AN ASPECT OF QUALITATIVE RESEARCH IN SPORT SCIENCE: ITS DATA AS VALUES OF VARIABLES

Petr Blahuš

Department of Kinanthropology, Division of Methodology
Charles University, Prague, Czech Republic

Original scientific paper
UDC: 001.5:796.012
Received: January 16, 1998
Accepted: March 1, 1998

Abstract:

Initiated by the practical research problems with an evaluation of categorized research material in several qualitative studies in physical and sport education a more general methodological problem is dealt with, viz. how to treat the data of such studies. Therefore, starting from the general methodological postulate of necessity of the "research variable" concept even in the qualitative research (Blahuš 1997) the author tries to bring further methodological support for the so-called integrative approach as a unified scientific strategy. The classification of categories produced by categorization during a qualitative research into purely qualitative, graded, and distanced categories is interconnected with the author´s previous classification of random variables (1993, 1996), with levels of measurement scales, and with the dichotomy of string vs. numeric variables of computational software. The Fundamental Postulate (cf. FPELSA in Sec. 8) for processing any kind of data, including the qualitative ones, was formulated on the basis that the data had to be justifiably recognized as values of variables prior to their meaningful evaluation.

Key words: kinanthropology, kinesiology, sport science, methodology, variable, data processing, measurement, qualitative data, qualitative research

Zusammenfassung:

EIN ASPEKT DER QUALITATIVEN WISSENSCHAFTLICHEN UNTERSUCHUNGEN IM SPORT: DEREN DATEN ALS VARIABLENWERTE

Ganz praktische Schwierigkeiten wissenschaflicher Untersuchung bei der Bewertung des kategorisierten Untersuchungsmaterials in einigen qualitativen Studien auf dem Gebiet der Sporterziehung und Sportpädagogik haben diese Studie von einem sehr allgemeinen methodologischen Problem angeregt: wie sollen die Daten aus den qualitativen Untersuchungen verarbeitet werden. Von dem allgemeinen methodologischen Postulat ausgehend, dass die Idee der "Untersuchungsvariable" auch bei den qualitativen Untersuchungen unvermeidbar sei (Blahuš, 1997), versucht der Autor, den sogenannten vollständigen Zugang als die einheitliche vollständige wissenschafliche Strategie methodologisch kräftiger zu unterstützen. Die Klassifizierung der während der qualitativen Untersuchung erschaffenen Kategorien in die völlig qualitativen graduierten und distanzierten Kategorien ist mit der früheren Klassifizierung der Zufallsvariablen desselben Autors (Blahuš, 1993; 1996) sowie mit den Messleiterebenen und mit der Dichotomie, die in den Computerprogrammen zwischen den numerischen und den Reihenvariablen entsteht, eng verbunden. Das Grundgesetz der Verarbeitung irgendwelcher, also auch qualitativer Daten, basiert auf der Tatsache, dass, bevor eine sinnvolle Bewertung und Interpretation unternommen wird, die Daten mit Recht und absoluter Zuversicht als die Variablenwerte erkannt und angenomen werden müssen.

Schlüsselwörter: Kinanthropologie, Kinesiologie, Sportwissenschaft, Methodologie, Variable, Datenverarbeitung, Messung, qualitative Daten, qualitative Untersuchung

Introduction

During recent years qualitative research as a rather new methodological approach has penetrated sport science and exercise pedagogy bringing discussions on its advantages and appropriateness under the specific conditions in that area of investigation. Several first applied studies have been carried out and among the plausible ones the following could be found as examples: Placek 1984 presented a combined casuistic study of physical education teachers, Bain 1985 studied the reactions of students to physical education lessons, a qualitative deep insight into the relationships between drug abuse and sports involvement was investigated by Stuck 1985, Gallmeier 1988 carried out participant observation in hockey, Fisher 1996 reported on his interpretative study to compare teacher vs. pupil opinions on physical education etc.

Later on, some theoretical methodological writings dealt with the evaluation and specification of the qualitative research features for its use in sport research, especially Earls 1986, Bain 1989, Locke 1989, Schutz 1989, Hendl 1997. Several more comprehensive and general papers and textbooks on research methodology in sport and human movement included a description and explanation of qualitative research and its techniques as a part of the wider spectrum of empirical research methods, cf. for example Schempp, and Choi 1994, but probably the most elaborated among them is the recent third edition of the well known **Research Methods in Physical Activity** by Thomas and Nelson 1996 which offers an innovated and extended chapter oriented towards qualitative research.

Among some of the qualitative researchers there exists an extreme opinion that the findings obtained by qualitative research lead inevitably to quite different or even opposing knowledge than traditional or classical research which is to be distinguished, then, by the label "quantitative". Contrary to such an opinion the author (1988, 1996a, 1997) stresses a unifying foundation of empirical sciences on the basis of a monistic branch of contemporary philosophy of science, namely the so-called methodological naturalism (Thomas 1979, Bhaskar 1979, Tudor 1989, Schwager 1991). The notion of methodological naturalism has to be strictly distinguished from positivism and its branches. In brief the basic idea of this doctrine is that in principle (!) there still exists a unity of "the scientific method" in all empirical sciences, especially in the natural as well as social ones. The latter, in a broader sense, includes behavioural science and kinesiology or kinanthropology[1], as well as sport science and sport and exercise pedagogy. As an example of a part of the unifying background the author (1996a) has developed the notion of "weak associative measurement" of theoretical concepts (constructs) to exhibit the similarity between measurement in the natural and behavioural sciences. Its application to the measurement of motor abilities as the basic explanatory notions of motor and sport performance in kinesiolgy was presented in this journal (1996b).

In any empirical science a borderline between its scientific theory and its specific empirical research methods is not sharp. The two areas influence each other mutually, the main task of empirical research being to contribute to the knowledge of legalities whose network creates the theory. Any legality is a kind of relationship between two (or more) attributes. The attribute which is to be recognized as scientifically meaningful must undergo some minimum conditions to avoid contradictions, from the point of view of logic at least. Then it can be accepted as a "research variable" (cf. author's 1997 paper for details). All different empirical research methods should follow this common mission to contribute to the knowledge of legalities and, thus, to build and develop jointly their scientific theory. The relative coherence of the empirical findings by different methods and their synthesis means a substantial mutual support for a higher degree of verification of the theory.

From the above point of view there is no principal difference in the use of participant observation in a physical education class, interpretation of a tape-recorded interviews of football coaches and/or testing sport talents by a battery of motor tests, administering standardized questionnaires on attitudes toward physical activity, or exercise physiology measurements. So, there is the mentioned common purpose of the qualitative as well as other methods of empirical research.

Although there are different opinions on the evaluation of qualitative research material some authors support an approach which uses certain formal methods for its analysis, especially the book by Miles and Hubermann 1994, which presents a highly elaborated description of the methods for qualitative data analysis. In the area of sport kinesiology and pedagogy special attention is paid to the analysis of data from the qualitative studies by Schutz 1989 and partly also from the textbook by Thomas and Nelson 1996.

In analyzing the data of a qualitative research the crucial stage is represented by categorization. This is a standard part of this kind of research which moves from primary

1 "Kinanthropology" is an officially recognized scientific discipline in the Czech Republic, more or less identical with the notion of "kinesiology", while the latter notion is understood as rather narrow, closer rather to the medical and physiotherapeutic meaning and with close relations to functional anatomy and biomechanics. More can be found in the following two papers, viz. Blahuš, and Dobrý, Hohler etal. 1993 and Renson 1990.

rough research material to its sorting and classification into categories. For the sake of convenience it usually ends with labelling the categories by abbreviated descriptions and different denotating symbols. Then a practical problem arises - how those can be used for qualitative data analysis. The author had some opportunities to meet confusions and misunderstandings during the consultations on using the suitable methods for data processing to the material and preliminary data obtained from qualitative studies. Therefore, the topic presented here is not only a continuation of the theoretical analysis of the notion of research variable (1997) but can also be seen as a practical memento for the researchers who wish to evaluate their data from the qualitative studies by data processing methods.

A minimum paradigm of the so-called non-paradigmatic qualitative research

In a former paper (1997) the author tried to bring together some evidence supporting the view that the sometimes claimed controversy between the so-called qualitative research and the so-called quantitative research is not necessarily so strict. On the contrary, some substantial common features of both approaches can be found to back the idea of "the integrative approach" to the empirical research in behavioural sciences including kinanthropology or sport science.

One of the supporting arguments was found in the fact that even if the qualitative research is desirable as free of any paradigm it still includes some minimum a priori assumptions to avoid logical contradictions and confusions. An example of such a "minimum paradigm" is the methodological concept of

research variable

which has been shown as necessary and fundamental even in the qualitative research.

Further, even qualitative research is supposed to identify relationships, the relationships relate to attributes, the attributes have to fulfill some requirements to avoid contradictions and therefore they are explicitly or implicitly considered in the form of research variables. It makes no difference that in qualitative research the attributes are usually formulated not before but during or after a research observation has been carried out. Moreover, the qualitative observation might be of a kind of quite subjectively perceived exploratory search experienced by a researcher during an ethnographic participated observation in the natural settings.

Categorization: From (primary) "research rough material" toward (secondary) research variables

In the practice of any exploratory qualitative research a post hoc, i.e. secondary, content analysis of the collected primary "research rough material" is carried out to generate "analytic categories". This process is well known as **categorization**, an important stage of a typical qualitative research. This secondary system of categories, however, should fulfill certain conditions so that it is able to represent real-world attributes. One group of the conditions deals with syntactic or logical requirements that are necessary so that an attribute is logically consistent, which is certainly necessary for its potential recognition as a "research variable" (Blahuš, 1997):

a) There are at least two categories in the set created by the categorization of the same attribute,

b) The categories are exhaustive, one of them must occur, and they are mutually exclusive (i.e. no event can occur in a set of categories, every event has to be classified into one category, and only into one).

Prior to the categorization many attributes of the primary "research rough material" are very weakly specified or even quite unknown. So, an attribute may be of a kind of rather unclear characterization which, step by step during the categorization, becomes more clear and its categories better specified in the standard procedure of qualitative research known as the "saturation of categories". Then, this

categorized attribute

whose categories fulfill the syntactic conditions (a), (b) and further theory ladden semantic requirements becomes

a research variable.

In some sense the categorization can be recognized generally as a kind of scientific systematization of the primary research material which is collected or which already has been collected. The systematization is implicitly or explicitly driven by the abovementioned a priori paradigm.

In the author´s paper (Blahuš, 1997, Section 4) a system of classification of transition levels in data assignment process was suggested. From its point of view the rise of a research variable expresses just the first level at which, although the variable was properly categorized, its categories need not be assigned to formal symbols yet. So, at this level the research variable is not yet formalized or symbolized. Therefore, at this stage, its categories are governed by a kind of non-formalized relations - empirical relations, namely.

The very fundamental empirical relation is

empirical coincidency,

a notion well known in the representation measurement theory. In the present context it means that a certain element of the rough research material is empirically classifiable into a category by the use of a certain empirical rule (or collection of rules) applied by the researcher in the process of categorization. Such rules must enable the classification, meaning that the relation of coincidency must be

empirically decidable.

In other words, the researcher has to be able to distinguish the elements which belong to different categories and recognize the elements which belong to the same category. Thus, under specific conditions of a concrete problem of given qualitative research study, the empirical coincidency relation takes on its specific practical instance in the process of categorization analysis.

From the point of view of formal logic the coincidency relation bears characteristic signs of equivalency relation thus being sometimes called relation of "a type of equivalency". Still, at this level of transition from rough research material, the coincidency relation which is carried out through the empirical processes cannot be fully recognized as identical with "the equivalency", i.e. as a relation of formal logic or set theory. Such a recognition would be made possible through the next step of transition to the further level when the non-formalized variable becomes formalized. From this point of view we could say that the empirical non-formalized relation is then "modelled" by a theoretical formal-logic relation. On the other hand the supporters of formalism among the logicians would say that an empirical model or a real-world interpretation has been found in the well known relation of formal logic.

Non-numeric coding of categories of research variables

The categories that are produced by categorization analysis and governed by the empirical coincidency relation are then coded by a variety of abbreviations, preferably by one-character symbols, typically by ASCII characters, but not by numbers in the "true sense". This process has been called "minimum data coding" (Blahuš, 1997) resulting thus in the

formalized research variable.

If, however, the numerals are used as symbols then the coding characters should still be recognized and interpreted as qualitative symbols, namely as numerals (and not as "true" numbers). Such a special method of data coding should be called

numeralization

but it should never be recognized as "quantification" or "measurement"or "scaling" or a "representation by numbers".

In terms of the set theory the coding of categories is the one-to-one mapping, i.e.

isomorphic mapping,

into formal symbols. (In this approach we prefer the isomorphic to homomorphic mapping which is also considered in the area of measurement theory.)

From the point of view of Stevens´ typology of measurement scales the isomorphic mapping into numerals produces the type of data which is close to his concept of nominal type. But the original Stevens´ terminology is different and quite inappropriate because his nominal data do not form any "scale" at all.

Even the advocates of broad conception of measurement usually agree on using the term "weak measurement" just to Stevens´ ordinal scale type and not to accept the contradictive term "nominal scale".

For the same reason the term "representation by numerals", perhaps borrowed from the common nonscientific language, would be quite contradictive since in the theory of representation measurement it is understood by definition that

representation is a mapping into ordinal numbers,

at least in the sense of the well known representation theorem (e.g. Coombs et al. 1970 and many others).

We suggested a general name for these research variables that have categories without any ordinal meaning (the author´s paper 1997), namely

purely qualitative variables.

A formalized research variable which is purely qualitative and whose formalization yields data coded by numerals can be called

numeralized research variable,

i.e. formalized by mapping into numerals, but not into numbers. To present an example let us consider the following illustration of a numeralized variable, say

"Favourite type of recreational sport activities",

with possible categories coded by the numeralization:

"1" for "indoors",

"2" for "outdoors",

"3" for "others".

So, to summarize, **numeralized research variable** is a special sub-type of **formalized** research variable, namely that one which is **purely qualitative**. The symbols used for its formalization by "minimum data coding" can be understood as numerals and treated only as

string variables

in computerized data processing.

On the other hand, if the symbols are to be treated not only as non-numerical characters then we confront the area of questions discussed in the following sections.

Three types of categories produced by categorization in qualitative research

Even in qualitative research an original empirical attribute under scope need not be necessarily

"purely qualitative".

From a general philosophical point of view it may be of another ontological-gnoseological kind. It can be of

"a weakly quantitative nature" and/ or

"a quantitative nature"

to use the terminology of Berka 1982. To express it in the measurement theory and classical Stevens´ terms, it can be represented by an ordinal scale or an interval scale.

The three types are sometimes called "levels of measurement". From the point of view of the representation measurement theory they are distinguished according to the so-called conditions of measurability which deals with the mapping of formal and abstract numeric relations and operations into the real world of empirical relations and operations. The conditions asking for empirical counterparts to the arithmetic relations "equal to" and "greater than" are called by Berka (1982)

topological conditions - they correspond to the "weakly quantitative nature" of attributes that can be represented by pseudomeasurement on an ordinal scale through the so-called semiquantification.

Together with the empirical relation of coincidency another one has to be found, viz empirical relation of predecessing.

If those two are augmented by empirical counterparts of the numeric operation of addition and of real world constant amount of the empirical content corresponding to the number "1" then the

metric conditions have been fulfilled and the quantitative attribute is measurable on an interval scale.

Here the empirical counterpart is called concatenation of the unit amount, to use the terminology of the representation measurement theory again.

The corresponding nature of a research variable and the possible level of its measurement becomes clear during the stage of qualitative research which deals with the process of categorization, i.e. during the expertise and analysis of the semantic content of primary "research rough material" into categories. Then, if the categories obtained are not purely qualitative they can be found to be:

- *graded* - allowing the ordering of the categories according to their semantic content
- *distanced* - expressing the magnitude of difference between the amounts of their semantic content

(cf. the author´s paper 1997, Prop. 10).

In some qualitative studies the distanced categories can be established directly during the subjective process of categorization, for example, when the categorization analysis of a semantic content of individual categories also includes a kind of weighted differentiation between their content as an inevitable and inherent part of the categorization. Say as in the case when a category "A" is identified by the content analysis as "twice more important" than category "B" or so. Such an approach is well known and is used in the so-called method of "direct scaling". It is widely used for the subjective evaluation of similarities in psychophysics, e.g. the similarity of colours, etc. In some other qualitative studies the distanced categories can be established not so directly but with the use of more sophisticated scaling methods that are typical for the analysis of expert evaluation. i.e. by such scaling methods as Thurstone´s pair comparisons, equally appearing intervals, judgment of ratio of common content, etc. (cf. for instance the author´s older 1984 review of scaling methods and also the series of more recent textbooks on scaling).

Numeric coding of categories of research variables

Naturally, categories of any type of research variables can also be coded by

numeric symbols,

i.e. the symbols that may not be treated only as numerals. Still, it does not mean that the symbols can be treated as "genuine" numbers automatically.

For example, let us remember the aforementioned coding of the numeralized purely qualitative research variable by "1" for "indoors", "2" for "outdoors", "3" for "other". Now let us consider the question whether the symbols "1", "2", "3" could be recognized not only as numerals. If we try to ascribe them at least some of the properties that apply to numbers then we would be able to create a very degenerate kind of numbers, i.e. numbers governed by only one numeric property, the first one of the two topological conditions - namely the arithmetic relation "equal to" (and/or its opposition). In a certain sense the arithmetic relation of equality may be considered as a special kind of formal-logic relation of equivalency, or as being of "a type" of it.

This type of mapping of the categories into such numeric symbols produces that very degenerate case which corresponds to the Stevens´ nominal type of data. To distinguish this type of mapping from representation - which produces mapping in ordinal numbers at least - we may call it perhaps

pseudo-representation.

Thus, the present system of notions on the isomorphic mapping of categories of variables can be organized into the following list with descending generality:

isomorphic mapping into **non-numeric** symbols:
formalization, symbolization

numeralization (as a special case of symbolization)

isomorphic mapping into **numeric** symbols:

- *pseudo-representation, qualitative labelling by numbers - numerization*
  - *-representation, ordinal pseudomeasurement, or semi - quantification*
    - *-metric representation, measurement, quantification*

N.B. the difference between

numeralization vs. numerization.

The former stands for symbolization by numerals as non-numeric characters. The latter means labelling qualities by numbers which, however, have their numeric properties limited only to the arithmetic relation of equality. Therefore in computerized data processing the numeralized ones have to be treated as string variables whilst the numerized ones may be processed as a numeric variable but with the strict limitation mentioned above. Certainly, there are no difficulties to convert the "computerized coincidence relation" comparing the identity of strings into the arithmetic relation of equality of numeric values and vice versa. From this point of view one could say that numerization - as pseudo-representation by numbers - yields numbers that are

"numerals in numeric disguise",

actually. Additionaly, especially for the purely qualitative research variables which are numerized (!) it also means that they can be treated in data processing on a computer quite **arbitrarily** either as

- string variables, and/or as
- numerical variables

(but with the numeric properties restricted only to the arithmetic relation of equality, as mentioned above).

The system described up to now can be summarized in Table 1 which, by the way, should also stress the following series of mutually tied notions:

(empirical) STATE - (research) CATEGORY - (numeric) VALUE.

Table 1: Three levels of formalization (I-III) and three levels up to quantification (1.-3.)

Ontological gnoseological nature:	Qualitative research categorization:		Measurement theory:
Unstructured STATES of undefined	CATEGORIES of		VALUES of
ATTRIBUTE of research material	RESEARCH VARIABLE		SCALE
	NON-FORMALIZED	SYMBOLIZED incl. numeralized	
	I.systemization	II. symbolization inc. numeralization	III. representation
1. non-quantitative, i.e. qualitative	purely qualitative categories	symbols (incl. numerals)	numerization[x] ("pseudo-representation")
2. weakly quantitative	graded categories		numbers on ordinal scale
3. quantitative	distanced categories		numbers on interval scale

[x] *This pseudo-representation through labelling the pure qualities by numbers corresponds to the Stevens´ term "nominal", but it may never be recognized as either a scale or as a genuine representation which is supposed to be the case of numerical mapping into ordinal scale at least.*

The dual status of the categories: events and/or variables ?

As regards any practical treatment of data that map the original categories there are two main conceptual approaches:

1) The first one is based on an implicit assumption that each **individual category** is a separate **event**, and that the event either occurs or it doesn't.

2) In the second conceptual background - let us borrow some set-theoretical terminology - **the complete set of all *n* categories** of the same research variable is taken into consideration as a domain of a one-to-one mapping of categories into the range of formal symbols that are then recognized as **data-values of a formalized research variable.**

Nevertheless, there is a two-way bridge between the two conceptions.

On the one hand any domain of categories in the second approach (2) can be decomposed into individual categories which are then considered as separate events. As a byproduct of such decomposition one would encounter the fact that one of the events then becomes a complement of the others because of the logical properties given by the definition of the notion of the research variable. Thus, necessarily, there is one fully redundant and avoidable event among the *n* events and the original *n* categories of the original variable should be substituted by *n - 1* events. Certainly, if the number of categories *n* is high then this approach is not only inconvenient but rather confusing or confronts the researcher with an insurmountable obstacle in his/her effort to find some regularities in such dissipated and intricate data.

On the other hand, in the first approach (1) the alternatives of a separate event, i.e. its occurrence and non-occurrence, may be interpreted as the complete set of categories of a research variable whose domain consists of n=2 categories only. In this way any separate event can be recognized as a special case of research variable, i.e. such that has only two possible categories. Such variables are sometimes called

dichotomous** or **binary.

It is very important to keep in mind that even such a dichotomous research variable can still have its two categories

purely qualitative

as well as

graded,

depending on the interpretation and the possible presence of an empirical relation of predecessing between the categories. So, the typical case of a so-called "yes-no" variable may be either purely qualitative and/or graded depending on the fact whether the "yes" (occurrence) means a better alternative than "no" (non-occurrence) or not, as implied by the context of the given research study.

Although the latter conception (2) transforms *n* events into *n* research variables, it does not bring the difficulties which were described for the former conception (1) as far as the higher number of *n* categories was considered. Therefore we definitely prefer the approach dealing with

the concept of "variable"

to the approach based on "events". So, in the following text exclusively the approach (2) based on the concept of variable will be considered as the chosen end of the aforementioned bridge.

The research variable vs. mathematical or random variable

Let us focus on the approach (2) of the preceding section and ask the question if, and under which condition, the research variable can be treated as a source of data for the purpose of data processing.

Firstly, let us consider the deterministic version which works with the concept of

*deterministic **quantitative mathematical variable**, i.e.*

mathematical quantity

which takes on

numbers** as **values

that express its quantitative magnitude on the metric scale. The relationships between the variables of this kind are typically expressed by exact mathematical functions without any error and with no necessity to take some probabilistic aspects into account.

Secondly, since there are almost no such exactly describable situations in behavioural research, let us weaken the strict assumptions and consider a probabilistic counterpart to the above mentioned deterministic quantitative mathematical variable as it is found in the area of the probability theory, namely the

*stochastic **random variable**[2] or*
random quantity

which takes on

***numbers** with assigned **probabilities** as its*
random values.

These numbers also express magnitudes on the metric scale but this time they occur with a certain probability, only.

Thus, from the point of view of the representation measurement theory, a research variable with distanced categories on an interval scale can be metrically represented by one of the two variable quantities - the deterministic and/or the stochastic one. Of course, the representation by deterministic version is rather rare in applications, more often in the areas like biomechanics and perhaps in some other areas where the stochastic component and random noise may be almost neglected.

From the viewpoint of theory of modelling we could express the same by saying that

the research variable is "modelled"

by one of the latter variable quantities.

Thus, on the one hand the basic focus of the measurement theory is concentrated on the meaningfulness or

the appropriateness of the representation.

On the other hand, the main problem of the modelling viewpoint is focused on the kind of

the "fit" of a model,

i.e. on the question whether the quantitative mathematical variable or random variable is a sufficiently good model of a given research variable.

The "appropriateness of representation", in one terminology, or "fidelity of modelling", in the other, should be evaluated at the most primary or primitive level of the representation and/or modelling. It seems that such a very elementary level can be expressed by the following necessary postulate:

THE FUNDAMENTAL POSTULATE

for processing numeric data
by quantitative methods:

The **numeric data** are recognized as **values** of variables, especially,

(a) either in the deterministic version
*as **values** of*

*quantitative **mathematical variable,***

(b) or in the stochastic version
*as **values**, accompanied*
by their probabilities, of

random variable.

As behavioural research rather deals with non-deterministic data the stochastic version of the Fundamental Postulate is more often under the scope of practical use. Then it can be interpreted, as the most elementary assumption to be accepted

prior to the application of any statistical method or model.

So, in any empirical research the practical imperative

"check if the data fulfill the conditions so that they may be recognizable as values of random variables"

is the basic starting level where any statistical modelling begins.

Therefore the stochastic version of the Fundamental Postulate could also be interpreted as

Fundamental Postulate

on an Elementary Level of Statistics Application

or abbreviated, FPELSA.

Its validity - at least as an assumption - is of the utmost necessity so that all the further machinery of modelling and computer operations be worthwhile.

So, FPELSA underlies any kind of statistical processing of data. This circumstance is a source of extraordinary importance of the mapping of inter-relationship between the type of numeric data as values generated by "their" random variable on the one hand, and the original categories of research variable on the other.

2 Another synonym for it being also "variate", cf. Kendall and Buckland 1975. Recently this term has not been used too frequently however it appears in the composed terms as "multivariate" or "univariate" etc.

Three types of random variables corresponding to the three types of categories

Expressed in terms of the set theory and the probability theory, the abovementioned mapping inter-relationship expressed in terms of FPELSA deals with the correspondence between the original events of probabilistic sample space and their numerical representation by the values of random variable. More precisely, it deals with the problem how well the possible non-numeric relations among the events of sample space are represented in terms of numeric relations among the number values which the random variable takes on.

In several writings of the author (especially 1988, 1993) three different types of random variable were introduced with respect to fulfilling the corresponding conditions for the numerical mapping of sample space to express the relations between the elements of the sample space in terms of numbers (also cf. Chuquai 1991 as one of the rare authors who have dealt with this problem). The three types have been called using the terminology of Berka´s 1983 classification of conditions of measurement (cf. also sections above):

numerized, topological, and metric random variable,

the last one to be also called

random quantity.

All of them were subsummed under a common general concept of the

"random variable-not-necessarily-quantity"

(cf. Blahuš 1993).

Now, in the light of the categorization analysis as a tool of qualitative research, the three types of random variables correspond to, and are to be assumed as numeric expressions of, the three kinds of research variables already discussed in this paper. That mutual correspondence is expressed in Table 2.

From this point of view a practical application of FPELSA in research and prior to data analysis means to appeal to the identification of a given research variable with one of the three types of the random variable-not-necessarily-quantity.

Illustration of some statistical methods corresponding to different types of variables

A practical application of FPELSA to different types of research variables in the sense of Table 2 will also help to ascribe the different statistical methods to the corresponding types of research variables and to process their data in an appropriate way.

In the frame of statistical data processing many stochastic or statistical methods are used for the evaluation of **purely qualitative** research variables which are treated in the form of **numerized random variables** quite typically.

There is a large pool of such methods at one's disposal and all of them are based on frequencies (absolute or relative ones). These frequencies can be used to estimate the **probability** and the joint probabilities. Among many others we find methods as discrete **stochastic methods**, e.g. Markov chains, and **statistical methods**, as contingency tables, goodness-of-fit tests, percentage tests, etc. There are several other methods especially recommended for analysis, evaluation and interpretation of this kind of qualitative data, e.g. see Schutz 1989, Hendl 1997, etc. Applications of these mathematical methods are generally accepted by the common

Table 2: Correspondence between the types of (a) research variables and (b) random variables

TYPES of	
a) **RESEARCH** VARIABLE	b) **RANDOM** VARIABLE NOT-NECESSARILY-QUANTITY
purely qualitative (incl. random events)	numerized
graded	topological
distanced	metric, random quantity

strategy of qualitative research, as by Thomas & Nelson 1990, p. 332:

"...qualitative research can use a wide range of quantitative analysis. This is especially appropriate in studies that use some type of observational instrument that codes designated categories. Frequency data in qualitative research are often nominal,... contingency tables are sometimes used...".

Additionaly, there are many classical parametric statistical methods based on the **metric random quantity concept.**

They include, for example, univariate as well as multivariate **tests of statistical hypotheses**, simple and multiple **regression,** analysis of variance - ANOVA, MANOVA, factor analysis and latent structure analysis (LSA, IRT), **structural equation models** (SEM) and many others.

There is a third group of statistical methods, which probably does not have its counterpart so elaborated among the deterministic methods, viz. non parametric rank-order **methods for topological random variables** that take on their values on an ordinal scale. They are usually called **semiquantitative methods.**

Among these we can find well known median tests, rank correlations (e.g. Kendall´s tau), and ordinal counterparts to the parametric methods mentioned above.

Still, a more informative list of statistical methods, which should represent their possibilities better, would be too long and out of the scope of the present communication.

Conclusion

In a continuation to the author´s former paper of 1997 the aim of the present article is to present evidence that data obtained by a qualitative research study bear some substantial features common to any kind of empirical research, namely, that they are to be recognized as (not necesarily quantitative) "values" of research variables. This is stressed here as a further integrative moment for qualitative and quantitative types of research.

So, the formerly shown minimum paradigm of the research variables (Blahuš, 1997) with its three types of categories produced by categorization analysis, i.e. purely qualitative, graded, and distanced categories, is now interconnected with another of the author´s former classification of random variables (Blahuš, 1993, 1996) into numerized, topological, and metric. This nature of data is shown to be basic and common to any kind of empirical research. Those classifications partly correspond to two kinds of variables that are typically distinguished in computerized data processing, viz. numeric vs. string (character) variables.

The approach which interprets the categories yielded by the categorization analysis of primarily collected "research rough material" as "values" of research variables is recommended to be preferred to the approach which could conceptualize the categories as individual separate events.

The coding of categories in the form of data is then hierarchically ordered as an isomorphic mapping of the categories into formal signs and strings of signs on several possible levels, i.e. mapping into non-numeric symbols: formalization or symbolization, numeralization (as a special case of symbolization), and/or mapping into numeric symbols: pseudo-representation which is qualitative labelling by numbers called numerization, ordinal representation called pseudomeasurement or semi-quantification, and, finally, metric representation called measurement or (genuine) quantification.

The Fundamental Postulate for data processing is formulated, on the basis that data have to be appropriately recognized as values of variables first. This should be checked prior to any practical data evaluation, even - and maybe especially - in qualitative research data to prevent methodological confusions as well as practical difficulties in the results interpretation which can be found in some qualitative studies in the area of physical education and sport research. Creating a clear, formally non-contradictive collection of data through categorization analysis and understanding typology of categories also enables the application of data processing methods.

References

1. Bain, L.L. (1985). *A naturalistic study of students response to an exercise class.* J. Teach. Phys. Educ., 5, 1:2-12.

2. Bain, L.L. (1989). *Interpretative and critical research in sport and physical education.* Res. Quart. Exer. Sport, 60, 1:21-24.

3. Berka, K. (1982). *Measurement: Its concepts, theories, and problems.* Dordrecht: Reidl.

4. Bhaskar, R. (1979). *The possibility of naturalism: A philosophical critique of the contemporary human sciences.* Brighton, UK: Harvester.

7. Blahuš, P. (1988). *K pojmu nahoda promenna ve společensko-vednich aplikacich statistiky z hlediska teorie mereni. (On the random vriable notion in applied statistics from the viewpoint of measurement theory).* Českosl. Psychol. 30:572-82.

8. Blahuš, P. (1993). *Random variable - a view of representation measurement theory.* In: R. Steyer et al. (Eds.), Psychometric methodology. Stuttgart: Fischer, pp. 84-91.

9. Blahuš, P. (1996). *K systemovemu pojeti statistickych metod v metodologii empirickeho vyzkumu chovani. (On system paradigm for statistics in the methodology of behavioural research.)* Praha: Carolinum, Charles University Press, 224 pp.

10. Blahuš, P. (1996). *Concept formation and latent variables modelling of motor abilities. Kinesiology,* 28, 2:12-21.

11. Blahuš, P. (1997). On the concept of "research variables" in qualitative and quantitative research. *Acta Univ. Carol. Kinanthropol.,* 33, 2: in press.

12. Blahuš, P., Dobry, L., Hohler, V., Hošek, V. et al. (1993). Kinanthropology - a new recognized scientific discipline. *Acta Univ. Carol. Kinanthropol.,* 29, 2:61-78.

13. Chuaqui, R. (1991). *Truth, possibility, and probability.* Amsterdam: North-Holland, Mathematical Studies.

14. Coombs, C.H., Dawes, R.M. and Tversky, A. (1970). *Mathematical psychology: An elementary introduction.* Englewood Clifs, NJ: Prentice Hall.

15. Earls, N. (1986). Naturalistic inquiry: Interpretative research and the insider-outsider perspective. *J. Teach. Phys. Educ.,* 6,1:1-108 (special monograph).

16. Fisher, R.J. (1996). A comparative study of teachers and pupils interpretation of physical education. *J. Comp. Phys. Educ. Soc.* 18:50-57.

17. Gallmeier, C.P. (1988). Methodological issues in qualitative sport research: Participant observation among hockey players. *Sociol. Spectr.* 8:213-235.

18. Hendl, J. (1997). *Uvod do kvalitativniho vyzkumu. (Introduction to qualitative research.)* Praha: Carolinum, the Charles University Press.

19. Kendall, M.G. and Buckland, W.R. (1975). *A dictionary of statistical terms.* London: Longman Group.

20. Locke, LF. (1989). Qualitative research as a form of scientific inquiry in sport and physical education. *Res. Quart. Exer. Sport,* 60,1:1-20.

21. Miles, M.B. and Hubermann, A.M. (1994). *Qualitative data analysis. A sourcebook of new methods.* London: Sage.

22. Placek, J.H. (1984). A multi-case study of teacher planning in physical education. *J. Teach. Phys. Educ.,* 4:39-49.

23. Schempp, P. and Choi, E. (1994). Research methodologies in sport pedagogy. *Sport Sci. Rev.,* 3:41-55.

24. Schutz, R.W. (1989). Qualitative research: coments and controversies. *Res. Quart. Exerc. Sport,* 60:30-35.

25. Stuck, M.F. (1985). *Adolescents, drugs, and sport involvment - a qualitative study.* Washington DC: National Institute of Justice monograph.

26. Schwager, K. (1988). *Theories of measurement in social science - a critical review.* Rotterdam, Erasmus University, Doctoral thesis. (Published by Wall-Thomson, Toronto 1991.)

27. Stevens, S.S. (1959). Measurement, psychophysics and utility. In: C.W. Churchman and P. Ratoosh, (Eds.), *Measurement: definitions and theories.* N. York: Wiley, pp. 18-63.

28. Thomas, D. (1979). *Naturalism and social science: a post-empiricist philosophy of social science.* Cambridge, UK: University Press.

29. Tomas, J.R. and Nelson, J.K. (1996). *Research methods in physical activity.* Champaign, IL: Human Kinetics.

30. Tudor, A. (1982). *Beyond empiricism: Philosophy of science in sociology.* Londom, UK: Routledge.