, ; \, \cdots \, ; \xi_r)$$

which, together with the corresponding limiting case of [56], leads to an important conclusion about the relation between a generating function and the interactions derived from it. From a domain d containing n elements, a total of $2^n - 1$ different non-null distinct-element q-sets can be formed, each of which has a G_α-value which formally is undetermined by any of the others. Hence for simplicity disregarding arguments containing duplicated elements, the function G_α generates $2^n - 1$ independent items of data from the elements $x_1, \cdots, x_n$. The element-interaction function $G_{I(\alpha)}$ (see [53]) likewise generates this same number of distinct-element data over elements $x_1, \cdots, x_n$, while [60] and the appropriate specialization of [56] show that each of these $G_{I(\alpha)}$-data can be computed from the G_α-data *and conversely.* Hence a generating function is equivalent data-wise to its associated element-interaction function, and the latter may thus be thought of as a "rotation" (to use an obvious factor-analytic metaphor) of the former which preserves all the information in G_α with no loss in economy of expression even while structuring this information for maximal visibility of whatever relational patterning it may contain.

The hierarchy of interaction components also gives rise to symmetric partitions of configural savings. Since $\mathrm{S}^1_{x_1,\cdots,x_n} I^{(0)}_\alpha(\xi_1 \,|Z) = -\sum_{i=1}^n G_\alpha(x_i|\, Z)$, it follows by [6] from [57] that

[61] $$C_\alpha(x_1 \cdots x_n \mid Z) = \sum_{r=2}^n \mathrm{S}^r_{x_1,\cdots,x_n} I^{(r-1)}_\alpha(\xi_1 \,;\, \cdots \,;\, \xi_r \mid Z)\cdot$$

Thus the configural savings, given Z, in a q-set X with respect to G_α is the sum of all the G_α-interactions, given Z, of level 1 and higher among the elements in X. It is also of interest, perhaps, to note from [59] and [21] that

[62] $$C_\alpha(X \mid YZ) = C_\alpha(X \mid Z) + \sum_{r=2}^n \mathrm{S}^r_{x_1,\cdots,x_n} I^{(r)}_\alpha(Y; \xi_1 \,;\, \cdots \,;\, \xi_r \mid Z) \qquad (X = x_1 \cdots x_n)$$

or equivalently,

[63] $$C_\alpha(X \mid YZ) = C_\alpha(X \mid Z) + C_{\alpha,\iota Y}(X \mid Z).$$

Since each interaction $I^{(r)}_\alpha(Y; \xi_1 \,;\, \cdots \,;\, \xi_r \,|Z)$ in [62] can be further expanded into interactions among only elements from X and Y, [62] and [58] show that the configural savings $C_\alpha(X \mid Y)$ of q-set X after q-set Y is partialled out equals the unconditional configural savings of X, plus certain interactions of level 2 and higher among the elements of X and Y.

Finally, there is still another extension of abstract-partials theory which is very useful for its analysis-of-variance interpretation. While generating function G was originally stipulated to be a *number*-valued function, there has been nothing in the ensuing axiomatic development which requires this restriction. The entire formal system remains unaltered if G maps its arguments into any commutative group whose composition operator and identity element are designated by "+" and "0", respectively. In particular, an abstract-partials system can in this way have as its domain a set $\mathbf{F}$ of number-valued *functions* over another domain $\mathbf{a}$ such that each element (function) $x_i \in \mathbf{F}$ maps each element $a \in \mathbf{a}$ into a number x_{ia}, while G maps q-sets from $\mathbf{F}$ into additional functions over $\mathbf{a}$. (For example, the x_i might be experimental variables whose values have been observed for a set $\mathbf{a}$ of subjects while the values of G are certain composite variables defined from the x_i .) Let $G_{\mathbf{F}}$

be some symmetric functional (i.e., a symmetric function of functions) such that for each q-set $x_1 \cdots x_n$ of functions in $\mathbf{F}$, $G_{\mathbf{F}}(x_1 \cdots x_n)$ is the function over $\mathbf{a}$ which maps each element $a \varepsilon \mathbf{a}$ into the number $G_{\mathbf{F}}(x_{1a} \cdots x_{na})$, while $G_{\mathbf{F}}(\phi)$ is the constant function whose value for each element of $\mathbf{a}$ is zero. Now, number-valued functions of the same arguments can be added and subtracted in the very same way as numbers—e.g., the function $G_{\mathbf{F}}(x_1 \mid x_2) \overset{\text{def}}{=} G_{\mathbf{F}}(x_1x_2) - G_{\mathbf{F}}(x_2)$ over $\mathbf{a}$ is the function whose value for an element $a \varepsilon \mathbf{a}$ is $G_{\mathbf{F}}(x_{1a}x_{2a}) - G_{\mathbf{F}}(x_{2a})$—while for all functions $x_i \varepsilon \mathbf{F}$, the function $G_{\mathbf{F}}(x_1 \cdots x_n) + G_{\mathbf{F}}(\phi)$ is identical with function $G_{\mathbf{F}}(x_1 \cdots x_n)$. Hence all definitions and theorems about scalar quantities derived from a number-valued generating function G over a domain $\mathbf{d}$ go over upon substitution of $G_{\mathbf{F}}$ for G and $\mathbf{F}$ for $\mathbf{d}$ into definitions and theorems about functions derived from a function-valued generating functional $G_{\mathbf{F}}$ over a domain $\mathbf{F}$ of common-argument functions. In particular, if the fact that $G_{\mathbf{F}}$'s domain consists of functions is made notationally explicit by writing each element $x_i \varepsilon \mathbf{F}$ as "$x_{i\alpha}$", in which α is an argument place-holder, partition [57] becomes

$$[64] \qquad -G_{\mathbf{F}}(X_{1\alpha} \cdots X_{n\alpha} \mid Z_\alpha) = \sum_{r=1}^{n} \mathrm{S}^r_{X_1,\ldots,X_n} I_{\mathbf{F}}^{(r-1)}(\xi_{1\alpha} ; \cdots ; \xi_{r\alpha} \mid Z_\alpha),$$

in which each $X_{i\alpha}$ (and similarly for Z_α) is some q-set $x_{h\alpha} \cdots x_{j\alpha}$ of functions in $\mathbf{F}$. The definition and combinatorial behavior of the interaction functions $I_{\mathbf{F}}^{(r-1)}(\xi_{1\alpha} ; \cdots ; \xi_{n\alpha} \mid Z_\alpha)$ follow by substitution of $X_{i\alpha}$ for each X_i in [45] *et seq.*

3. *Interpretations*

The entry problem.

To give the formal system of abstract partials a substantive embodiment, it is necessary to identify a set of measures which stand in relations isomorphic to the equations developed above. Since all the abstract functions $G(X \mid Z)$, $C(X \mid Z)$, $R(X; Y \mid Z)$ and their ramifications were introduced above by explicit definitions built upon G, finding a substantive measure G_s which satisfies the axiomatic properties of G suffices to establish an interpretation for the entire system (possibly excluding the fusion equations), since measures $G_s(X \mid Z)$, $C_s(X \mid Z)$, $R(X; Y \mid Z)$ etc. can then be defined from G_s to parallel the uninterpreted system erected upon G. But it is not at all *necessary* that an interpreted abstract-partials system have this particular definitional structure—all that matters mathematically is that the system satisfy equations [1]–[3], [5]–[8], [45], and [46] (from which, excepting the fusion equations, all else follows), irrespective of why this is true. Thus it may well be that one or more of the interpreted measures $G_s(X \mid Z)$, $C_s(X \mid Z)$, $R_s(X; Y \mid Z)$, etc. derived from substantive generating function G_s in accord with the definitive abstract structure is actually a quantity whose most fundamental definition is independent of this quantity's abstract-partials relation to G_s even though it has the latter as a mathematical consequence. For example, all serious interpretations examined below begin with a substantive statistic $Q(X, Y)$ on ordered pairs of q-sets from a certain domain and of which $Q(X, \phi)$ is conceptually no more than a limiting case, but which has the important property that if generating function G_Q is introduced as $G_Q(X) \overset{\text{def}}{=} Q(X, \phi)$, it follows as a theorem that $G_Q(X \mid Z) = Q(X, Z)$. Any abstract-partials quantity for which a given interpretation of the system provides a substantive definition which is *not* merely a construction out of other abstract partials already given substance may be thought of as a *system-entry* position for that interpretation. The more abundantly a model of the abstract-partials system makes contact with its substantive substratum in this way, the more "significant" or "meaningful" (in a sense as intuitively important as it is difficult to define) the interpretation is. In fact, as will be seen, the various abstract-partial quantities in a particular model of the system are not, in general, equally meaningful, and their degree of significance appears to be strongly determined by their nearness to a system-entry position. To understand different substantive instances of the system in depth, therefore, it is necessary to appreciate not merely how one is isomorphic to another via the abstract structure but also how they differ in respect to system-entry properties.

The Information-theoretical interpretation.

The reader who is acquainted with Uncertainty analysis (Information theory) has doubtlessly already recognized the isomorphism between the uninterpreted system of abstract partials developed above and the structure of partial Uncertainties previously articulated by McGill [1954], Watanabé [1960], Garner [1962] and others. The fundamental measure of Uncertainty analysis, first introduced by Shannon [1948] under the name "Information," is defined as follows: Let $x_1, x_2, \cdots, y_1, y_2, \cdots$, etc. be scientific variables ("variates") each of which has a finite number of values and which have a joint probability distribution in some background population P. Let "$x_{i(j)}$" designate the jth value of variable x_i and similarly for $y_{h(k)}$ etc. Then the (unconditional) Uncertainty over the joint distribution of variables $y_1, \cdots, y_n$ in P is defined

$$[65] \qquad \mathrm{U}(y_1, \cdots, y_n) \overset{\text{def}}{=} -\sum^{y} \Pr[y_{1(i)} \cdots y_{n(j)}] \log \Pr[y_{1(i)} \cdots y_{n(j)}],$$

where $\Pr[y_{1(i)} \cdots y_{n(j)}]$ is the probability in P of a particular combination $y_{1(i)} \cdots y_{n(j)}$ of values on variables $y_1, \cdots, y_n$, respectively, summation is over all such combinations, and the base of the logarithm is an arbitrary parameter. More generally, the joint Uncertainty over $y_1, \cdots, y_n$ given a particular combination $x_{1(h)} \cdots x_{m(k)}$ of values on predictor variables $x_1, \cdots, x_m$, respectively, is

$$[66] \quad U_{x_{1(h)}\cdots x_{m(k)}}(y_1, \cdots, y_n) \overset{\text{def}}{=} -\sum^{y} \Pr[y_{1(i)} \cdots y_{n(j)} \mid x_{1(h)} \cdots x_{m(k)}] \cdot \log \Pr[y_{1(i)} \cdots y_{n(j)} \mid x_{1(h)} \cdots x_{m(k)}],$$

where $\Pr[y_{1(i)} \cdots y_{n(j)} \mid x_{1(h)} \cdots x_{m(k)}]$ is the probability that a random member of population P has configuration $y_{1(i)} \cdots y_{n(j)}$ of y-values when his configuration of x-values is $x_{1(h)} \cdots x_{m(k)}$. Finally, the conditional joint Uncertainty (in P) over variables $y_1, \cdots, y_n$ given *variables* $x_1, \cdots, x_m$ is the Uncertainty statistically expected to remain for the joint scores on $y_1, \cdots, y_n$ of a random member of P after his scores on variables $x_1, \cdots, x_m$ are given, i.e.,

$$[67] \quad U_{x_1\cdots x_m}(y_1, \cdots, y_n) \overset{\text{def}}{=} \sum^{x} \Pr[x_{1(h)} \cdots x_{m(k)}] U_{x_{1(h)}\cdots x_{m(k)}}(y_1, \cdots, y_n),$$

where summation is over all combinations of values on $x_1, \cdots, x_m$. If the set of variables $x_1, \cdots, x_m$ is considered also to include an additional null-variable x_0 whose value is the same for all members of P (a convention which does not affect the right-hand sides of [66] and [67]), [65] becomes the special case of [67] in which $m = 0$. A simple but vital consequence of [66] and [67] is that

$$[68] \quad U(x_1, \cdots, x_m, y_1, \cdots, y_n) = U(x_1, \cdots, x_m) + U_{x_1\cdots x_m}(y_1, \cdots, y_n),$$

or, more generally,

$$[69] \quad U_Z(XY) = U_Z(X) + U_{XZ}(Y),$$

where X, Y, and Z are ordered sets of categorical variables $x_1, \cdots, x_m$; $y_1, \cdots, y_n$; and $z_1, \cdots, z_p$, respectively, jointly distributed in P. Uncertainty over a set of *no* variables is left undefined by [67], so we are free to adopt

$$[70] \quad U_Z(\phi) \overset{\text{def}}{=} 0,$$

which allows [69] to hold without restrictions on the number of variables in X, Y, and Z. Alternatively, if ϕ is construed to be constant-variable x_0, [70] is a consequence of [65]–[68].

The argument of unconditional Uncertainty measure U is any ordered set of zero or more of the categorical (i.e., finite-valued) variables jointly distributed in population P, the value of U for a given argument $x_1, \cdots, x_n$ is invariant under all permutations of the x_i, and the value of U for the null argument is zero; hence U satisfies the formal requirements to be the generating function for an abstract-partials system whose domain, $\mathbf{d}_{cv}$, is some set of these variables. Accordingly, let

$$\text{[71]} \qquad \mathrm{G}_{\mathrm{U}}(X) \overset{\text{def}}{=} \mathrm{U}(X)$$

for any q-set X of variables from $\mathbf{d}_{cv}$, and let $\mathrm{G}_{\mathrm{U}}(X \mid Z)$ etc. be defined for q-sets from $\mathbf{d}_{cv}$ in accord with the definitions of their uninterpreted counterparts. In view of [69] we then have

$$\text{[72]} \qquad \mathrm{G}_{\mathrm{U}}(X \mid Z) = \mathrm{U}_Z(X)$$

$$\text{[73]} \qquad \mathrm{C}_{\mathrm{U}}(X \mid Z) = \sum_{i=1}^{n} \mathrm{U}_Z(x_i) - \mathrm{U}_Z(X) \qquad (X = x_1 \cdots x_n)$$

$$\text{[74]} \qquad \mathrm{R}_{\mathrm{U}}(Y; X \mid Z) = \mathrm{U}_Z(Y) - \mathrm{U}_{XZ}(Y)$$

which, together with the G_{U}-interactions, are familiar concepts in Uncertainty analysis. In particular, $\mathrm{R}_{\mathrm{U}}(y; X)$—i.e., $\mathrm{U}(y) - \mathrm{U}_X(y)$—is readily appreciated to be the Information-theoretical analog of a criterion variable's multiple correlation with a set of predictor variables (a parallel which will shortly be seen to be much more than *mere* analogy), since it describes the average amount of Uncertainty in criterion y which is eliminated through knowledge of data on predictors X. The quantity $\mathrm{C}_{\mathrm{U}}(X)$, known as the "total constraint" [Garner, 1962] or "total correlation" [Watanabé, 1960] in the joint distribution of variables X, is a symmetric measure of total relatedness within an n-tuple of categorical variables and as such is a challenging new concept for multivariate analysis to play with, especially in light of its provocative partitions. Finally, the Information-theoretical "fusion" of variables $x_1, \cdots, x_n$ is their cartesian product—i.e., $^{|}\overline{x_1 \cdots x_n}^{|}$ is the single categorical variable whose values are the various combinations of values jointly possible on $x_1, \cdots, x_n$—which obviously satisfies $\mathrm{G}_{\mathrm{U}}(^{|}\overline{X}^{|}) = \mathrm{G}_{\mathrm{U}}(X)$.

When parallels are sought between Information theory and other systems of multivariate analysis, those theorems about U which command the greatest interest are the ones describing the partition of a criterion variable's total Uncertainty into components attributable to various sources. Primary among these is

$$\begin{aligned}
\text{[75]} \qquad \mathrm{U}(y) &= \mathrm{U}_X(y) + \mathrm{R}_{\mathrm{U}}(y; X) \\
&= \mathrm{U}_X(y) + \sum_{i=1}^{n} \mathrm{R}_{\mathrm{U}}(y; x_i) + \sum_{r=2}^{n} \mathrm{S}^{r}_{x_1,\cdots,x_n} \mathrm{I}^{(r)}_{\mathrm{U}}(y; \xi_1; \cdots; \xi_r) \\
&= \mathrm{U}_X(y) + \sum_{r=1}^{n} \mathrm{S}^{r}_{x_1,\cdots,x_n} \mathrm{I}^{(r)}_{\mathrm{U}}(y; \xi_1; \cdots; \xi_r) \\
&= \mathrm{U}_X(y) + \sum_{r=1}^{n} \mathrm{S}^{r}_{x_1,\cdots,x_n} \mathrm{I}^{(r-1)}_{\mathrm{U}.iy}(\xi_1; \cdots; \xi_r) \qquad (X = x_1 \cdots x_n)
\end{aligned}$$

(from [26], [59], and [52]), in which $U_X(y)$ is the residual Uncertainty in y unaccounted for by predictor variables X while the Uncertainty $R_U(y; X)$ of y jointly accounted for by predictors X is further analyzed into single-predictor contributions plus higher-order interaction components. This is the Uncertainty partition introduced by McGill in 1954 and later [Garner and McGill, 1956] shown to be isomorphic to the familiar Fisherian partition of a metrical criterion's variance (cf. [84], below).

It is instructive at this point to ask what, specifically, is contributed to Uncertainty analysis by Shannon's definition of Information (i.e., Uncertainty) as given by [65], in contrast to other available statistics which likewise assess categorical dispersion. For example, a measure $U^*(x)$ of categorical unpredictability which makes a good deal of intuitive sense is the number of errors statistically expected when guessing variable x's value by a sequential procedure in which the value guessed first is the one whose probability is highest, the one tried second (if the first guess is wrong) has the second-highest probability, etc. This measure is clearly symmetric in its arguments when extended to the joint unpredictability $U^*(X)$ of a set of variables, and can be taken to define an Unpredictability model of the abstract-partials system, including Unpredictability components which precisely parallel the Uncertainty components based on definition [65]. Why, then, should U be preferred to U*? The answer lies in [67] and [68]. Conditional Uncertainties are not defined *from* unconditional Uncertainties but in parallel to them. That is, the meaning of $U_Z(X)$ is not derivative from that of $U(X)$ but stands on an equal conceptual footing with it. It is consequently a happy accident, so to speak (though Shannon carefully contrived for this "accident" to occur), that the measure $G_U(X \mid Z)$ defined from $G_U(X)$ in accord with [3] happens to coincide with $U_Z(X)$. In contrast, while $G_{U^*}(X \mid Z)$ is similarly derived from Unpredictability measure $G_{U^*}(X) = U^*(X)$, $G_{U^*}(X \mid Z)$ does *not* stand in any fixed relation to G_{U^*}-values for the conditional distributions of X, given specific values of Z, but depends entirely upon generating function G_{U^*} by way of definition-form [3] for its meaning. The potency of Shannon's measure thus lies not in there being anything special about $U(X)$ as a generating function—$U^*(X)$ and many others are equally qualified—but in $U(X)$'s being a special case of a more comprehensive substantive statistic $U_Z(X)$ in virtue of which the Information-theoretical interpretation achieves simultaneous entry to the system across the entire sheet of partial G_U-values of all orders, rather than at just the zero-order baseline as would be true of an interpretation based on U*.

It should be added that the definition of U also confers upon it another useful property, namely,

[76] $$U_Z(x_1, \cdots, x_n) \leq \sum_{i=1}^{n} U_Z(x_i),$$

with equality holding when the x_i are all statistically independent of one another. Consequently, $\mathrm{C_U}(X \mid Z)$ and $\mathrm{R_U}(Y; X \mid Z)$ are always non-negative, and signify an absence of relationship by a value of zero. This convenience does not extend to interactions higher than the first level, however, nor does it provide any additional system-entry positions. As it is, Uncertainty analysis appears to have entry to the abstract-partials system *only* through its array of partial Uncertainties, with the result that the higher-order interaction terms in [75] have at present only dubious significance for empirical research.

The analysis-of-variance interpretation.

In a Fisherian analysis-of-variance design, we are given (*a*) a metrical criterion variable y, (*b*) a set $\mathbf{d}_v$ of predictor variables, the number of which is unbounded in principle though remarkably finite in practice, and (*c*) a population $\mathbf{a}$ of subjects within which y and the variables in $\mathbf{d}_v$ have a joint probability distribution. It is usually assumed also that (*d*) each predictor variable has only a finite number of values and that (*e*) the predictors are all fully independent of one another. However, (*d*) is superfluous here, while we shall abstain from (*e*) until we are in position to see precisely what this crucial condition achieves.

Fisherian analysis of variance partitions both the criterion variable and the criterion's variance. The distinction between these two partitions is fundamental, for abstract-partials theory makes clear that each can be developed independently of the other. To derive the first partition, let X be any q-set of zero or more predictor variables in $\mathbf{d}_v$ and define $\mu_y(X)$ to be the variable whose value for each member a of population $\mathbf{a}$ is the statistically expected value of criterion variable y among members of $\mathbf{a}$ whose scores on predictors X are the same as a's. That is, $\mu_y(X)$ is the multiple curvilinear regression of y upon predictors X in $\mathbf{a}$. For $X = \phi$, $\mu_y(\phi)$ is the constant variable whose value for each $a \varepsilon \mathbf{a}$ is the unconditional expectation (i.e., grand mean) μ_y of y in $\mathbf{a}$. Now write

[77] $$e_y \overset{\mathrm{def}}{=} y - \mu_y(X)$$

for the component of y unaccounted for by y's curvilinear regression upon predictors X and let

[78] $$\mathrm{G}_{\mu_y}(X) \overset{\mathrm{def}}{=} \mu_y - \mu_y(X),$$

whence by [77],

[79] $$\begin{aligned} y - \mu_y &= e_y(X) + \mu_y(X) - \mu_y \\ &= e_y(X) - \mathrm{G}_{\mu_y}(X). \end{aligned}$$

Since the regression of y in $\mathbf{a}$ upon variable $\mu_y(\phi)$ has the constant value μ_y ,

$G_{\mu_y}(\phi)$ has the value zero for all members of **a**. Thus while G_{μ_y} maps q-sets of variables in $\mathbf{d}_y$ into number-valued functions over **a**, it still satisfies the conditions to be a generating function over $\mathbf{d}_y$ in the extended sense described above. Accordingly, to each ordered set $X_1, \cdots, X_n, Z$ of q-sets of variables in $\mathbf{d}_y$, there corresponds an $(n-1)$th level conditional interaction variable $I^{(n-1)}_{\mu_y}(X_1 ; \cdots ; X_n \mid Z)$. By [64], these determine symmetric partitions of $G_{\mu_y}(X \mid Z)$ for any q-set X of $\mathbf{d}_y$-predictors. In particular, when $Z = \phi$,

$$[80] \qquad -G_{\mu_y}(X_1 \cdots X_n) = \sum_{r=1}^{n} S^r_{X_1,\ldots,X_n} I^{(r-1)}_{\mu_y}(\xi_1 ; \cdots ; \xi_r),$$

which by [79] thus partitions y as

$$[81] \qquad y - \mu_y = e_y(X_1 \cdots X_n) + \sum_{r=1}^{n} S^r_{X_1,\ldots,X_n} I^{(r-1)}_{\mu_y}(\xi_1 ; \cdots ; \xi_r).$$

In the special but most familiar case of [81] wherein each $X_i = x_i$ (i.e., the X_i are of unit length) and all predictors are fully independent of one another, the components $I^{r-1}_{\mu_y}(\xi_1 ; \cdots ; \xi_r)$ of zero order (i.e., for which $r = 1$) are called the "main effects" of the various predictors $x_1, \cdots, x_n$ upon y, while the rest are known as "interaction effects" of various orders from 1 to n.

Equation [81] partitions criterion y into component *variables*. To partition y's *variance* in this same fashion, we introduce a second generating function over $\mathbf{d}_y$:

$$[82] \qquad \begin{aligned} G_{V_y}(X) &\overset{\text{def}}{=} -\text{Var}\,[\mu_y(X)] \\ &= -\text{Var}\,[G_{\mu_y}(X)]. \end{aligned}$$

That is, for any q-set X of predictors in $\mathbf{d}_y$, $-G_{V_y}(X)$ is the variance of y (in **a**) accounted for by y's curvilinear regression upon predictors X. Since $\mu_y(X)$—and hence $G_{\mu_y}(X)$—is orthogonal to $e_y(X)$ (a basic theorem of regression theory), [79] and [82] entail

$$[83] \qquad \begin{aligned} \text{Var}\,(y) &= \text{Var}\,[e_y(X)] + \text{Var}\,[\mu_y(X)] \\ &= \text{Var}\,[e_y(X)] - G_{V_y}(X), \end{aligned}$$

whence by abstract partition [57],

$$[84] \qquad \text{Var}\,(y) = \text{Var}\,[e_y(X_1 \cdots X_n)] + \sum_{r=1}^{n} S^r_{X_1,\ldots,X_n} I^{(r-1)}_{V_y}(\xi_1 ; \cdots ; \xi_r).$$

The formal parallel between criterion partition [81] and criterion-variance partition [84] is obvious. Moreover, it is not difficult to show (though we shall not do so here) that if each variable in q-set $X_1 \cdots X_n$ is fully independent of the rest, as always contrived in analysis of variance either by experimental control over allocation of subjects to the various treatment cells or hypothetically by analyzing the configuration of cell means as though this were so, then the component variables in [81] are all orthogonal to (though not in general independent of) one another and hence

$$\text{[85]} \qquad \mathrm{Var}\,[\mathrm{I}_{\mu_y}^{(r-1)}(\xi_1\,;\,\cdots\,;\,\xi_r)] = \mathrm{I}_{\mathrm{V}_y}^{(r-1)}(\xi_1\,;\,\cdots\,;\,\xi_r)$$

(full predictor independence)

for each r-selection $\xi_1\,,\,\cdots\,,\,\xi_r$ from $X_1\,,\,\cdots\,,\,X_n$. That is, under the standard (but artificial) analysis-of-variance stipulation of predictor independence, each variance component in [84] is the variance of the corresponding criterion component in [81].

Variance partition [84] is not only formally isomorphic to the last line of Uncertainty partition [75] but, insomuch as variance can also be construed as a measure of uncertainty, is remarkably close to it conceptually as well. Even so, the abstract-partials isomorphism between analysis-of-variance and Uncertainty analysis contains an important asymmetry. In a system of abstract partials based on generating function G_α , the primary measure of relationship is $R_\alpha(y;\,X)$, i.e., the G_α-contingency of element y upon the elements in q-set X. In the Information-theoretical interpretation, $\mathrm{R_U}(y;\,X)$ is the amount of Uncertainty eliminated (on the average) by knowledge of X-scores, and for an intuitively meaningful isomorphism the analysis-of-variance counterpart of this measure should be the amount of criterion variance jointly accounted for by the predictors, namely, Var $[\mu_y(X)]$. But to put Var $[\mu_y(X)]$ into correspondence with $\mathrm{R_U}(y;\,X)$, as done when partitions [75] and [84] are considered to be isomorphic, $\mathrm{G}_{\mathrm{V}_y}$ must be coordinated not with $\mathrm{G_U}$ but with $\mathrm{G_U}$'s linear development $\mathrm{G}_{\mathrm{U}.\iota y}$. Admittedly, this is a perfectly good isomorphism, but the counterpart of $\mathrm{I}_{\mathrm{V}_y}^{(k-1)}(x_1\,;\,\cdots\,;\,x_k)$ is then $\mathrm{I}_{\mathrm{U}.\iota y}^{(k-1)}\cdot(x_1\,;\,\cdots\,;\,x_k) = \mathrm{I}_{\mathrm{U}}^{(k)}(y;\,x_1\,;\,\cdots\,;\,x_k)$ and there is no analysis-of-variance counterpart of $\mathrm{G_U}$ at all. In particular, whereas both $\mathrm{U}(y)$ (i.e., $\mathrm{G_U}(y)$) and $\mathrm{U}_X(y)$ (i.e., $\mathrm{G_U}(y \mid X)$) in [75] belong to the same interpreted abstract-partials system as the other components in [75], there exists no expression in the $\mathrm{G}_{\mathrm{V}_y}$-system for the criterion's total variance Var (y) and residual variance Var $[e_y(X)]$ in [84]. On the other hand, if $\mathrm{G_U}$ is isomorphically coordinated with $\mathrm{G}_{\mathrm{V}_y}$, $\mathrm{R_U}(y;\,x)$ becomes coordinated not with main-effect variance Var $[\mu_y(x)]$ but with a first-level interaction variance of form $\mathrm{R}_{\mathrm{V}_y}(x_i\,;\,x_j)$, while more generally the analysis-of-variance counterpart of $\mathrm{R_U}(y;\,X)$ becomes a sum of interaction terms which do not include any main-effect variance.

A further disfiguring complication to the formal similarity between Uncertainty analysis and analysis of variance is injected by the duality of partitioning in the latter. The fact that each criterion-variance component in [84] is normally the variance of a corresponding component of the criterion gives the analysis-of-variance interaction components an extra significance *not* shared by the interaction components in an Uncertainty partition. That is, given predictor independence, the variance-partition sector of analysis of variance, unlike Uncertainty analysis, has entry to the abstract-partials system through all interaction terms. Thus to think uncritically of the Uncertainty components in [75] as analogous to the variance of main effect and interaction components in an analysis-of-variance design promotes a false sense of understanding by tempting us to think that Uncertainty components are the Uncertainties of criterion components. On the other hand, it is worth stressing that while variance partition [84] always holds whether the predictor variables are fully independent or not, so that the isomorphism between [75] and [84] does not require any distributional assumptions, if the analysis-of-variance predictors are *not* fully independent then the variance components $\mathrm{I}_{V_v}^{r-1}(\xi_1 ; \cdots ; \xi_r)$ in [84] for which $r > 1$ are no longer the variances *of* anything and do not stand on the same conceptual footing as an honest-to-god variance. In this more general case, analysis of variance has as much to learn about the meaning of its variance components from their isomorphism to Uncertainty components as there is to be learned about the latter by sighting along the isomorphism in the other direction.

The conditional-probability interpretation.

This time, let x_1, x_2, etc., be various *attributes** which may or may not be possessed by a member of background population P, and let the quantity

$$-\log \Pr (x_1 \cdots x_n \mid z_1 \cdots z_m)$$

be called the (conditional) *Implausibility* of attribute-combination $x_1 \cdots x_n$ (duplications not excluded) given attribute-combination $z_1 \cdots z_m$, where $\Pr (x_1 \cdots x_n \mid z_1 \cdots z_m)$ is the conditional joint probability of $x_1 \cdots x_n$ given $z_1 \cdots z_m$ in P and the logarithm's base is an arbitrary parameter. The unconditional Implausibility of attribute-cluster $x_1 \cdots x_n$ in P is of course $-\log \Pr (x_1 \cdots x_n)$. The Implausibility of joint attributes $x_1 \cdots x_n$ given attributes $z_1 \cdots z_m$ has a lower bound of zero, attained when $z_1, \cdots, z_m$ jointly imply cluster $x_1 \cdots x_n$ with certainty, and increases without limit as the probability of cluster $x_1 \cdots x_n$ given $z_1 \cdots z_m$ approaches zero.

Now let $\mathbf{d}_A$ be some set of attributes which can meaningfully be ascribed to members of population P. Since the unconditional Implausibility (in P) of any cluster $X = x_1 \cdots x_n$ of attributes in $\mathbf{d}_A$ is symmetric in the x_i and may appropriately be stipulated to have a value of 0 when $n = 0$ (this being alternatively available as a theorem if ϕ, i.e., x_0, is taken to be a universal attribute possessed by every member of P), Implausibility qualifies as a generating function over $\mathbf{d}_A$. Hence setting

$$\text{[86]} \qquad \mathrm{G}_{cp}(X) \overset{\text{def}}{=} -\log \Pr (X)$$

*For clarification of the difference between attributes (i.e., properties) and scientific variables (i.e., "variates"), see Rozeboom [1966b].

invests the abstract-partials system with an interpretation which analyzes the dependency structure of attributes in P with perfect isomorphism to the pattern of analysis applied to categorical variables by Information theory. Moreover, while any other transformation of Pr (X) followed by an appropriate zero-adjustment likewise defines a generating function over $\mathbf{d}_A$, Implausibility has the special virtue that, like Uncertainty, it achieves simultaneous entry to the abstract-partials system at conditional G-values of all orders. Specifically, since $\Pr(X \mid Z) = \Pr(XZ)/\Pr(Z)$, it follows from [86] that

$$\text{[87]} \qquad \mathrm{G}_{cp}(X \mid Z) = -\log \Pr(X \mid Z),$$

so $\mathrm{G}_{cp}(X \mid Z)$ is the Implausibility of attribute-cluster X given attributes Z. Configural savings and G-contingencies are readily intuited as measures of relationship in the conditional-probability model, for

$$\text{[88]} \qquad \mathrm{R}_{cp}(Y; X) = \log \left[\frac{\Pr(Y \mid X)}{\Pr(Y)} \right]$$

$$\text{[89]} \qquad \mathrm{C}_{cp}(x_1 \cdots x_n) = \log \left[\frac{\Pr(x_1 \cdots x_n)}{\prod_{i=1}^{n} \Pr(x_i)} \right],$$

and similarly for conditional values of R_{cp} and C_{cp}. $\mathrm{R}_{cp}(Y; X)$ compares the probability of attribute-cluster Y given attributes X to the unconditional probability of Y and hence assesses how Y's likelihood is affected by X, while $\mathrm{C}_{cp}(X)$ compares the joint probability of the attributes in cluster X to the probability this cluster would have if its constituents were independent of one another while retaining their present marginal probabilities.

It is evident that the conditional-probability model of the abstract-partials system also provides an interpretation for the fusion equations, namely, when $'\overline{x_1 \cdots x_n}'$ is taken to be the conjunction of attributes $x_1, \cdots, x_n$.

Insomuch as a categorical variable's Uncertainty is a weighted average of the Implausibilities of the alternative attributes which compose its values, the relation between Information theory and the analysis of conditional probabilities is actually more intimate than just an abstract-partials isomorphism. In view of this close substantive overlap between the two systems, it should be worth inquiry whether there may not be applications of Information theory which would be served as well or better by a formally equivalent analysis of the Implausibilities of attribute clusters. Implausibility theory (if we may so call the pattern of abstract-partials analysis based on G_{cp}) is to an extent handicapped by the fact that C_{cp} and R_{cp}, in contrast to C_U and R_U, can assume negative values as well as positive ones, though how serious a disadvantage this may be remains to be seen. On the other hand, whereas what

Uncertainty analysis reveals about the relations among variables includes nothing about *how* they are related—i.e., how one varies as a function of another—, conditional Implausibilities make explicit what probabilistic conclusions can be drawn about criterion attributes given the predictor evidence. It is far from impossible that Uncertainty analysis will turn out to be the more powerful tool for detecting gross patterns of relationship in categorical data while Implausibility theory is then the precision instrument with which these relations are best analyzed in detail.

Incidently, there is another generating function over domain $\mathbf{d}_A$ which also has mathematically interesting and conceivably useful combinatorial properties. Let G_{d_p} (the subscript signifies "disjunctive probability") be defined

[90] $$G_{d_p}(X) \overset{\text{def}}{=} \Pr({}^{|}\overline{X}{}^{|}),$$

where ${}^{|}\overline{X}{}^{|}$, the fusion of q-set X, is defined in this model to be the *disjunction* of the attributes in X—i.e., ${}^{|}\overline{x_1 \cdots x_n}{}^{|} \overset{\text{def}}{=}$ the attribute of possessing at least one of the attributes x_1 or $\cdots$ or x_n. It then follows that

[91] $$\mathrm{I}_{d_p}^{(n-1)}(x_1 ; \cdots ; x_n \mid z_1 \cdots z_m)$$
$$= (-1)^n \Pr(x_1 \,\&\, x_2 \,\&\, \cdots \,\&\, x_n \,\&\, {\sim}z_1 \,\&\, {\sim}z_2 \,\&\, \cdots \,\&\, {\sim}z_m),$$

where & and $\sim$ are the logical connectives *and* and *not*, respectively. Hence for any class definable by a Boolian algebra on a set of attributes $x_1, \cdots, x_n$, there exists a sum of terms of form $(-1)^{k+1}\mathrm{I}_{d_p}^{(k)}(x_0 ; x_1' ; \cdots ; x_k' \mid x_{k+1}' \cdots x_n')$ $(0 \leq k \leq n)$, where $x', \cdots, x_1'$ are some permutation of $x_1, \cdots, x_n$, which equals the probability of that class. (The reason for including universal attribute x_0, for which $\Pr(x_0) = 1$, is to allow the possibility $k = 0$.) If any of these sums vanish, a pattern of incompatibility or entailment among the x_i stands thereby revealed. In this connection, it is also of interest that $C_{d_p}(x_1 \cdots x_n)$ has a lower bound of zero, attained when and only when attributes $x_1, \cdots, x_n$ are mutually exclusive.

The correlational interpretation.

In light of the close affinities between Uncertainty components and conditional-probability measures, their isomorphism is not particularly startling. That the system of linear correlation coefficients should also share this common structure is much less evident, and that the theory of abstract partials reveals this to be so is a major testimonial to its systematizing potential. To be sure, previous writers [e.g., Attneave, 1959, Appendix I; Ross, 1962; Fhanér, 1966] have noted the appearance of linear correlations in more-or-less literal applications of Information theory to the joint distribution of metrical variables. However, the relation between Information theory and correlational analysis so disclosed (*a*) employs not strictly the categorical

Uncertainty measure, the value of which is infinite for any continuous distribution, but a modification [Shannon, 1948, Section 20; Ross, 1962] applicable to metrical variables, and (*b*) requires the latter's joint distribution to be normal; hence the equivalence established between the two systems by the Shannon–Ross metrical Uncertainty measure is at best a precarious one. In contrast, the to-be-described abstract-partials *isomorphism* between Information theory and linear correlational analysis requires no distributional assumptions or measure modifications whatsoever.

Once again let lower-case letters x, y, etc., denote variables—this time metrical variables—which are jointly distributed in background population P, while upper-case letters X, etc., denote ordered sets of these variables. Then we may write $\mathrm{R}_{y(X)}$ for the multiple linear correlation of criterion y with predictor variables X, $\mathrm{r}_{xy\cdot Z}$ ("partial" correlation) for the correlation between the residuals of x and y after their linear regressions upon variables Z have been extracted, and $\mathrm{R}_{y(X)\cdot Z}$ ("multiple-partial" correlation) for the multiple correlation of the residual of y upon the residuals of variables X after variables Z have been partialled out. (r_{xy} is, of course, the zero-order correlation between x and y.) It will be recalled that all these correlational statistics (or more precisely their magnitudes) are functions of residual standard deviations of form $\sigma_{y\cdot X}$, i.e., the standard deviation of the component of y which remains after y's linear regression upon predictors X has been extracted.* (In the special case where X is the null set ϕ, $\sigma_{y\cdot X} = \sigma_y$—i.e., extraction of y's regression upon no predictors does not reduce its variance.) Also, to each correlation coefficient there corresponds a coefficient of alienation, namely,

[92] $$\mathrm{k}_{xy} \overset{\text{def}}{=} \sqrt{1 - \mathrm{r}_{xy}^2} = \frac{\sigma_{y\cdot x}}{\sigma_y} = \frac{\sigma_{x\cdot y}}{\sigma_x}$$

[93] $$\mathrm{K}_{y(X)} \overset{\text{def}}{=} \sqrt{1 - \mathrm{R}_{y(X)}^2} = \frac{\sigma_{y\cdot X}}{\sigma_y}$$

[94] $$\mathrm{k}_{xy\cdot Z} \overset{\text{def}}{=} \sqrt{1 - \mathrm{r}_{xy\cdot Z}^2} = \frac{\sigma_{y\cdot xZ}}{\sigma_{y\cdot Z}} = \frac{\sigma_{x\cdot yZ}}{\sigma_{x\cdot Z}}$$

and most generally of all

[95] $$\mathrm{K}_{y(X)\cdot Z} \overset{\text{def}}{=} \sqrt{1 - \mathrm{R}_{y(X)\cdot Z}^2} = \frac{\sigma_{y\cdot XZ}}{\sigma_{y\cdot Z}}.$$

$\mathrm{K}_{y(X)\cdot Z}$ is a monotonic decreasing function of $\mathrm{R}_{y(X)\cdot Z}$ and while alienation is perhaps a less familiar statistic than is correlation, K is more directly meaningful as a measure of predictive control than is R in that if a variable's standard

*See, e.g., DuBois [1957] or Rozeboom [1966a] for detailed development of linear correlation theory.

deviation is construed as our "metrical uncertainty" (not Uncertainty) about its value for a random member of P, the ratio $\sigma_{y \cdot XZ}/\sigma_{y \cdot Z}$ states directly how much uncertainty about y remains after scores on predictors XZ are available, compared to the uncertainty about y given only data on predictors Z.

Now let the statistic Π_X for the joint distribution of metrical variables X be defined

[96] $$\Pi_X \overset{\text{def}}{=} \prod_{i=1}^{n} \sigma_{x_i \cdot x_1 \cdots x_{i-1}} = \left[\prod_{i=1}^{n} \mathrm{K}_{x_i(x_1 \cdots x_{i-1})}\right]\left[\prod_{i=1}^{n} \sigma_{x_i}\right] \qquad (X = x_1 \cdots x_n)$$

for $n \geq 1$, while if X is null,

[97] $$\Pi_\phi \overset{\text{def}}{=} 1.$$

To give quantity Π_X a name, we may as well call it the "Pi-value" of multivariate configuration X, while the quantity

[98] $$\pi_X \overset{\text{def}}{=} \frac{\Pi_X}{\prod_{i=1}^{n} \sigma_{x_i}} = \prod_{i=1}^{n} \mathrm{K}_{x_i(x_1 \cdots x_{i-1})}, \qquad (X = x_1 \cdots x_n)$$

which is what Π_X becomes when all the variables in X are standardized to unit variance, is the "Pi-coefficient" of configuration X. Although Π_X appears in [96] to be hopelessly dependent upon the order of variables in X, it can be shown [Rozeboom, 1965] that Π_X^2 is the *generalized variance* [cf. Anderson, 1958, p. 166ff.] of configuration X—i.e., that

[99] $$\Pi_X = \prod_{i=1}^{n} \sigma_{(X)i} = \sqrt{|\mathrm{C}_{XX}|},$$

where n is the number of variables in X, $\sigma_{(X)i}$ is the standard deviation of the ith principal component of configuration X, and $|\mathrm{C}_{XX}|$ is the determinant of the X-configuration's covariance matrix. Since $|\mathrm{C}_{XX}|$ is unaffected by permutation of the variables in X, the same is true of Π_X and π_X; hence Π and π may be described as functions which take q-sets of jointly distributed metrical variables for their arguments.

Just as Π_X is the generalized standard deviation (i.e., the square root of the generalized variance) of configuration X, so may Pi-coefficient π_X, which symmetrically summarizes the coefficients of alienation holding among the variables in X, be thought of as the *generalized alienation* within multivariate distribution X. Its maximum value, unity, is attained when X is an orthogonal configuration whereas $\pi_X = 0$ (and similarly $\Pi_X = 0$) implies that at least one of the variables in X is an errorless linear function of the remainder.

The concepts of Pi-value and Pi-coefficient also obviously apply to configurations of residual variables. Specifically, $\Pi_{X\cdot Z}(\pi_{X\cdot Z})$ is the Pi-value (Pi-coefficient) of the joint distribution of the linear residuals of the variables in X after the variables in Z have been partialled out.

Since the Pi-statistic assigns a numerical value (relative to the background population) to each q-set $X = x_1 \cdots x_n$ from any domain $\mathbf{d}_{mv}$ consisting of jointly distributed metrical variables, it may be used to define any number of generating functions over $\mathbf{d}_{mv}$. Most of these, like measures of categorical uncertainty other than U, merit little if any attention; but one—again like U—is of outstanding interest. Specifically, let

[100] $$\mathrm{G_K}(X) \overset{\mathrm{def}}{=} \log \Pi_X ,$$

where again the base of the logarithm is a parameter. (Stipulation [97] insures that $\mathrm{G_K}(\phi) = 0$ as required by axiom [2].) The quantity $\mathrm{G_K}(X)$ might appropriately be called the "linear uncertainty" over joint distribution X—in fact, when the X-distribution is normal, $\mathrm{G_K}(X)$ differs from the Shannon–Ross metrical Uncertainty in X only by an additive constant. As consequences of [100] we have

[101] $$\mathrm{G_K}(X \mid Z) = \log \Pi_{X\cdot Z} ,$$

[102] $$\mathrm{C_K}(X \mid Z) = -\log \pi_{X\cdot Z} ,$$

of which the zero-order case is

[103] $$\mathrm{C_K}(X) = -\log \pi_X ,$$

and finally

[104] $$\mathrm{R_K}(Y; X \mid Z) = -\log \left[\frac{\Pi_{YX\cdot Z}}{\Pi_{Y\cdot Z}\Pi_{X\cdot Z}}\right] = -\log \left[\frac{\pi_{YX\cdot Z}}{\pi_{Y\cdot Z}\pi_{X\cdot Z}}\right],$$

various special cases of which are

[105] $$\mathrm{R_K}(y; x) = -\log \mathrm{k}_{yx} = -\tfrac{1}{2} \log [1 - \mathrm{r}_{yx}^2]$$

[106] $$\mathrm{R_K}(y; x \mid Z) = -\log \mathrm{k}_{yx\cdot Z} = -\tfrac{1}{2} \log [1 - \mathrm{r}_{yx\cdot Z}^2]$$

[107] $$\mathrm{R_K}(y; X) = -\log \mathrm{K}_{y(X)} = -\tfrac{1}{2} \log [1 - \mathrm{R}_{y(X)}^2]$$

[108] $$\mathrm{R_K}(y; X \mid Z) = -\log \mathrm{K}_{y(X)\cdot Z} = -\tfrac{1}{2} \log [1 - \mathrm{R}_{y(X)\cdot Z}^2]$$

and most remarkable of all,

[109] $$\mathrm{R_K}(Y; X) = -\log \prod_{i=1}^{m} \mathrm{k}_{(X,Y)i} = -\tfrac{1}{2} \log \prod_{i=1}^{m} [1 - \mathrm{r}_{(X,Y)i}^2],$$

in which $r_{(X,Y)1}, \cdots, r_{(X,Y)m}$ are the non-zero canonical correlations between configurations X and Y while the $k_{(X,Y)i}$ are the corresponding alienation coefficients [Rozeboom, 1965]. (Cononical theorem [109] also holds for general case [104] if $k_{(X,Y)i}$ and $r_{(X,Y)i}$ are replaced with the corresponding statistics for the X-residual and Y-residual configurations after variables Z are partialled out.) The abstract concept of G-contingency in the uninterpreted theory thus subsumes, under interpretation [100] of G, all linear correlations of various orders and complexities (or more precisely a certain monotonic transformation of their magnitudes), while the analogy between linear correlation and Uncertainty reduction turns out to be a sweeping isomorphism.

To be sure, not *all* abstract partials have intuitive significance in their correlational interpretation: Interactions $I_K^{(n-1)}(x_1 ; \cdots ; x_n \mid Z)$ for $n > 2$ correspond to no meaningful multivariate properties now known to correlation theory. The same is true, however, for Uncertainty analysis—both interpretations coordinate certain terms of the fundamental system with measures whose significance lies in the external merits of their substantive definitions, while the remaining components of the ramified system, interactions higher than first level in particular, acquire whatever significance they may have through their derivation from these system-entry positions. As it is, linear correlation theory's entry to the abstract-partials system is extraordinarily massive—*all* measures in the fundamental system are entry positions for the K-interpretation—and, rather than analysis of variance, is the proper analogy to consult when attempting to make sense out of Uncertainty components.

The parallel between linear correlation theory and Uncertainty analysis extends even beyond their abstract-partials isomorphism in that, like [76],

$$[110] \qquad G_K(x_1 \cdots x_n \mid Z) \leq \sum_{i=1}^{n} G_K(x_i \mid Z).$$

Hence $C_K(X \mid Z)$ and $R_K(Y; X \mid Z)$, like their Information-theoretical counterparts, have a lower bound of zero. There is, though, one important respect in which the parallel is less than perfect: Whereas G_U is unaffected by duplications of elements in its argument—i.e., $G_U(XXY) = G_U(XY)$ for any q-sets X and Y from the system's domain—element duplications introduce linear dependencies within a q-set $(XXY)(X \neq \phi)$ of metrical variables, whence it follows that $\Pi_{XXY} = 0$ and $G_K = -\infty$. Neither for that matter is there any element in domain $\mathbf{d}_{mv}$ which provides an inter-pretation for the fusion of a q-set from $\mathbf{d}_{mv}$. The concept of "fusion" can, however, be naturally introduced into the correlational interpretation—and for some purposes usefully so—by construing the latter's domain to be the set of all vectorial variables defined as ordered sets of variables in $\mathbf{d}_{mv}$, while the value of G_K for a q-set $X_1 \cdots X_n$ of ordered sets $X_1, \cdots, X_n$ from $\mathbf{d}_{mv}$ is $G_K(X_1 \cdots X_n)$ as in [100].

4. *Appraisal and Summary*

This paper has attempted no more than to articulate the formal abstract-partials structure and point out its embodiment by several of the statistical systems which behavioral scientists hold most dear. Whatever merit, if any, a particular segment of this structure may have in a given substantive interpretation is not our present concern. Even so, while abstract-partials theory has enough inherent mathematical appeal to warrant study for its own sake, a challenge which must eventually be faced is "What good is it?"

It must cooly be recognized at the outset that while the system of abstract partials is prepared to invest a substantive discipline with a ready-made array of intricately interlaced analytic measures the instant some aspect of its data is found to have the properties of a generating function, there is no guarantee that these measures will be at all useful. Each different interpretation of each abstract partial must be judged anew, with special concern for its relation to system-entry positions, and what is the key to hidden treasure in one need be no more than debris in another. In particular, the probable frequency of insightful interpretations for the ramified parts of the system is not especially encouraging, not even for the hierarchy of interaction terms.

Even with an appropriately sceptical guard posted against great expectations, however, it is still possible to recognize ways in which abstract-partials theory holds methodological promise. For one, the abstract analysis brings heightened mathematical power to established systems such as Information theory where a particular embodiment of the abstract-partials structure has already acquired scientific stature. By pruning derivations to their formal essentials, it is generally possible to establish theorems with greater elegance and generality, and to exhibit logical connections among the concepts more perspicuously, than can in practice be achieved when the material under study is fogged over with irrelevant substantive detail. (Thus the theorems in Sections 1 and 2 above are considerably more comprehensive than any previous development of Information theory, though how useful these additional results may be is of course another question. Similarly, the abstract development of the interaction hierarchy is both more general and more succinct than previous expositions of interaction terms in the analysis-of-variance literature.) Conversely, by considering what properties of a given substantive measure follows from its abstract-partials character alone, we can investigate whether it has any methodologically significant features which do *not* reside wholly in the abstract-partials structure, and if so, what precisely these additional properties accomplish.

More importantly, by invoking isomorphism across a variety of conceptual systems, abstract-partials theory transduces our familiarity with one into deepened understanding and provocative new twists of development in another. For example, while the degree of association between two attributes a and b has long been assessed by comparing the joint probability of a and

b to the joint probability they would have, given their marginal probabilities, in the absence of any relationship, all such measures developed to date have been based on the *difference* between Pr (ab) and Pr $(a) \cdot$ Pr (b) [cf. Lazarsfeld, 1961, p. 112]; whereas the strikingly successful conditional-probability interpretation of the abstract-partials system suggests that the *ratio* of these quantities may also be an analytically fruitful measure of attribute association. Again, abstract-partials theory discloses that the fundamental statistic of linear correlational analysis is the little-known *generalized variance* of a multivariate distribution, and urges that we investigate whether this measure may not have important applications which lie beyond our present limited vision. Moreover, while the significance of all the components which appear in symmetric partitions of Pi-coefficients and multiple coefficients of alienation (or in partitions of their negated logarithms) is far from clear, we know that these terms must have at least the same sort of meaning as the Uncertainty components to which they are isomorphic, since in view of the system-entry patterns in the two cases, the Uncertainty interpretation can claim no significance for any Uncertainty component which is not equally warranted for its correlational counterpart. In particular, no matter how much like gibberish the higher correlational interaction terms may seem, they must be fully as meaningful as the corresponding Uncertainty interactions—or conversely, the latter must be as meaningless as the former. This last consideration illustrates nicely how the theory of abstract partials can sharpen our comprehension of otherwise unrelated measures by pointing out formal identities in the conceptual routes by which they are derived. To be sure, an abstract-partials isomorphism between two substantive measures does not guarantee that they are essentially alike in their significance, for one can enjoy interpretive depths acquired from substantive details not possessed by the other (e.g., the interaction variances in analysis-of-variance, which under full predictor independence are not merely components of the criterion's variance but also variances of the criterion's components). An interpreted partial can acquire its meaning either from its position in the abstract-partials structure (relative, ultimately, to the interpretation's system-entry measures), from additional properties not inherent in the abstract-partials axioms, or from both. Hence when different quantities are identified as alternative interpretations of the same abstract partial, either the significance which is thought to invest one must be acceded to all or we must be able to make clear what is importantly distinctive about the one that is lacking in the others.

Finally, after all the words of caution have been spoken and the sophisticated illuminations of isomorphism properly extolled, there still remains the simple-minded fact that the system of abstract partials, wherever applicable, can be put to work computationally to grind out abundant assessments on endless combinations of the data elements which constitute the generating function's domain. Whenever a substantive measure G_s having the properties of a generating function is suspected to be at all relevant to the phenomenon under study, a very real possibility also exists that patterns of regularity can be found among the numerical values of the interpreted partials erected upon G_s which in one way or another implicate hypotheses or conclusions about natural principles operative in these data. For example, it may be possible to discover informative groupings ("clusters") of the data elements by studying how their total configural savings can be partitioned into within-group and between-group components according to equation [28] or [31].* Or we might look for traces of causal structure by partialling various q-sets of data elements out of the G_s-relations among the remainder by means of [20] and [21] in order to see which residual relations vanish. Or general regularities may appear, such as a tendency for the G_s-contingencies among the data elements to be a simple function of certain observable features of the latter, or a trend for interactions to diminish with increasing level, or etc., which cannot be written off to mathematical artefact and hence demand empirical explanation. There is even an outside chance that something akin to inferential factor analysis might be built upon equation [62], though this would probably require too many implausible assumptions to warrant serious concern. Whatever type of patterning, if any, may lie within the G_s-data will undoubtedly depend critically upon the G_s-function's substantive content. Even so, once we have conceived a portfolio of regularities which, if found, would command our respect, and have written computer programs which search for them, it should be little extra trouble (assuming computer availability) when the data are analyzed by more conventional methods, also to run the abstract-partials programs both for the sake of the data analysis itself and for whatever this shotgun approach may disclose about the empirical applicability of abstract-partials theory. The system of abstract partials is an intriguing new fine-mesh seine which is well worth a few test drags through interpretations additional to those reviewed above to see what it may catch.

REFERENCES

Anderson, T. W. *Introduction to multivariate statistical analysis.* New York: Wiley, 1958.

Attneave, F. *Applications of information theory to psychology.* New York: Holt–Dryden, 1959.

DuBois, P. H. *Multivariate correlational analysis.* New York: Harper & Bros., 1957.

Fhanér, S. Some comments in connection with Rozeboom's linear correlation theory. *Psychometrika*, 1966, 31, 267–270.

Garner, W. R. *Uncertainty and structure as psychological concepts.* New York: Wiley, 1962.

Garner, W. R., and McGill, W. J. The relation between information and variance analysis. *Psychometrika*, 1956, 21, 219–228.

Lazarsfeld, P. F. The algebra of dichotomous systems. In: Solomon, H. (ed.) *Studies in item analysis and prediction.* Stanford, Calif.: Stanford University Press, 1961.

*Application of the Information-theoretical interpretation of [28] for identifying groupings has already been pioneered by Watanabé [1961]. The possibility of its use for cluster identification in linear correlational analysis is raised in Rozeboom [1965].

McGill, W. J. Multivariate information transmission. *Psychometrika*, 1954, **19**, 97–116.

Ross, J. Informational coverage and correlational analysis. *Psychometrika*, 1962, **27**, 297–306.

Rozeboom, W. W. Linear correlations between sets of variables. *Psychometrika*, 1965, **30**, 57–71.

Rozeboom, W. W. *Foundations of the theory of prediction*. Homewood, Ill.: Dorsey Press, 1966. (a)

Rozeboom, W. W. Scaling theory and the nature of measurement. *Synthese*, 1966, **16**, 170–233. (b)

Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal*, 1948, **27**, 379–423, 623–656.

Watanabé, S. Information theoretical analysis of multivariate correlation. *IBM Journal of Research and Development*, 1960, **4**, 66–82.

Watanabé, S. A note on the formation of concept and of association by information-theoretical correlation analysis. *Information and Control*, 1961, **4**, 291–296.

Manuscript received 8/18/66

Revised manuscript received 9/7/67