



# Quantitative methods in psychology: inevitable and useless

Aaro Toomela\*

Institute of Psychology, Tallinn University, Tallinn, Estonia

**Edited by:**

Jean-Luc Patry, *Fachbereich  
Erziehungswissenschaft der Universität  
Salzburg, Austria*

**Reviewed by:**

Shevaun D. Neupert, *North Carolina  
State University, USA*  
Richard Gonzalez, *University of  
Michigan, USA*

**\*Correspondence:**

Aaro Toomela, *Institute of Psychology,  
Tallinn University, Narva Road 25,  
Tallinn 10120, Estonia.*  
e-mail: aaro.toomela@ut.ee

Science begins with the question, what do I want to know? Science becomes science, however, only when this question is justified *and* the appropriate methodology is chosen for answering the research question. Research question should precede the other questions; methods should be chosen according to the research question and not vice versa. Modern quantitative psychology has accepted method as primary; research questions are adjusted to the methods. For understanding thinking in modern quantitative psychology, two epistemologies should be distinguished: structural-systemic that is based on Aristotelian thinking, and associative-quantitative that is based on Cartesian–Humean thinking. The first aims at understanding the structure that underlies the studied processes; the second looks for identification of cause–effect relationships between the events with no possible access to the understanding of the structures that underlie the processes. Quantitative methodology in particular as well as mathematical psychology in general, is useless for answering questions about structures and processes that underlie observed behaviors. Nevertheless, quantitative science is almost inevitable in a situation where the systemic-structural basis of behavior is not well understood; all sorts of applied decisions can be made on the basis of quantitative studies. In order to proceed, psychology should study structures; methodologically, constructive experiments should be added to observations and analytic experiments.

**Keywords:** quantitative methodology, epistemology, causality, mathematics, constructive experiment

Science begins with questions. Everybody can have questions, and even answers to them. What makes science special is its method of answering questions. Therefore a scientist must ask questions both about the phenomenon to be understood and about the method. There are actually not one or two but four principal questions that should be asked by every scientist when conducting studies (Toomela, 2010b):

1. What do I want to know, what is my research question?
2. Why I want to have an answer to this question?
3. With what specific research procedures (methodology in the strict sense of the term) can I answer my question?
4. Are the answers to three first questions complementary, do they make a coherent theoretically justified whole?

First, there should be a question about some phenomenon that needs an answer. Next, the need for an answer should be justified – in science it is quite possible to ask “wrong” questions, which answers do not help understanding the studied phenomena. Vygotsky (1982a) gave in his colorful language an ironic example of answering scientifically wrong questions:

One can multiply the number of citizens of Paraguay with the number of versts [an obsolete Russian unit of length] from Earth to Sun and divide the result with the average length of life of an elephant and conduct this whole operation without a flaw even in one number; and yet the number found in the operation can confuse anybody who would like to know the national income of that country

(p. 326; my translation).

It can be said that modern psychology is more advanced than science of Vygotsky’s time; perhaps the questions asked in the modern science are meaningful. This opinion, however, may be wrong. One source of wrong questions about the studied phenomena is unsatisfactory answer to the last question – when answers to first three questions do not agree one with another. In this paper I am going to suggest that psychology asks “wrong” questions far too often. The problem is related to the mismatch in answers to the first and third question. Specifically, quantitative methodology that dominates psychology of today is not appropriate for achieving understanding of mental phenomena, psyche.

The number of substantial problems with quantitative methods brought out by scholars is increasing every year. Already one observation could make scientists cautious. The questions provided above are in a certain order – first we should have a question about the phenomenon and only then the appropriate method for finding an answer should be looked for. Substantial part of modern psychology follows the opposite order of decisions – first it is decided to use quantitative methods, and the question about the phenomenon is already formulated in the language of data analysis. Between 1940 and 1955, statistical data analysis became *the* indispensable tool for hypothesis testing; with this change of scientific methodology, statistical methods began to turn into theories of mind. Instead of looking for theory that perhaps can be elaborated with the help of statistical tools, statistical tools began to determine the shape of theories (Gigerenzer, 1991, 1993).

For instance, a researcher may ask, how many factors emerge in the analysis of personality or intelligence test results. But why to look for the number of *factors* if personality or intelligence is

studied? We would guess here that the original question may be something like, is it possible to identify distinguishable components in the structure of personality or intelligence? However, the decision to use factor analysis for that purpose must be justified before this method is chosen. This justification seems to be missing; it is only a hypothesis – ungrounded hypothesis – that factor analysis is an appropriate tool for identifying distinct mental processes that underlie behavioral data (filling in a questionnaire is behavior). The problems emerge already with the determination of the number of factors to retain. There are formal and substantial criteria for that. Formal decisions are based on Kaiser’s criterion, Cattell’s scree test, Velicer’s Minimum Average Partial test, or Horn’s parallel analysis. There is no evidence that any of these criteria is actually suitable for distinguishing the number of distinct processes that underlie behavior. Researchers also decide the number of factors on the basis of comprehensibility – the solution which generates the most comprehensible factor structure is chosen. But this substantial criterion is always used after applying formal criteria; nobody starts from the possibility that all, say, 248 items of an inventory correspond to 248 distinct mental processes. The number of factors usually retained – from two to six or seven – seems to correspond to processing limitations of the researcher’s working memory rather than to true structure of the mind.

In this paper, epistemological issues that underlie quantitative methods used in psychology are discussed. I suggest that, regardless of the research area in psychology, mathematical procedures of any kind cannot answer questions about the structure of mind. The discussion focuses primarily on the statistical methodology as used in psychology today; yet there are fundamental problems inherent to other kinds of mathematical approaches as well. My intention is not to suggest that scientific studies of mind should reject mathematical approaches. Rather, it should be made clear, which questions can be answered with the help of mathematical methods and which cannot.

### WHICH QUESTIONS CAN AND WHICH CANNOT BE ANSWERED BY THE STATISTICAL DATA ANALYSIS PROCEDURES?

We should look for reasons to use statistical data analysis into the works of those, who introduced quantitative methodology into sciences in general and psychology in particular. Today, as a rule, users and developers of statistical data analysis procedures do not ask any more which questions can and which cannot be answered with the help of those procedures. Scholars who introduced mathematical procedures, however, made it clear, what kinds of answers they are looking for. We will see that these scholars would reject the questions answered by statistical procedures today for reasons that are largely ignored without any scientific reason by modern researchers. One of the most influential figures in introducing factor analysis into psychology was Thurstone. There are several ideas in his fundamental work *The vectors of mind* that are worthy of attention (Thurstone, 1935). These ideas, in the most part, underlie the use of not only factor analysis but the use of all forms of covariation-based data analysis procedures.

### WHAT ARE THE (STATISTICAL) CAUSES OF RELATIONSHIPS BETWEEN VARIABLES?

Thurstone suggested that the object of factor analysis is to discover the mental faculties. It is interesting that for him factor analysis alone would have never been sufficient for proving that a new faculty has been discovered – he held a position that results of factor analysis must be supported by the experimental observations. In another work he found that in some fields of studies tests are used that are not tests at all – “They are only questionnaires in which the subject controls the answers completely. It would probably be very fruitful to explore the domain of temperament with experimental tests instead of questionnaires.” (Thurstone, 1948, p. 406). So, Thurstone would very likely reject the modern practice to study many psychological phenomena – personality, values, attitudes, mental states, etc. – with questionnaires alone as it is often done now. There seems to be no theory that would justify studies of the structure of mind only by questionnaires. Without a theory that links subjectively controlled patterns of answers to objective structure of mind the results of all such studies are not grounded. Thorough analysis of this issue is beyond the scope of the current paper.

We can ask, what was the general question Thurstone aimed at answering with the help of factor analysis? Thurstone was not asking how mental faculties operate; he was looking for identification of what he called *abilities*, i.e., traits (which are attributes of individuals) which are defined by what an individual can *do* (Thurstone, 1935, p. 48).

The same questions about identification of “abilities” underlie the use of not only factor analysis but other covariation-based statistical data-analysis procedures as well. Here it is feasible to go deeper into the roots of introducing quantitative data-analysis into sciences. Methods for calculating correlation coefficients entered sciences somewhere in the middle of the 19th century but became popular with the works of Pearson (cf., 1896). He formulated the tasks of statistical data analysis in the following way:

One of the most frequent tasks of the statistician, the physicist, and the engineer, is to represent a series of observations or measurements by a concise and suitable formula. Such a formula may either express a physical hypothesis, or on the other hand be merely empirical, i.e., it may enable us to represent by a few well selected constants a wide range of experimental or observational data. In the latter case it serves not only for purposes of interpolation, but frequently suggests new physical concepts or statistical constants.

(Pearson, 1902, p. 266).

I think it is especially noteworthy – formula that is searched for, represents observations or measurements – i.e., *variables*. This fact is so obvious that consequences that follow from it are usually not thought through. The main problem related to use of observations and measurements is that they do not necessarily reflect the reality objectively; they are subjective interpretations of the world by the researcher. This is especially true in the situation when the observation – of external behavior, in psychology – is supposed to reflect operation of the hidden from direct observation construct, mental faculty. Thurstone acknowledged that externally similar behaviors can be based on internally different mechanisms; and he

was only interested in finding formulas that express regularities in the external behavior. Pearson essentially did the same; he assumed that with the help of correlations, it is possible to get closer to identification of different causes of external regularities. Pearson also did not aim at describing how these causes operate. He, similarly with Thurstone, was looking for identifying regularities (faculties in Thurstone's terms) in the observable cause → effect chains without claiming that unique cause, hidden from direct observation, is necessarily identified. This limitation for the aim of statistical analyses can be found in many of his works, as in the following passage, for instance:

We shall now assume that the sizes of this complex of organs are determined by a great variety of *independent* contributory causes, for example, magnitudes of other organs not in the complex, variations in environment, climate, nourishment, physical training, various ancestral influences, and innumerable other causes, which cannot be individually observed or their effects measured.

(Pearson, 1896, p. 262).

When Pearson correlated sizes of organs he was, thus, aware that mathematical formulas that reflect certain commonalities in the variation of two variables do not reflect unique roles of individual contributory causes; these causes determine the measured sizes in ways that are not known.

The general form of the question Pearson answered with statistical analyses can be formulated: What is the value of a certain variable when we know a value of another variable that is correlated to the first? An example of this kind of use was provided by Pearson when he reconstructed “the parts of an extinct race from a knowledge of a size of a few organs or bones, when complete measurements have been or can be made for an allied and still extant race.” (Pearson, 1899, p. 170).

Correlation, in this case, can be understood as a representation of some abstract cause which “makes” variables to covary. Thus, the same question can be reformulated: Is it possible to discover an abstract cause-like communality of different variables that is expressed as covariation? I think Pearson was very clear in understanding that correlation reflects covariations of appearances; the true underlying causes of covariation, the mechanisms that determine how the covariation emerges, cannot be known with statistical procedures – there are many independent causal agents operating, “which cannot be individually observed or their effects measured.” At the same time, he could interpret covariations between variables in non-mathematical terms; he interpreted them as reflecting common cause. For instance, he concluded on the basis of statistical analyses that fertility and fecundity are inherited characteristics (Pearson et al., 1899) – he, thus, suggested that some non-mathematical factor, inheritance, underlies the correlations he discovered.

Thurstone went a step further and suggested – it is possible to find formulas for expressing patterns of covariations; factor analysis identifies “faculties” or “abilities” that underlie correlation among several variables simultaneously. He was also clear that factor analysis expresses relationships between appearances; possible differences in internal mechanisms that may underlie externally similar behaviors are not reflected in the results of factor analysis.

## LIMITS ON THE QUESTIONS THAT CAN BE ANSWERED WITH THE HELP OF STATISTICAL DATA ANALYSIS PROCEDURES

### *Statistical theories reflect regularities only in appearances*

Pearson was fully aware of the limits of statistical theories. Theory that looks for mechanisms should be clearly distinguished from pure descriptions of regularities in superficial observations; statistical laws... “have nothing whatever to do with any physiological hypothesis” (Pearson, 1904, p. 55). According to him, “the statistical view of inheritance is not at basis a theory, but a description of observed facts, with which any physiological theory must be in accord” (Pearson, 1903–1904, p. 509). We know from the modern biological theory of inheritance, how correct he was: the statistical laws discovered by him had really nothing to do with the discovery of the structure of DNA, even though they may have directed biologists to look for possible substrate of inheritance. It is also noteworthy that, contrary to Pearson, after discovering the structure of DNA and explaining the biological mechanisms of inheritance, there was no need at all to check whether this theory of structure accords with Pearson's laws or not; these laws became irrelevant for the theory.

Thurstone was looking for discovering *mental faculties* with the help of the factor analysis. Similarly with Pearson, he did not assume that discovered faculties can be directly related to mental operations; he did not assume one-to-one correspondence between observed behaviors and mechanisms that underlie them:

The attitudes of people on a controversial social issue have been appraised by allocating each person to a point in a linear continuum as regards his favorable or unfavorable affect toward the psychological object. Some social scientists have objected because two individuals may have the same attitude score toward, say, pacifism, and yet be totally different in their backgrounds and in the causes of their similar social views. If such critics were consistent, they would object also to the statement that two men have identical incomes, for one of them earns when the other one steals. They should also object to the statement that two men are of the same height. The comparison should be held invalid because one of the men is fat and the other is thin.

(Thurstone, 1935, p. 47)

It was shown above that Thurstone was not asking how mental faculties operate; he was looking for just identification of abilities. There can be, thus, an ability to make money; and this ability is treated the same independently of whether the income is made by earning or by stealing. Statistical procedures used by Thurstone aim at discovering an ability to make income, for instance, but there would be no clue as to the mechanisms of the income-making. So, if there is a phenomenon like income in the world, perhaps factor analysis would be helpful to discover it.

There are, however, reasons to disagree partly with Thurstone's interpretation of this procedure. He suggested, for example, that incomes based on different sources should be considered to be different if social scientists were fully consistent; he disagreed with this idea. Essentially, Thurstone seems to assume that it is possible to isolate the phenomenon from the world and perhaps study it after isolation as a thing in itself. In the real world, however, a thing that exists completely isolated from the world would be unknowable in principle, because we know the world only being in relation with it. Income is by definition the amount of money or its equivalent received during a period of time. If we analyze the phenomenon of money, we discover that it is

a relational phenomenon. Money is a medium of exchange and unit of account; money, thus, is a phenomenon that mediates certain economic relationships. Outside the society, money ceases to be money; it becomes just a physical object. Societies determine relations toward money in much more complex ways than just in economic terms. For instance, in modern democratic societies it would be legally possible to confiscate money if the money turns out to be stolen; but there are no societies where money would be confiscated because it was earned. So, the incomes of two men, for one who earns and for the other who steals, are not the same indeed.

Thurstone would likely – and fairly – reject this critique by telling that “Every scientific construct limits itself to specified variables without any pretense to cover those aspects of a class of phenomena about which it has said nothing” (Thurstone, 1935, p. 47). By saying that a factor represents some isolated characteristic of the studied phenomenon, Thurstone retains consistency of his approach. And this is exactly where the weakness of statistical theories lies: these theories are about regularities in appearances with no necessary connection to the underlying mechanisms. Thurstone, similarly with Pearson, was fully aware of this limitation:

This volume is concerned with methods of discovering and identifying significant categories in psychology and in other social sciences. [...] It is the faith of all science that an unlimited number of phenomena can be comprehended in terms of a limited number of concepts or ideal constructs. [...] The constructs in terms of which natural phenomena are comprehended are man-made inventions. To discover a scientific law is merely to discover that a man-made scheme serves to unify, and thereby to simplify, comprehension of a certain class of natural phenomena. A scientific law is not to be thought of as having an independent existence which some scientist is fortunate to stumble upon. A scientific law is not a part of nature. It is only a way of comprehending nature. [...] While the ideal constructs of science do not imply physical reality, they do not deny the possibility of some degree of correspondence with physical reality. But this is a philosophical problem that is quite outside the domain of science.

(Thurstone, 1935, p. 44).

If biologists would have accepted this view, there would be no modern science of inheritance, for example. Modern biological theories do not look for “some degree of correspondence” between theories and physical reality; these theories aim at full correspondence.<sup>1</sup> In other words, scientific theories are not assumed to represent human-made generalizations based on covariations between appearances with no necessary connection to the reality that underlies these connections. On the contrary, the aim of sciences has become to understand exactly what Thurstone, Pearson and other statistical theorists did not aim at – to understand phenomena as they exist, not as they seem to us.

### ***Statistical theories of mechanisms depend on postulates that are not grounded and on conditions that are not satisfied***

Modern quantitative psychology may sometimes claim that its aims are similar to modern biology or physics – the discovery of the mechanisms that underlie the appearances, the observable behaviors.

<sup>1</sup>I am not going here into philosophical question whether such full correspondence can be known in principle. I agree that we can never be sure that our theories are correct (cf., e.g., Engels, 1996; Kant, 2007). I only suggest that theories can be in full correspondence with the physical reality even though we cannot demonstrate it.

Founders of the statistical theorizing denied such possibility by means of quantitative data analysis that is based on analysis of covariations between variables; perhaps they missed something fundamental that makes possible what they declared not to be possible? Perhaps it became possible to discover, for instance, by means of factor analysis the structure of mind as it is, not just as a man-made law that reflects only superficial covariations between observed events?

There are reasons to suggest that quantitative tools are not appropriate for this aim. Statistical data analysis procedures used in modern quantitative psychology are based on postulates that do not contradict the aims of statistical theorizing Thurstone, Pearson and their followers had. The same postulates, however, are incompatible with the aims of those who look for properties of mind as it is and not only for the generalizations that can be made about any kind of observations.

**Postulate of quantitative measurement.** Modern psychology must postulate that variables that are entered into analyses can be interpreted in terms of underlying mechanisms. Otherwise interpretation of the results of analyses in terms of those mechanisms would not be valid. Reasons to doubt whether this postulate is actually true, emerge already from the Pearson’s works. Namely, he extended statistical theorizing to characteristics that cannot be quantified (Pearson and Lee, 1900). For Thurstone and Pearson, it was not a problem that measured variables represent events with essentially unknown underlying causes because they did not aim at understanding those causes; they just looked for descriptions of statistical regularities in different observations. So for them even the question whether a variable represents something that can be quantified or not, was not an issue. But it must be one of the first problems to solve if physical or psychological reality is aimed to understand – what exactly is encoded in variables?

Psychology of today can be called pathological – many hypotheses are accepted as true without attempts being made to test them; the hypothesis that psychological attributes are quantitative is not tested in psychology of today (Michell, 2000). Worse, there are all reasons to suggest that attributes that are “measured” in psychology cannot be measured, because they are not quantitative (e.g., Essex and Smythe, 1999; Michell, 2010). Therefore, covariations between variables have no meaningful interpretation as to the underlying mechanisms in principle, because different levels of variables may denote qualitatively different phenomena. This alone would be sufficient for rejecting interpretations of quantitative analyses about underlying mechanisms. But there is more – as Thurstone also pointed out – externally similar behaviors can rely on internally different mechanisms. Thus even the same level on some variable may represent qualitatively different phenomena in different cases. It follows that under such circumstances no quantitative procedure can distinguish qualitatively different mechanisms that may underlie externally the same behavior – the variable that encodes behavior independently of differences in (psychic) mechanisms simply does not contain information about mechanism (Toomela, 2008).

If a researcher would be interested in distinguishing the psychological mechanisms of behavior, other procedures would be needed. A researcher would invent different methods to reveal differences in externally similar behaviors. For instance, in many situations it could be possible just to ask directly from the person a justification for his or her behavior. It is important that the methods that must

be created for discovering potential differences in externally similar behaviors are only qualitative because, as we saw, variables entered into quantitative data analyses lack the necessary information.

**Postulate of continuity.** There is another postulate, which underlies quantitative data analysis procedures. Thurstone, for instance, postulated:

The standard scores of all individuals in an unlimited number of abilities can be expressed, in first approximation, as linear functions of their standard scores in a limited number of abilities  
(Thurstone, 1935, p. 50).

So, there is a postulate that linear functions characterize relationships between the abilities (i.e., mental faculties) and individual acts of behavior. The main question that should be answered here is not only about postulate of linearity – the same problem would be related to non-linear relationships between variables – but about the postulate of continuity that is also made with this postulate of linearity. If it would turn out that some relationships between events are in essence qualitative then no factor analysis, or any other kind of quantitative data manipulation, can reveal those qualitative aspects of changes.

Often qualitative relationships hold between events. Lack of one nucleotide in a gene may be related to qualitatively different processes of protein synthesis, related to that gene. One extra chromosome does not just end up with more proteins; it ends up with qualitatively different pathologies, depending on the chromosome. It is also not meaningful to postulate a continuous quantitative series of events in the following continuum: one chromosome missing – the normal number of chromosomes – one extra chromosome in addition to the normal set.

**Postulate of correspondence between inter-individual and intra-individual levels of analysis.** In modern psychology, it is often assumed that *intra-individual* faculties can be revealed by studying *inter-individual* differences. This can be a major problem with all theories about individual attributes that are based on studies of differences between individuals: differences between individuals do not reflect distinctions inside individual minds (e.g., Lewin, 1935; Epstein, 1980; Toomela, in press-b).

Several quantitative scholars have provided substantial reasons why inter-individual differences cannot ground interpretations at the intra-individual level. They propose that quantitative analyses should be conducted with variables that encode intra-individual variability (e.g., Molenaar, 2004; Hamaker et al., 2005; Molenaar and Valsiner, 2005; Nesselroade et al., 2007; Boker et al., 2009). This approach, however, still assumes continuity and quantification. Before analyses of intra-individual variabilities, it must be demonstrated that the attributes, which are encoded as variables, can be quantified at all. The variables that are used in intra-individual analyses, however, are usually based on the scores of the same tests and inventories as used in inter-individual analyses. Therefore conducting analyses at the intra-individual level still cannot ground interpretations about attributes of mind. Another problem related to intra-individual quantitative approach follows from their assumption that data collected over time, reflect qualitatively the same processes. This assumption is in many cases wrong.

A person answering the same question repeatedly does not necessarily rely on the same mental operations – already second time the same question is asked, a person can answer in a certain way because he remembers answering the same question before. Data encoded as variables, again, do not reflect such qualitative changes of mental operations that underlie externally similar answers (see also Toomela, 2008, in press-b).

**Postulate of interpretability of covariations between variables.** Modern quantitative psychology also assumes that components of mental attributes can be discovered by analyzing covariations of variables. This postulate is questionable as well. As a rule, qualitatively different wholes emerge from the same elements in qualitatively different relationships. Quantitative data analysis, however, is not suitable for taking quality of relationships into account.

Human language, for instance, is based on units – words – that are composed from a limited number of sounds or letters in different relationships. We can take a series of events, words, and find perfect covariation between variables, sounds, in those events. Let us take, for instance, a series of events – words – this-shit-hits-pool-loop-polo. We create the following data-file from our observation of those six cases so that variables represent presence or absence of letters in each event/word:

	h	i	s	t	l	o	p
this	1	1	1	1	0	0	0
shit	1	1	1	1	0	0	0
hits	1	1	1	1	0	0	0
pool	0	0	0	0	1	2	1
loop	0	0	0	0	1	2	1
polo	0	0	0	0	1	2	1

We could make many different statistical analyses with those data and would not get any closer to understanding what is happening. Perhaps we would discover that all variables are perfectly correlated; we would discover that this data set can be perfectly “explained” by one factor, etc. Statistically, such results would be a perfect dream for a quantitative scientist. And yet all this would have no meaning. The data in the table show where the problem is – first three and last three qualitatively different cases are identical after quantification. Here we know that the cases are not identical; we do not know it when solving usual scientific problems. In any case, quantification of data into variables where the possibility of qualitatively different relationships between variables is ignored ends up with non-sense if qualitatively different wholes emerge from the same attributes encoded as variables.

Now it can be objected that such phenomena perhaps are not common. Nothing would be further from truth – the world around us provides massive amount of examples where the same elements in different relationships “cause” the emergence of qualitatively different wholes. The structure of DNA and its relationship to protein synthesis in a cell is an example; all chemical substances that are composed from the same elements in different relationships would be examples; different tools that can be made from the same material; different houses that can be built from the same stones; money that is earned and money that is stolen is also not the same, etc.

**Conditions that are not satisfied.** Over the last decade or two, an increasing number of substantial problems with statistical data analysis have been revealed. Some of them I have already mentioned above. But the list is definitely not complete with this. For instance, there are fundamental problems of interpreting variables that encode behavioral data (Toomela, 2008). The problems with interpretations emerge when (1) variables contain information about events at different levels of analysis; (2) wrong attribute from many that characterize the observed event is chosen for encoding into a variable; (3) measurement tool is not sufficiently sensitive (i.e., certain behaviors and mental phenomena underlying it exist but are not represented in the tools that are supposed to “measure” this mental phenomenon); (4) the studied phenomenon is absolutely necessary and therefore it does not vary; (5) variables represent variability that emerges because of the properties of the test or questionnaire rather than because the phenomenon really varies; and (6) the variable does not encode variability at the causally relevant range.

Results of the statistical data analyses cannot be interpreted in terms of the processes that underlie observed behaviors unless the meaning of the variables is clear. This condition is not satisfied in psychology. If the meaning of a variable is not clear then statistical data analysis may end up with demonstrating misleading dependencies or misleading independencies. Common textbooks of statistical data analysis all agree that discovery of a dependency between variables cannot be interpreted causally; these textbooks usually do not mention that absence of dependence also cannot be unequivocally interpreted – statistical independence of variables does not demonstrate absence of causal connections. If neither dependence nor independence can be unequivocally interpreted, the results of statistical data analyses cannot be taken as evidence for or against causal connections.

#### REMARK ON OTHER KIND OF QUESTIONS – QUESTIONS THAT CANNOT BE ANSWERED STATISTICALLY

Quantitative psychology asks questions about patterns of relationships between variables; the main question to be answered by such analyses is whether it is possible to identify some faculty, some ability, some cause that underlies observed behavior. In the discussion above I brought again and again examples from biology and chemistry, where the format of questions is different. In addition (not instead!) to asking whether a certain cause can be identified, questions are asked about the structure of the studied phenomena – what elements in which particular relationships underlie the emergence of a whole phenomenon that is aimed to understand. Quantitative data manipulations cannot reveal structure because in structures qualities of elements *and* qualities of relationships between elements determine the whole.

Altogether, there is not one epistemology that underlies science but two; one is looking for identification of cause → effect relationships and the other is aiming at structural-systemic description of the phenomena under study (see more on these two epistemologies, e.g., Toomela, 2009, 2010a, in press-a). These two epistemologies are rooted in philosophy. Next a very short description of the philosophical roots of these epistemologies is provided. It turns out that modern quantitative psychology is based on Cartesian–Humean epistemology whereas modern biology, chemistry, and several other

sciences are based on Aristotelian epistemology. Furthermore, psychology pretends to be like other sciences and superficially aims at understanding reality that underlies appearances. This, however, is impossible. We will see that in psychology there is a fundamental mismatch between questions asked and methods used to answer these questions.

## TWO EPISTEMOLOGIES

Two epistemologies that underlie different views on science are first of all distinct in their understanding of what is cause and causality. History of the notion of causality is complex; philosophers, and scientists have formulated a wide variety of theories of causation, each substantively different from the others. A nice summary of different definitions of causality can be found in Chambers’ *Cyclopaedia* (Chambers, 1728a,b). Under the entry “CAUSE” there is First Cause and Second Causes and many more. Under the “Causes in the School Philosophy,” there are: (1) Efficient causes; (2) Material causes; (3) Formal causes; (4) Final causes; (5) Exemplary causes. In the other way, again, “Causes” are distinguished into Physical, Natural, and Moral. Or yet another way, “Causes” are considered as Universal or Particular; Principal or Instrumental; Total or Partial; Univocal or Equivocal, etc. Two prominent views on causes and causality are relevant in the context of this paper.

### ARISTOTLE

Aristotle suggested that to know causes means to explain, to know “why” (e.g., Aristotle, 1941c, p. 240, Bk.II, 194<sup>b</sup>). This knowledge of causes is not just knowledge, it is scientific knowledge: “We think we have scientific knowledge when we know the cause (Aristotle, 1941b, p. 170, Bk.II, 94<sup>a</sup>). So, we can say that the aim of sciences is understanding what the causes of the studied phenomena are.

Aristotle distinguished four kinds of causes. In different works he described them from different perspectives. I am suggesting that Aristotelian philosophy of causality rooted structural-systemic epistemology that is followed by many sciences today. Shortly, according to this epistemology, scientific understanding implies description of the distinguishable elements, their specific relationships, the qualities that characterize the novel whole that emerges in the synthesis of those elements, and dynamic processes of the emergence of the whole (Toomela, 2009, 2010a). The connection of this kind of epistemology to Aristotelian becomes evident with the following quote:

All the causes now mentioned fall under four senses [...] some are cause as the substratum (e.g., *the parts*), others as the essence (*the whole, the synthesis, and the form*). The semen, the physician, the adviser, and in general the agent, are all sources of change or of rest. The remainder are causes as the end [...] (Aristotle, 1941a, p. 753, Bk.V, 1013<sup>b</sup>, my emphasis)

So, here we find concepts of parts, relationships or synthesis, and whole or form. We also find here another important notion for structural-systemic epistemology – emergence or causes of change. These four causes are called by tradition that was established long after Aristotle’s time, *material, formal, efficient, and final cause*, respectively.

## DESCARTES AND HUME

Two thousand years after Aristotle, we find considerably more limited views on causality. Instead of four complementary kinds of causes only one – efficient causality – is taken.

### *Descartes and efficient causality*

Descartes' view on causality is fundamentally different from the Aristotelian. First of all, he accepts only efficient causes and second, these efficient causes are very different from Aristotelian. For Descartes, *cause* is: independent, simple, universal, single, equal, similar, straight, etc.; *effect*, in turn, is: relative, dependent, composite, particular, many, unequal, dissimilar, oblique, etc. (cf. Descartes, 1985c). According to Descartes, effects can be deduced from causes in a series of steps. The cause–effect relationship, therefore, is unidirectional.

Another noteworthy idea in Cartesian epistemology was that “cause and effect are correlatives” (Descartes, 1985c, p. 22). In most cases, cause–effect relationships are essentially correlations, just covariations of events; there is, however, the First Cause – God – on whose power all causal relationships depend (cf. Descartes, 1985a,b). As God's plans cannot be known by less perfect humans (Descartes, 1985b), humans can know only correlations between appearances.

Cartesian description of cause contains terms and ideas that we also recognize in modern statistical data analysis. Here we find: *independent* and *dependent* variables; we find an idea of *linear* (or at least continuous) relationships – correlations; the idea that effects can be understood by knowing (efficient) causes – dependent variables or variability is statistically “*explained*,” etc. There are two noteworthy ideas more. First, the notion of “relationship” has only one meaning, that between cause and effect; no other kind of relationship is important. And second, there is no suggestion that qualitatively novel wholes emerge from the synthesis of parts. This idea is also similar to quantitative thinking in modern psychology. The overlap between Cartesian philosophy and modern quantitative epistemology, I suggest, is not just a coincidence; it reflects fundamental agreement between Cartesian causality and modern quantitative approaches to science.

### *Hume and efficient causality*

Slightly different approach to causality, even though similar to Cartesian in looking for efficient causality only, was taken by Hume. According to him,

Similar objects are always conjoined with similar. Of this we have experience. Suitably to this experience, therefore, we may define a cause to be *an object, followed by another, and where all the objects, similar to the first, are followed by objects similar to the second*. Or in other words, *where, if the first object had not been, the second never had existed*

(Hume, 2000, pp. 145–146).

So, cause is an object which appearance is related to the appearance of the other object. Space limitations do not allow going into detailed description of Hume's ideas. So I only mention them together with references to specific parts in his works where the corresponding ideas have been expressed by him. First, the relationship between causes and effects is characterized by contiguity

(Hume, 2000, p. 54). Second, causes relative to effects have priority in time; cause must precede the effect (Hume, 2000, p. 54). Third, the number of causes is smaller than the number of effects; therefore many observations of effects can be reduced to a few identified causes (Hume, 2000, p. 185). Fourth, the relationship between cause and effect reflects only relationships between appearances; no conclusion about reality that necessarily underlies the connection can be made (Hume, 1999, p. 136). Therefore conclusions about relations between causes and effects concern only matter of fact; they concern only the existence of objects or of their qualities (Hume, 2000, p. 65). Finally, according to Hume, the relationship between causes and effects is only probable (Hume, 1999, p. 115). The more often we observe an effect following the cause and the less often we observe effect not following the cause, the stronger is the impression of causality between the observed events (Hume, 2000, p. 105).

Taken together, it turns out that Humean epistemology is practically identical with the modern quantitative science – in both the succession of continuous events ground impressions about cause–effect relationships that can be observed with some probability; in both the impression of causal connection is perceived stronger when the proportion of observations that agree with one direction of events (from the supposed cause to the supposed effect) is higher than the proportion of observations that disagrees with this assumed direction of relationship; in both there can be no evidence that absolutely disagrees with some hypothetical causal relationship because cause–effect relationships can be observed only in degrees and not in necessary all-or-none relationship; and in both it is assumed that large number of observations can be “explained” by knowing small number of causes.

There is one interesting correspondence more between Humean epistemology and modern quantitative psychology. According to Hume, discovery of cause–effect relationships is based not on deductions or thinking but on “some instinct or mechanical tendency” (Hume, 1999, p. 130). The same can be said about quantitative science – the ways by which man-made causes (if to use Thurstone's words) are discovered, are highly mechanical. There are algorithms that are strictly followed in calculations of probabilities, effect sizes, and all other statistical descriptors of the variables in the analyses; there is no adjustment of each particular case of study to particular statistical calculations, for instance. In scientific inquiry, mechanization leads to dead end because it puts constraints on what can be understood in principle.

There is yet one point where the scholars who introduced statistical data analysis into sciences agreed with Hume but the modern researchers tend (at least implicitly) to disagree. It was already discussed above that both Pearson and Thurstone were fully aware that statistical theories are about appearances, about relationships between observed events; no conclusion can be made about the essence of the reasons why the statistical relationships between variables emerge. Hume had identical understanding of the state of affairs – efficient causality is about appearances and not about what he called “secret powers” that underlie the observed relationships:

It must certainly be allowed, that nature has kept us at a great distance from all her secrets, and has afforded us only the knowledge of a few superficial qualities of objects; while she conceals from us

those powers and principles, on which the influence of these objects entirely depends. [...] there is no known connection between the sensible qualities and the secret powers

(Hume, 1999, pp. 113–114).

Here modern quantitative psychology seems to disagree – on the basis of different kinds of statistical data analyses often conclusions are made about exactly those “secret powers.” Psychologists today attribute often the statistically “discovered” causes not just to man-made generalizations that leave an impression of causality but rather directly to “secret powers,” to mental attributes that are supposed to underlie the behaviors. It is ignored that behavior is not in one-to-one correspondence with psychic reality that underlies the behavior; externally identical behaviors may emerge from mentally qualitatively different operations and vice versa. So, all quantitative theories are only about appearances and not about underlying mechanisms because quantification of data into variables already excludes the information that is necessary for discovering the mechanisms that underlie observed covariations.

#### **Why only efficient causality?**

Aristotelian causality distinguished four complementary causes; Descartes and Hume, nevertheless, proposed only one. It is also important that neither Descartes nor Hume proposed entirely new concepts of causality; they took one Aristotelian cause out of his four. The reasons why they treated causality only in terms of efficient causes are relevant here.

**Descartes.** It was already discussed above that, according to Descartes, understanding of causality is about correlations between observed events; correlations do not imply necessity – every appearance can be correlated with every other appearance in principle. For Aristotle, causes were essentially constraints – in order to make a statue, bronze is used; there are many substances out of which it is not possible to make statues. If things have been made according to plan, then plan constrained the possible course of events; the result did not come out by accident or by chance but was constrained by plan before the event took place. Descartes, in order to be coherent with his philosophy, could not accept any kind of cause as a constraint.

Descartes believed in God, and not just some God, but God who is “infinite, eternal, immutable, omniscient, omnipotent [...] all the perfections which I could observe to be in God.” (Descartes, 1985a, p. 128). Therefore, logically, there can be no causes that are constraints because God has no constraints; God is omnipotent. God is the First Cause of everything that is. Effects follow from cause by necessity in principle because effects follow from God’s omnipotence. Humans, however, cannot know necessity that relates causes to effects; for them only knowledge about correlations is available:

When dealing with natural things we will, then, never derive any explanations from the purposes which God or nature may have had in view when creating them. For we should not be so arrogant as to suppose that we can share in God’s plans. We should, instead, consider him as the efficient cause of all things, and starting from the divine attributes which by God’s will we have some knowledge of, we shall see, with the aid of our God-given natural light, what

conclusions should be drawn concerning those effects which are apparent to our senses

(Descartes, 1985b, p. 202).

Taken together, humans can only know what is given for them through senses – appearances and correlations of them; they cannot know reasons that connect causes to effects because they are imperfect. Correlations do not allow going beyond observations of events; there is no way to know what are the reasons for observed correlations but one – God’s will.

**Hume.** For Hume, too, efficient causality was not related to necessity: “tis possible for all objects to become causes or effects to each other [...]” (Hume, 2000, p. 116). If there is no necessary relationship between events then it is not possible to know, why the events are related because it is actually not even possible to prove that the events are related essentially and not by accidents or by mistakes of observation. But his reasons for this view were different from Descartes’. Hume suggested that God is not knowable in principle and therefore the idea of God should not be taken into account in philosophy.

Hume suggested, similarly with Descartes, that humans have no access to knowledge beyond appearances; they cannot know why observed causes are related to observed effects. Human knowledge is actually even more limited – it is also not possible to be sure in discovered laws; the laws of nature can change and what we thought to be a cause may turn out to be the effect, or no connection between events would be discovered eventually (Hume, 1999, p. 115). The reasons of human limitations of understanding the world lied for Hume in limitations of the human (and animal) mind; world is not knowable beyond appearances because the mind is unable to go beyond appearances. Here Hume’s psychology becomes central for understanding his views. According to him, the mind works only on the principles of association:

[...] principles of association [...] To me, there appear to be only three principles of connection among ideas, namely, *Resemblance*, *Contiguity* in time or place, and *Cause or Effect*. [...] But the most usual species of connection among the different events, which enter into any narrative composition, is that of cause and effect;

(Hume, 1999, pp. 101–103).

So, the only operation available for mind is to form associations between observed events as they appear to us. If this would be the case, then Humean rejection of the possibility to have knowledge beyond senses – his proposition that only efficient causes as they appear to us can be known – would be well grounded. Humean psychology, however, was acceptable in his time, but not any more. The inability of associationism to be sufficient for explaining the human mind was established and grounded with empirical studies almost a century ago. Not only humans, but even apes were demonstrated to be able to think in a way that is not based on associations alone (Köhler, 1925; see also Koffka, 1935; Köhler, 1959; Vygotsky, 1982b). The idea that animal mind is based only on reflexes and conditioned reflexes, discovered by Pavlov (1927, 1951) was actually rejected by scholars from his own laboratory (Anokhin, 1975; Konstantinov et al., 1978).



## MODERN QUANTITATIVE PSYCHOLOGY – MIX OF TWO INCOMPATIBLE EPISTEMOLOGIES

Psychology today often aims at understanding structures that underlie observed behaviors. This aim is borrowed from that Aristotelian structural-systemic epistemology. Methods chosen for studies, however, are based on Cartesian–Humean cause–effect epistemology. Both philosophers who limited understanding of causality to efficient causality – Descartes and Hume – and scholars who introduced quantitative methods into sciences – Pearson and Thurstone – agreed that method of associating events by contiguity and covariation cannot ground interpretations in terms of underlying necessary reasons that connect observed causes to observed effects. They all also agreed that what is represented in observed associations between events or variables is subject to doubt. Interpretation of those associations can be only weaker or stronger depending on the relative frequency of events that correspond to certain idea of causality to the frequency of observations that contradict it. Laws discovered by such procedures are therefore not absolute but relative; laws cannot be refuted by observations that contradict it – in psychology effect sizes 1.0 are practically never observed; it is actually conveniently accepted that far-going conclusions can be made when 10–30% of data variability is statistically “explained.” It is ignored that in such situations substantial number of cases disagrees completely with the conclusions of the study. After conducting some “meta-analysis” it often turns out that the laws of association discovered in different studies contradict; and a new law can be proposed to replace those from the analyzed studies. It would not become a surprise when some meta-meta-analysis would yet lead to different generalization. Laws, in this epistemology, are not absolute; they can change without destroying the theory that is built from the collection of associative generalizations. Some philosophers would suggest that this kind of activity is not what science should do:

For it is an important postulate of scientific method that we should search for laws with an unlimited realm of validity. If we were to admit laws that are themselves subject to change, change could never be explained by laws. It would be the admission that change is simply miraculous. And it would be the end of scientific progress; for if unexpected observations were made, there would be no need to revise our theories: the *ad hoc* hypothesis that the laws have changed would “explain” everything

(Popper, 2002, p. 95).

### **Statistical methods in psychology are useless**

Taken together, there are reasons to suggest that quantitative methods are useless for psychology – IF the aim of psychology is to develop knowledge about mind, about “secret powers” that underlie observed behaviors. Such understanding would require qualitative approaches that allow distinguishing between externally similar behaviors based on internally different mental processes; and between externally different behaviors that are based on similar mechanisms.

Modern quantitative psychology is based on the epistemology where the questions are asked about efficient causality; explanation is reduced to identification of cause–effect relationships

between events. Such approach could be fully consistent if it would be accepted – as did Thurstone and Pearson – that discovery of such relationships cannot be connected to underlying structures in principle. Modern quantitative psychology, however, takes methods from Cartesian–Humean efficient causality epistemology and aims from incompatible with it Aristotelian-structural epistemology. Structural-systemic description of the studied phenomena cannot be based on quantitative methodology. The histories of biology or chemistry which are based on systemic-structural epistemology, also shows that majority of discoveries in these sciences have been made without statistical methods.

### **Statistical methods in psychology are inevitable**

The suggestion to reject quantitative methodology, I made, is conditional – IF the aims of studies would correspond to methods, quantitative methodology would turn out to be extremely valuable, almost inevitable ... for applied psychology. Now we need to turn the discussion upside-down. Instead of asking what cannot be accomplished with quantitative methods we ask, what it can bring to us? The world around us is constantly changing and always unique. How to live in the world of unique events? This would be impossible – in order to live, all life-forms must be able to react to future changes of the environment before these changes actually take place (Anokhin, 1978; Toomela, 2010a); foresight must be based on generalization and abstraction.

Coherent systemic-structural theories, as modern applications of physics, chemistry, and biology amply demonstrate, are extremely practical. But how to behave if the theory about underlying processes has not been created yet? Here quantitative methods become valuable: it is possible to create useful generalizations without knowing the processes that underlie the events. This was exactly what Thurstone, for instance, aimed at:

It is the faith of all science that an unlimited number of phenomena can be comprehended in terms of a limited number of concepts or ideal constructs. Without this faith no science could ever have any motivation. To deny this faith to affirm the primary chaos of nature and the consequent futility of scientific effort

(Thurstone, 1935, p. 44).

Thurstone, as we saw above, aimed explicitly and only at discovering ways to comprehend nature by describing regularities among observed events; these discoveries would be just man-made schemes, and yet they would help to manage otherwise unmanageable amount of information. A lot could be learned in this way – it would be possible to discover behaviors that should be avoided and behaviors that should be repeated in appropriate conditions – and all this without necessarily knowing, why. Until systemic-structural theories replace associative quantitative theories, psychology can create increasingly strong ground to applied uses of it. Quantitative science is inevitable for applied purposes until a theory about structures that underlie behavior is sufficiently developed for grounding applied uses. If, however, quantitative science continues to look for what it cannot find – the “secret powers” – then it ends up where Hume warned us not to go:

We are got into a fairy land, long ere we have reached the last steps of our theory [...]

(Hume, 1999, p. 142).

### Some notes on mathematical psychology in general

Mathematical psychology is not based exclusively on statistical methods. Perhaps non-statistical mathematical psychology is better suited for discovering the structure of mind? Indeed, from a certain perspective, it seems that mathematical psychology is doing well – there are fields of studies where mathematical psychology is prospering: foundational measurement theory, signal detection theory, decision theory, psychophysics, neural modeling, information processing approach, and learning theory (Townsend, 2008). Sometimes it almost seems that the only true science is based on mathematics; so Townsend suggests that psychology undergraduate training should change toward “solid-science” education and in order to do that, “The only practical solution I can espouse is for psychology departments to offer a *true scientific psychology track*, with mandatory courses in the sciences, *mathematics and statistics*” (Townsend, 2008, p. 275, my emphasis).

A small problem can be that achievements of mathematics, such as axiomatic measurement theory and computer-based, non-metric model fitting techniques, do not have an impact on psychology these “revolutions” deserve (Cliff, 1992). It might be that many problems will be solved with some developments in mathematics which, for instance, would explicate relationships between ways of describing randomness and ways of describing structure (Narens and Luce, 1993; Luce, 1999). It might be, however, that mathematics as such is inappropriate for answering questions psychology aims at answering. The most fundamental issue is not *how* mathematics should be applied in psychology but rather *whether* it can be applied for answering the core question of the science of psyche – what is mind? No development in any kind of measurement theory, for instance, will be helpful if psychological attributes cannot be measured in principle; there are strong reasons to suggest that they are not indeed (Valsiner, 2005; Trendler, 2009; Michell, 2010).

In order to proceed, a definition of mathematics is needed. According to Luce (1995, p. 2):

Mathematics studies structures and patterns described by systems of propositions relating aspects of entities in question. Deriving logically true statements from sets of assumed statements (often called axioms), uncovering symmetries and patterns, and evolving and understanding general structures are the concerns of mathematicians.

It is noteworthy that the term “structure” does not apply directly to the things and phenomena studied by physics, biology, psychology, or any other science. Rather, mathematics studies descriptions of objects and phenomena – systems of propositions – and “structure” refers to the system of descriptions; in that sense mathematics is an *abstract* science (Veblen and Young, 1910); it is a body of theorems deduced from a set of axioms (Veblen and Whitehead, 1932).

It is important that, as an abstract science, mathematics is based on assumptions, its “starting point” is

a set of *undefined elements and relations*, and a set of *unproven propositions involving them*; and from these all other propositions (theorems) are to be derived by the methods of formal logic (Veblen and Young, 1910, p. 1, my emphasis).

So, mathematics is a system of propositions that *begins* with a set of undefined assumptions, called axioms or postulates; and there are rules of deduction or a system of logic. Thus, mathematics defines *a priori* certain principles which are not derived from the studies of the world but attributed to it before studies. Mathematical description of the concrete real-world phenomena is successful only if the concrete system of things satisfies the fundamental assumptions of mathematics (Veblen and Young, 1910). Even though axioms can be postulated on the basis of scientific studies of the world and added to the basic set of abstract axioms, the abstract basis of mathematics is nevertheless determined before the studies. Taken together, it can be said that mathematics does not study the world but rather searches for events where the world corresponds to abstract mathematical principles – principles that cannot be proven or even defined.

Mathematics studies only formal aspects of the world. For Poincare, “mathematics is the art of giving the same name to different things” (Poincare, 1914, p. 34; also: “Mathematics teaches us, in fact, to combine like with like,” Poincare, 1905, p. 159). Similarities of things can be discovered by studying relationships:

Mathematicians do not study objects, but the relations between objects; to them it is a matter of indifference if these objects are replaced by others, provided that the relations do not change. Matter does not engage their attention, they are interested by form alone (Poincare, 1905, p. 20).

Thus, similarities of things discovered by mathematical studies are purely mathematical – things are similar if the mathematical relationships that describe them formally are similar. And the essence of these mathematical relationships, we saw above, is defined *a priori* and not derived from scientific studies.

Here are the reasons why not only statistics but any mathematical approach – if mathematics is defined as described above – is unable to reveal the structure of the things and phenomena studied; mathematics cannot in principle answer the questions structural-systemic science aims to answer – what is the studied thing or phenomenon? First, in the world, externally similar things and phenomena can be based on different underlying structures; for mathematics these structural differences do not exist. If, for instance, in a similar threatening situation one person reacts aggressively because he has made a conscious choice and the other impulsively, then for mathematics these two reactions are the same even though in the first case psychological structure included rational processes and in the other it did not. Mathematical prediction of future in such case cannot be very accurate because a person who chooses rationally to react aggressively in certain situations is also able to control his reactions whereas impulsive behavior is directed by the situation. This problem of mathematics has been recognized by mathematicians themselves; Poincare, for instance, suggested:

It is not enough that each elementary phenomenon should obey simple laws: *all those that we have to combine must obey the same law*; then only is the intervention of mathematics of any use. [...] It is therefore, thanks to the approximate homogeneity of the matter studied by physicists, that mathematical physics came into existence. In the natural sciences the following conditions are no longer to be found: – homogeneity, relative independence of remote parts, simplicity of the elementary fact; and that is why *the student*

*of natural science is compelled to have recourse to other modes of generalisation*

(Poincare, 1905, pp. 158–159, my emphasis).

In fact, Trendler (2009) proposed essentially the same reason why psychological attributes cannot be measured – in case of psychological attributes there are too many sources of systematic errors that cannot be controlled experimentally; in other words – psychological attributes are not independent but depend on each other. Therefore they cannot be measured.

Next, mathematics is a secondary science; the successful application of mathematics to the phenomena of the world depends on the experiments conducted in other sciences:

Experiment is the sole source of truth. It alone can teach us something new; it alone can give us certainty

(Poincare, 1905, p. 140).

Mathematics can help to organize the results of experiments; it can direct generalization; but it does not provide any new knowledge (Poincare, 1905). There are two different problems here. One problem is related to the essence of generalization. In mathematics, generalization is related to *relationships* between events, but in order to understand what a thing *is*, it is not sufficient to know with what else it can be related. In principle, there is no constraint on the number of other things with which any given thing in the world can be related; but what the thing *is*, is defined qualitatively and is fully constrained. A thing is what it is. Mathematical generalizations are not useful for discovering what things are. Another problem with mathematical generalizations can also be related to Poincare's discussion of the role of mathematics in sciences. According to him,

It is clear that any fact can be generalised in an infinite number of ways, and it is a question of choice. The choice can only be guided by considerations of simplicity

(Poincare, 1905, p. 146).

In case of phenomena that can be – in psychology they are – externally similar but yet based on qualitatively different psychic structures, the assumption of simplicity is not just wrong – it is fundamentally misleading. The assumption of simplicity forces scientists who rely on mathematics to ignore what should be studied and understood – complexity. Parenthetically, it should be mentioned here that mathematical models can be extremely complex; but they are fundamentally oversimplified if there is an assumption that externally similar events are all based on similar structures.

Third, mathematics can model only what is given by experiments and observations conducted in other sciences. It follows that mathematics is not able to provide any understanding of becoming, of emergence of qualitatively novel things. If these other sciences have described the emergence of novel qualities, this emergence can be modeled mathematically in principle. But again, what is modeled is not the novel thing or phenomenon itself – that model would be structural description of *the* thing or phenomenon – but relationships between events, i.e., what is modeled mathematically is always external to the thing itself. Here mathematicians perhaps could object and suggest that structural theory is also mathematical model. This suggestion, however, would require fundamental

redefinition of what mathematics is; because structural model based on studies of things, such as a model of atom, gene, wristwatch, or mind, is not based on a set of *a priori* given assumptions-axioms but rather on studies of the world. As Poincare suggested, science can be based on different kinds of generalizations and mathematical generalization that fits so well into physics, is considerably less useful in other sciences. I will discuss briefly the methods of scientific generalization in the next section.

Altogether, there are reasons to suggest that mathematics is not appropriate tool if the aim of science is to understand what the studied thing or phenomenon is. For mathematical psychologist, naturally, mathematics is almost *the* most important tool for science:

Mathematical psychology has arguably accelerated the evolution of psychology and allied disciplines into rigorous sciences many times over their likely progress in its absence. Let's nurture and strengthen it

(Townsend, 2008, p. 279).

I suggest that mathematics has actually had the opposite role for psychology – it has oversimplified theories, blinded scientists, and directed their attention to the study of relationships between things and phenomena instead of guiding them to study what these things and phenomena are. Physics has been most successful not where something can be exactly calculated but where the theory has defined what the things are, atoms, for instance. Yes, mathematics is useful tool here and there – as it can be also for psychology for grounding practical decisions – but no machine has been ever built on the basis of mathematical formulas alone whereas many of them have been constructed completely without any aid from mathematics. The same can be said about biology – it is very powerful science not because of applications of mathematics in some peripheral matters of biology but because of the theories about what are cells, components of cells, organs, organisms, etc. Perhaps what should be “nurtured and strengthened” in psychology is not mathematical psychology but studies that aim to understand what mind is.

## IF NOT MATHEMATICS, THEN WHAT?

So, mathematics is a useful tool for generalizations about relationships between events. The value of mathematics, though, is not the same in all sciences. Poincare, for instance, suggested that mathematics is very useful tool for generalization of the results of experiments in physics:

It might be asked, why in physical science generalisation so readily takes the mathematical form. The reason is now easy to see. It is not only because we have to express numerical laws; it is because the observable phenomenon is due to the superposition of a large number of elementary phenomena which are all similar to each other; and in this way differential equations are quite naturally introduced

(Poincare, 1905, p. 158).

He also suggested, as was discussed above, that mathematical generalization is appropriate only in cases when the matter studied by the scientists is homogeneous; when parts are relatively independent and elementary facts are simple. These conditions are not

met in psychology. Therefore another way for generalization must be found. Another way for scientific generalization is a special kind of experiment.

Usually it is thought that there is one kind of experiments. This kind is Cartesian–Humean; the question answered in the experiment is whether certain event is or is not an efficient cause of another event. In order to answer this question, the artificial situation is created where, ideally, all conditions are kept equal but one that is manipulated or “controlled.” If the expected effect follows when manipulated event is present and does not follow when the manipulated event is absent, then it is concluded that the cause of the event has been identified.

There is, however, another kind of experiment that, to the best of my knowledge, was first brought into the theory of scientific experimentation by Engels (even though in several respects similar idea can be found already in Aristotle’s works; cf. Aristotle, 1941a, p. 690, Bk.I, 981<sup>a</sup>–981<sup>b</sup>). Engels discussed the role of induction in scientific discoveries and proposed that there is much more powerful way for scientific proofs than induction:

A striking example of how little induction can claim to be the sole or even the predominant form of scientific discovery occurs in thermodynamics: the steam-engine provided the most striking proof that one can impart heat and obtain mechanical motion. 100,000 steam-engines did not prove this more than *one*, but only more and more forced the physicists into the necessity of explaining it. [...] The empiricism of observation alone can never adequately prove necessity. *Post hoc* but not *propter hoc*. [...] But the proof of necessity lies in human activity, in experiment, in work: if I am able to *make* the *post hoc*, it becomes identical with the *propter hoc*.

(Engels, 1987, pp. 509–510)

This kind of experiment that follows from the principles outlined by Engels, I have called “constructive” (Toomela, in press-a). In constructive experiments it is attempted to create the thing or phenomenon that is studied. If the phenomenon or thing can be constructed on the basis of knowledge about hypothetical elements and specific relationships between them, the experiment has provided corroborating evidence for the theory. Here the result of the experiment – constructed thing or phenomenon – follows from theory. It is important that there is no logical necessity that a whole with certain emergent properties must emerge when theoretically defined elements are put into certain relationships. Instead of logical deduction, the necessity is proven in the real construction of the phenomenon that is attempted to understand. Mathematics derives logically true statements from assumptions that cannot be proven. It is important that the truth of logical derivations depends fully on the truth of the assumptions. If even one assumption cannot be proven, there is no proof possible for the scientific theory as a whole as well. In constructive experiments, on the other hand, the proof is obtained by actual construction of the studied thing.

## REFERENCES

- Anokhin, P. K. (1975). Oчерки по физиологии функциональных систем. Moscow: Medicina.
- Anokhin, P. K. (1978). “Operezhajuscheje otrazhenije deistvitel’nosti,”

(Anticipating reflection of actuality. In Russian. Originally published in 1962.). in *Filosofskije aspekty teorii funkcional’noi sistemy*, eds F. V. Konstantinov, B. F. Lomov, V. B. Schvyrkov and P. K. Anokhin,

Such actual construction does not contain any assumptions that cannot be proven – these assumptions are proven by the result of the experiment.

Constructive experiments can be found in different fields of science. Atomic theory is well corroborated by the construction of the nuclear bombs and reactors; chemical theories are corroborated with the synthesis of new molecules. Just now biologists have reported a major breakthrough in constructive experiments on the cells; a bacterial cell with chemically synthesized genome has been created (Gibson et al., 2010). There are also examples of constructive experiments in psychology. Neuropsychological rehabilitation based on Luria’s theory is grounded on structural-systemic theory; numerous psychological functions have been artificially created with special educational programs that were designed on the basis of theories about the elements and relationships of elementary psychological processes of those complex functions (Luria, 1948; Tsvetkova, 1985).

Taken together, mathematics is not useful for discovering what things are. For such discoveries, observations and analytic experiments should be combined with constructive experiments.

## CONCLUSIONS

Science begins with the question, what do I want to know? Science becomes science, however, only when this question is justified *and* the appropriate methodology is chosen for answering the research question. Research question should precede the other questions; methods should be chosen according to the research question and not vice versa. Modern quantitative psychology, though, has accepted method as primary and research questions are adjusted to the methods. It would not be a problem if methods would fit the questions about the studied phenomena; but they do not. The crucial question that needs to be asked, is – do the answers to the questions what, why, and how I want to know, make a coherent theoretically justified whole? All psychology that aims at understanding the structure of mind with any kind of mathematical tools has to admit that the methods do not correspond to the study questions.

For understanding thinking in modern quantitative psychology, two epistemologies should be distinguished: structural-systemic that is based on Aristotelian thinking; and associative-quantitative that is based on Cartesian–Humean thinking. The first aims at understanding the structure that underlies the processes leading to events observed in the world, the second looks for identification of apparent cause–effect relationships between the events with no claim made about processes that underlie the appearances.

Quantitative methods are useful for generalizations about the relationships between things and events. What the studied things or phenomena are cannot be revealed by such methods. Structural-systemic science, which aims at understanding structures, relies on qualitative methodology that includes, in addition to the observations and analytic experiments, constructive experiments.

- Izbrannyje trudy. (Moscow: Nauka), pp. 7–26.
- Aristotle. (1941a). “Metaphysics,” (Metaphysica.). in *The Basic Works of Aristotle*, ed. R. McKeon (New York: Random House), pp. 681–926.

- Aristotle. (1941b). “Organon,” in *The Basic Works of Aristotle*, ed. R. McKeon (New York: Random House), pp. 7–212.
- Aristotle. (1941c). “Physics,” (Physica.). in *The Basic Works of Aristotle*, ed. R.

- McKeon (New York: Random House), pp. 213–394.
- Boker, S. M., Molenaar, P. C. M., and Nesselroade, J. R. (2009). Issues in intraindividual variability: individual differences in equilibria and dynamics over multiple time scales. *Psychol. Aging* 24, 858–862.
- Chambers, E. (1728a). *Cyclopaedia, or, An universal dictionary of arts and sciences*. In two volumes. Volume the first. London: Ephraim Chambers Retrieved 21 July 2009 from <http://digicoll.library.wisc.edu/HistSciTech/subcollections/CyclopaediaAbout.html>
- Chambers, E. (1728b). *Cyclopaedia, or, An universal dictionary of arts and sciences*. In two volumes. Volume the Second. London: Ephraim Chambers Retrieved 21 July 2009 from <http://digicoll.library.wisc.edu/HistSciTech/subcollections/CyclopaediaAbout.html>
- Cliff, N. (1992). Abstract measurement theory and the revolution that never happened. *Psychol. Sci.*, 3, 186–190.
- Descartes, R. (1985a). “Discourse on the method of rightly conducting one’s reason and seeking the truth in the sciences,” (Originally published in 1637) in *The Philosophical Writings of Descartes, Vol. 1*, eds J. Cottingham, R. Stoothoff and D. Murdoch (New York: Cambridge University Press), pp. 111–151.
- Descartes, R. (1985b). “Principles of philosophy,” (Originally published in 1644 in Latin and in 1647 in French) in *The Philosophical Writings of Descartes, Vol. 1*, eds J. Cottingham, R. Stoothoff and D. Murdoch (New York: Cambridge University Press), pp. 179–291.
- Descartes, R. (1985c). “Rules for the direction of the mind,” (Originally written in about 1628, published in 1684) in *The Philosophical Writings of Descartes, Vol. 1*, eds J. Cottingham, R. Stoothoff and D. Murdoch (New York: Cambridge University Press), pp. 9–78.
- Engels, F. (1987). “Dialectics of nature,” in *Collected Works, Vol. 25*, eds K. Marx and F. Engels (New York: International Publishers), pp. 313–590.
- Engels, F. (1996). *Ludwig Feuerbach and the End of Classical German Philosophy* (Originally published in 1888). Beijing: Foreign Language Press.
- Epstein, S. (1980). The stability of behavior. II. Implications for psychological research. *Am. Psychol.* 35, 790–806.
- Essex, C., and Smythe, W. E. (1999). Between numbers and notions. A critique of psychological measurement. *Theory Psychol.* 9, 739–767.
- Gibson, D. G., Glass, J. I., Lartigues, C., Noskov, V. N., Chuang, R.-Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z.-Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, III, C. A., Smith, H. O., and Venter, J. C. (2010). Creation of a bacterial cell controlled by a chemically synthesized genome. *Science* 329, 52–56.
- Gigerenzer, G. (1991). From tools to theories: a heuristic of discovery in cognitive psychology. *Psychol. Rev.* 98, 254–267.
- Gigerenzer, G. (1993). “The superego, the ego, and the id in statistical reasoning,” in *A Handbook for Data Analysis in the Behavioral Sciences: Methodological Issues*, eds G. Keren and C. Lewis (Hillsdale, NJ: Lawrence Erlbaum Associates), pp. 311–339.
- Hamaker, E. L., Dolan, C. V., and Molenaar, P. C. M. (2005). Statistical modeling of the individual: rationale and application of multivariate stationary time series analysis. *Multivariate Behav. Res.* 40, 207–233.
- Hume, D. (1999). “An enquiry concerning human understanding,” (Originally published in 1748) in *An Enquiry Concerning Human Understanding*, ed. T. L. Beauchamp (Oxford: Oxford University Press), pp. 81–211.
- Hume, D. (2000). “A treatise of human nature,” (Originally published in 1739–1740) in *A Treatise of Human Nature*, eds D. F. Norton and M. J. Norton (Oxford: Oxford University Press), pp. 1–418.
- Kant, I. (2007). “Critique of pure reason,” (Originally published in 1787) in *Critique of Pure Reason*. Immanuel Kant. Revised 2nd Edn., ed. N. K. Smith (New York: Palgrave MacMillan).
- Koffka, K. (1935). *Principles of Gestalt Psychology*. London: Routledge & Kegan Paul.
- Köhler, W. (1925). *The Mentality of Apes*. New York: Harcourt, Brace & Co.
- Köhler, W. (1959). *Gestalt Psychology. An Introduction to New Concepts in Modern Psychology*. New York: Mentor Books.
- Konstantinov, F. K., Lomov, B. F., and Shvyrykov, B. V. (eds). (1978). *P. K. Anokhin. Izbrannyye trudy. Filosofskije aspekty teorii funktsional’noi sistemy*. Moscow: Nauka.
- Lewin, K. (1935). *A Dynamic Theory of Personality. Selected Papers*. New York: McGraw-Hill.
- Luce, R. D. (1995). Four tensions concerning mathematical modeling in psychology. *Annu. Rev. Psychol.* 46, 1–26.
- Luce, R. D. (1999). Where is mathematical modeling in psychology headed? *Theory Psychol.* 9, 723–737.
- Luria, A. R. (1948). *Vosstanovleniye funktsii mozga posle vojennoi travmy*. (Restoration of Brain Functions After War Trauma. In Russian). Moscow: Izdatel’stvo Akademii Medicinskih Nauk SSSR.
- Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory Psychol.* 10, 639–667.
- Michell, J. (2010). “The quantity/quality interchange: a blind spot on the highway of science,” in *Methodological Thinking in Psychology: 60 Years Gone Astray?* eds A. Toomela and J. Valsiner (Charlotte, NC: Information Age Publishing).
- Molenaar, P. C. M. (2004). A Manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Measurement* 2, 201–218.
- Molenaar, P. C. M., and Valsiner, J. (2005). How generalization works through the single case: a simple idiographic process analysis of an individual psychotherapy. *Int. J. Idiographic Sci.* Article 1. Retrieved October 25 2005 from <http://www.valsiner.com/articles/molenvals.htm>
- Narens, L., and Luce, R. D. (1993). Further comments on the “nonrevolution” arising from axiomatic measurement theory. *Psychol. Sci.* 4, 127–130.
- Nesselroade, J. R., Gerstorff, D., Hardy, S. A., and Ram, N. (2007). Idiographic filters for psychological constructs. *Measurement* 5, 217–235.
- Pavlov, I. P. (1927). *Lekcii o rabote bol’shikh polusharii golovnogo mozga*. (Lectures on the work of the hemispheres of the brain.). Moscow: Gosudarstvennoje Izdatel’stvo.
- Pavlov, I. P. (1951). *Dvatcatilet’ nii opyt ob’ektivnogo izuchenija vyshei nervnoi dejatel’nosti (povedeniya zhivotnykh)*. Moscow: Medgiz.
- Pearson, K. (1896). Mathematical contributions to the theory of evolution. III. Regression, heredity, and panmixia. *Philos. Trans. R. Soc. Lond. A* 187, 253–318.
- Pearson, K. (1899). Mathematical contributions to the theory of evolution. V. On the reconstruction of the stature of prehistoric races. *Philos. Trans. R. Soc. Lond.* 192, 169–244.
- Pearson, K. (1902). On the systematic fitting of curves to observations and measurements. *Biometrika*, 1, 265–303.
- Pearson, K. (1903–1904). Mathematical contributions to the theory of evolution. XII. On a generalised theory of alternative inheritance, with special reference to Mendel’s laws. *Proc. R. Soc. Lond.* 72, 505–509.
- Pearson, K. (1904). Mathematical contributions to the theory of evolution. XII. On a generalised theory of alternative inheritance, with special reference to Mendel’s laws. *Philos. Trans. R. Soc. Lond. A* 203, 53–86.
- Pearson, K., and Lee, A. (1900). Mathematical contributions to the theory of evolution. VIII. On the inheritance of characters not capable of exact quantitative measurement. *Philos. Trans. R. Soc. Lond. A* 195, 79–150.
- Pearson, K., Lee, A., and Bramley-Moore, L. (1899). Mathematical contributions to the theory of evolution. VI. Genetic (reproductive) selection: Inheritance of fertility in man, and of fecundity in thoroughbred racehorses. *Philos. Trans. R. Soc. Lond. A* 192, 257–330.
- Poincaré, H. (1905). *Science and Hypothesis*. London: Walter Scott Publishing.
- Poincaré, H. (1914). *Science and Method*. New York: Cosimo.
- Popper, K. (2002). *The Poverty of Historicism*. (First English edition published in 1957). London: Routledge Classics.
- Thurstone, L. L. (1935). *The Vectors of Mind: Multiple-Factor Analysis for the Isolation of Primary Traits*. Chicago: University of Chicago Press.
- Thurstone, L. L. (1948). Psychological implications of factor analysis. *Am. Psychol.* 3, 402–408.
- Toomela, A. (2008). Variables in psychology: a critique of quantitative psychology. *Integr. Psychol. Behav. Sci.* 42, 245–265.
- Toomela, A. (2009). “How methodology became a toolbox - and how it escapes from that box,” in *Dynamic Process Methodology in the Social and Developmental Sciences*, eds J. Valsiner, P. Molenaar, M. Lyra and N. Chaudhary (New York: Springer), pp. 45–66.
- Toomela, A. (2010a). Biological roots of foresight and mental time travel. *Integr. Psychol. Behav. Sci.* 44, 97–125.
- Toomela, A. (2010b). “Modern mainstream psychology is the best? Noncumulative, historically blind, fragmented, atheoretical,” in *Methodological Thinking in Psychology: 60 Years Gone Astray?* eds A. Toomela and J. Valsiner (Charlotte, NC: Information Age Publishing), pp. 1–26.
- Toomela, A. (in press-a). “Guesses on the future of cultural psychology: past, present, and past,” in *Oxford Handbook of Culture and Psychology*, ed. J. Valsiner (Oxford: Oxford University Press).
- Toomela, A. (in press-b). “Methodology of idiographic science: limits of single-case studies and the role of typology,” in *Yearbook of Idiographic Science*, Vol. 2, eds S. Salvatore, J. Valsiner,

- A. Gennaro, and J. B. Travers Simon (Rome: Firera & Liuzzo Group).
- Townsend, J. T. (2008). Mathematical psychology: prospects for the 21st century: A guest editorial. *J. Math. Psychol.* 52, 269–280.
- Trendler, G. (2009). Measurement theory, psychology and the revolution that cannot happen. *Theory Psychol.* 19, 579–599.
- Tsvetkova, L. S. (1985). *Neiropsikhologičeskaja reabilitatsija bol'nykh. Rech i intellektual'naja dejatel'nost.* (Neuropsychological rehabilitation of a sick person. Speech and intellectual activity. In Russian.). Moscow: Izdatel'stvo Moskovskogo Universiteta.
- Valsiner, J. (2005). Transformations and flexible forms: where qualitative psychology begins. *Qual. Res. Psychol.* 4, 39–57.
- Veblen, O., and Whitehead, J. H. C. (1932). *The Foundations of Differential Geometry. Cambridge Tracts in Mathematics and Mathematical Physics.* Cambridge: Cambridge University Press, 29.
- Veblen, O., and Young, J. W. (1910). *Projective Geometry, Vol. I.* New York: Ginn & Co.
- Vygotsky, L. S. (1982a). "Istoricheski smysl psikhologičeskogo krizisa. Metodologičeskoje issledovanije," (Historical meaning of the crisis in psychology. A methodological study. Originally written in 1927; First published in 1982) in *Sobranije sochinenii. Tom 1. Voprosy teorii i istorii psikhologii.* eds A. R. Luria and M. G. Jaroshevskii (Moscow: Pedagogika), pp. 291–436.
- Vygotsky, L. S. (1982b). "Problema razvitija v strukturnoi psikhologii. Kritičeskoje issledovanije," (Problem of development in structural psychology. A critical study. Originally published in 1934) in *Sobranije sochinenii. Tom 1. Voprosy teorii i istorii psikhologii.* eds A. R. Luria and M. G. Jaroshevskii (Moscow: Pedagogika), pp. 238–290.
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 23 February 2010; paper pending published: 26 March 2010; accepted: 25 June 2010; published online: 30 July 2010.

Citation: Toomela A (2010) Quantitative methods in psychology: inevitable and useless. *Front. Psychology* 1:29. doi: 10.3389/fpsyg.2010.00029

This article was submitted to *Frontiers in Quantitative Psychology and Measurement*, a specialty of *Frontiers in Psychology*. Copyright © 2010 Toomela. This is an open-access article subject to an exclusive license agreement between the authors and the *Frontiers Research Foundation*, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.