Integr Psych Behav (2008) 42:245–265
DOI 10.1007/s12124-008-9059-6

REGULAR ARTICLE

Variables in Psychology: A Critique of Quantitative Psychology

Aaro Toomela

Published online: 5 June 2008

Abstract Mind is hidden from direct observation; it can be studied only by observing behavior. Variables encode information about behaviors. There is no one-to-one correspondence between behaviors and mental events underlying the behaviors, however. In order to understand mind it would be necessary to understand exactly what information is represented in variables. This aim cannot be reached after variables are already encoded. Therefore, statistical data analysis can be very misleading in studies aimed at understanding mind that underlies behavior. In this article different kinds of information that can be represented in variables are described. It is shown how informational ambiguity of variables leads to problems of theoretically meaningful interpretation of the results of statistical data analysis procedures in terms of hidden mental processes. Reasons are provided why presence of dependence between variables does not imply causal relationship between events represented by variables and absence of dependence between variables cannot rule out the causal dependence of events represented by variables. It is concluded that variable-psychology has a very limited range of application for the development of a theory of mind—psychology.

Keywords Variable · Statistical data analysis · Causality · Psychology · Research methods

Psychology is mostly aiming to discover causal relationships between environmental and mental events or trying to reveal hidden structure of the mind. First of these aims is related to the experimental stream of psychology where treatment effects are statistically compared. The second aim is related to the study of individual differences (Cronbach 1957). Differently from the psychology of the pre-WWII era (cf. Toomela 2007a, 2008), the studies of mind in modern mainstream psychology rely heavily, often exclusively, on the interpretation of information

A. Toomela (✉)
Department of Psychology, Tallinn University, Narva mnt. 25, Tallinn 10120, Estonia
e-mail: aaro.toomela@ut.ee

encoded in variables. Variables are used in both streams of psychology, experimental and individual differences. In both streams, also, the variables are usually analyzed with some statistical data analysis procedures.

The two streams of psychology are qualitatively different in the way the data are searched for. Experimental psychology manipulates with study conditions and collects data related to different treatment effects whereas individual differences psychology uses data collected without directly interfering with the study situation. In the way in which the data are interpreted, however, the two streams of psychology are similar. Both of them search with the help of statistical data analysis procedures for patterns of covariations or dependencies between variables. It is obvious in all cases where data analysis procedures reveal patterns of correlations/covariations (correlation coefficients, multiple regression analysis, factor analysis, structural equation modeling procedures, etc.). Data analysis procedures where dependencies between nominally encoded variables are searched for also belong to this group. So, for example, contingency tables can be represented in terms of correlation matrix so that chi-square turns out to provide the same information as canonical correlation (Dunlap et al. 2000).

Superficially, the statistical analysis of treatment effects ("independent" variables) where differently treated groups are compared seems to rely on a kind of statistical analysis different from correlational procedures. Comparison of mean levels between groups, for example, may seem for researchers to be different from correlational procedures. The similarity between the procedures becomes obvious when we understand that comparison of mean levels of a variable between two groups can at the same time be understood as a correlation between two variables, one representing the group and the other representing the measured construct. So, in case of two groups, difference between the groups on some variable can be understood as a correlation between the dichotomous group belongingness variable and the measured variable. In case of mean models with more than two groups or in case of factorial designs the analysis can always be broken down into a set of comparisons between two groups (Keppel 1991).

So, the data analysis procedures used in experimental and individual differences psychology are basically identical. The relationship between the two streams of thought is actually deeper. Experimental psychology searches explicitly for causal information. Individual differences psychology has a similar aim implicitly. In psychology of intelligence and personality, for example, factor analysis is used for revealing the structure of mind. It is assumed that factor structure of relationships between variables is underpinned by hidden structural characteristics of mind. Neuroticism items in a personality questionnaire, for example, load on one factor and Extraversion items on another because there are hidden Neuroticism and Extraversion dimensions of mind that cause systematic response differences to questionnaire items. Both streams of psychology, therefore, aim to discover causal relationships and give information about the structure of mind.

Usually statistical data analysis procedures are used with the understanding represented in most statistical data analysis textbooks that no causality can be inferred from these procedures. In modern philosophy causality is understood as a certain relationship between events (e.g., Campbell et al. 2007; Russo and

Williamson 2007; Sosa and Tooley 1993). There is an increasingly influential school of thought in psychology according to which it is possible to reveal causality with the help of statistical data analysis of information about events encoded into variables (e.g., Pearl 2000; Sloman 2005; Spirtes et al. 2000). If two variables are correlated, then difficulties for inferring causality are related to the problem that any two variables, say A and B, can be correlated because A causes B, B causes A, or both A and B are influenced by a third variable X. Those who claim that causality can be revealed through statistical data analysis believe that it is possible because certain patterns of probabilistic dependencies imply unambiguous causal dependencies. It is important that independently of whether statistical data analysis is believed to reveal true causality or not, it is unequivocally assumed that the fact of presence or absence of a reliable relationship—statistical/probabilistic dependence or independence of variables—is meaningful. Even though dependence of variables can have different causal interpretations, independence is understood as an unambiguous support to the inference that events represented by variables are also causally independent or not related.

The aim of a science of psychology is not the study of relationships between variables—this is only a tool—but the study of relationships between events represented as variables. It is possible to find an enormous amount of discussion on the possibilities and applications of more and more complex statistical data analysis techniques in the modern mainstream psychology. There is one fundamental issue, however, that is relatively (but not completely, see, e.g., Blumer 1956; Borsboom et al. 2004; Essex and Smythe 1999; Michell 2000, 2003) ignored in this context—the ontology and epistemology of a variable. Without clearly understanding what information is represented in a variable—how exactly variables represent events—it is not possible in principle to meaningfully interpret events and relationships between events on the basis of any procedure of analysis of variables.

In this paper I suggest that, theoretically, variables may contain different kinds of information about the event they represent. In case of causally related events, what kind of information about events is represented in variables determines whether relationship or dependence is discovered or not with the statistical data analysis procedures. I will show that variables representing events can be statistically independent even when underlying the variables events are causally related. Therefore, statistical data analysis procedures cannot always provide theoretically meaningful information about causal relationships between events not only because statistical dependence of variables can have different causal interpretation ($A \rightarrow B$; $B \rightarrow A$; $X \rightarrow A$ and B) but also because the statistical independence of variables does not imply independence of events represented by variables. Statistical data analysis procedures are sometimes used to discover true causal relationships between events (e.g., Pearl 2000); in other cases it is assumed that (Sloman 2005; Spirtes et al. 2000) at least some clues about presence or absence of causal relationships can be found through the analyses. If neither the statistical dependence nor the statistical independence of variables is unambiguously informative about the presence or absence of causal relationships between events studied, then it must be concluded that no causal inferences should be based on statistical data analysis procedures alone.

Information Encoded in Variables

In case of directly observable events it is sometimes possible to define the content of a variable unambiguously. Psychology studies phenomena that are hidden from direct observation. In this situation, definition of the content of a variable becomes a problem to be faced by the researchers. Behaviors or results of behaviors must be *interpreted* in order to find the relationship between a behavioral event and the hidden mental event that underlies the behavior. So, there are not one but two steps from a studied event to a variable. First there is a hidden mental process that manifests in a behavior. Second, this behavior is encoded as a variable. The major problem is that one and the same variable may contain more than one kind of information about an event at the same time. What kinds of information can be represented in variables is discussed next.

Qualitative Information—Yes or No

One kind of information encoded in all variables is about whether a certain thing or phenomenon exists or not. If a psychologist would only need to encode directly observable behaviors into a variable, this task would not be problematic. For a behaviorist, therefore, there would be no difficulties with understanding a variable. But in modern mainstream psychology variables are usually supposed to encode information about hidden mental processes that manifest in behaviors. Major conceptual difficulties arise because of this fact.

First, *it is not obvious whether a phenomenon encoded as 'existing' does exist in reality or not* (see also Borsboom et al. 2004). It is possible to find regularities in behaviors that do not stem from a hidden mental cause or structure. A phenomenon can instead be created, for example, when group data are analyzed. Galton's works with 'composite photography' (e.g., Galton 1878) can be used as an example. Galton created photos where types of appearances were created by exposing a number of individual portraits of chosen groups of people on a photographic plate to obtain collective representations. In a composite picture not an individual but an average image where individual physiognomic qualities vanish and common to the group characteristics accentuate is represented. Composite picture shows a face that exists only as a picture, no human face of this kind does exist in reality; it is a result of construction, a prototype. A factor analysis, analogously, may create a phenomenon—such as *g*-factor, neuroticism, extraversion, collectivism, etc.—that does not necessarily characterize mental organization of any person taken individually. This is a general principle that results of studies based on data about variation between cases cannot be extended to individual cases because psychological processes are not ergodic (Molenaar 2004; Molenaar and Valsiner 2005, see also Sohn 1999). In other words, if one abstracts from individual differences, there is no logical way back from these generalities to the individual case (Lewin 1935, 1997b). But mind is not a characteristic of a group, it is an individual phenomenon. Correspondingly, it is possible that a variable encodes information about a phenomenon that does not exist in persons studied. To search for 'causal' relations between variables that encode nonexisting at that level of analysis phenomena can be theoretically misleading.

Second, variables encode information about events. But *externally similar events may emerge from different internal mechanisms*. For example, it is possible to create a variable 'discrimination' on the basis of whether certain group of people is discriminated or not. Discrimination is not a static object but a process that characterizes interaction between two populations. Discrimination can refer to refusals and permissions, orderings and yieldings, which indicate open and closed possibilities for various individuals in their daily living (cf. Lewin 1997c). The problem is that the same behavior—encoded as 'discrimination'—can result from different hidden mental processes. Somebody is refused a job because of a race, the other—possibly from the same race—because of incompetence, the third because of age, gender, criminal record, or many other reasons. In this way an aggregate behavioral phenomenon is created that is caused by many different underlying causes. Which of the many possible causes led to the externally similar act encoded as 'discrimination' cannot be differentiated after the variable is encoded. The same 'phenomenon exists' code is attributed to several different phenomena in different cases.

Third, *variables can contain information about events at different levels of analysis*. Confusion of levels of analysis may lead to serious theoretical problems (see also Munro 1992). Many elements can in certain situations act collectively so that their behavior can be characterized by one or a few so-called 'order parameters' or 'collective variables' (e.g., Smith and Thelen 1993; Thelen and Smith 1994). One problem related to this situation is already mentioned; externally the same collective phenomenon may arise because of different hidden causal processes. The other problem related to collective variables is that the real causal factor can still be related to one element that happens to behave collectively together with other elements in certain situations. In a collective variable a behavior of an individual element cannot be distinguished from the behavior of the whole system in which the individual element is involved. In this case the real causal event may be masked; instead of a single element a complex phenomenon is misleadingly understood as a cause. The opposite is also possible: when we observe a behavior of a single element without knowing that behavior of this element is actually being determined collectively, we may attribute causal power to an element instead of the whole system part of which the element is. In psychology there are many obviously collective variables that are used as representing one and the same phenomenon. Sex can be an example. Psychological sex differences have been reported in thousands of studies. But what is represented by a variable 'sex'? There are numerous possibilities: level of certain hormones, presence or absence of certain body characteristics such as genitals, differences in the number and kind of genes, size and efficiency of muscles, differences in cultural values, etc. Conversely, we may encode into variable information about the level of hormones without realizing that 'sex' as a whole is actually relevant instead of the level of a hormone particularly. Even more, as elements behave collectively, we may attribute relevance to one element even though the other element of the same whole is actually important.

Finally, the problem of encoding a quality of a phenomenon is further complicated by the fact that *the same phenomenon can be classified differently because events can be described by more than one attribute at the same time*. Instead of phenotypical similarity or dissimilarity the facts of psychology may be classified on the basis of interdependence, for example. (e.g., Lewin 1997a). Thus, it might be theoretically justified to *encode* variables on the basis of causal information: events should be analyzed in terms of causal relations instead of phenotypical qualities. For example, a person can be described by qualities of being a man, a Caucasian, an alcoholic, etc. In this case a person is classified together with other men, Caucasians, or alcoholics, respectively. The same person, however, can be classified together with women (daughter and wife, a family), non-Caucasians (colleagues), non-alcoholics (some members of a therapy group), respectively. Choosing the relevant basis for classification of the same event or phenomenon becomes crucial in order to find causal relationships between events encoded as variables. The problem is that in quantitative psychology usually it is not known before encoding what basis is appropriate. The number of possible bases, however, is not limited in principle. Therefore, causal information must be used already at the stage of encoding variables, not at the stage of analyzing them.

'Zero' Does Not Necessarily Signify Absence; 'One' Does Not Necessarily Signify Presence

Problems with encoding qualitative information in a variable are not exhausted with the reasons provided above. In a dichotomous variable, value 'zero' does not necessarily indicate the absence of a phenomenon. *A phenomenon may exist but not manifest in a behavior*. There are two possibilities. First, there is a possibility that *a measurement tool is not sufficiently sensitive*. For example, a participant of a study may not understand an instruction. In that case the zero result—wrong or missing answer—does not correspond to the lack of a process a measurement tool is supposed to reveal. Another possibility is that instructions are not sufficiently unambiguous. For instance, an intelligence test may contain a task to continue the series of numbers: 2…4…8…? An expected by the compiler of a test answer would be 16 (the previous number multiplied by 2). There is, however, an unlimited number of other possibly correct answers not unambiguously ruled out by the usual instruction. Other rules applied to the series could be: every next is just bigger than the previous; a series of even numbers; a series of unconnected numbers; etc. So, the answer 46 could be entirely permissible by the instruction under many rules (bigger than the previous, even number, etc.) but still 'incorrect' for a test. Consequently, the answers that actually indicate possibly high—and creative—intelligence would be encoded as an absence of an ability to solve the problem.

Second, there is always a possibility that *a phenomenon manifests, but in an unexpected way* that is not taken into account in a measurement tool. In a study of short-term memory of persons with Traumatic Brain Injury (Toomela et al. 1999) some persons did not recall a single correct item from a list presented to them. However, these persons could at least try to give a meaningful response by reporting some words, though incorrect. Therefore, the patients remembered the instruction that must have been available for them in their short-term memory in order to provide any answer at all. Zero performance in this situation shows that persons are not able to perform *the* task, but it does not follow that they lack the underlying process that is supposed to be measured by a task.

On the other hand—for the reason discussed above that externally the same behavior may rely on different mental mechanisms—*observation of a behavior does not imply that a certain phenomenon exists*. In a short-term memory task, for example, 'correct' answers may result from echolalia, the involuntary parrotlike repetition (echoing) of a word or phrase just spoken by another person.

Finally, in some situations *the studied phenomenon cannot vary in principle*. There are many occasions where a whole system cannot exist without several subprocesses. Mind, for example, is impossible without the processes of perception and long-term memory. Long-term memory is a cause of mental acts as much as perception is. Invariable presence is required in order the complex process to emerge at all. Inability to perform on some perceptual or long-term memory tasks, therefore, does not suggest the absence of these processes but only the absence of an ability to utilize the processes in a particular task situation.

Quantitative Information—How Much or How Many?

Another kind of information encoded in variables is about a quantity of an observed phenomenon. The meaning of the quantitative information encoded in a variable depends on what quality is quantified. As discussed above, first distinction between qualities encoded as variables was between whether a phenomenon encoded as existing exists in reality or it is constructed by a researcher. Each of these kinds can further be differentiated according to (a) whether externally the same behavior encoded as a variable emerges from one or more underlying mental mechanisms, (b) at what level of analysis a phenomenon is encoded, and (c) what particular characteristic of the event out of infinitely many possible characteristics is the basis of encoding. Additional problem related to all these possible combinations of information encoded in a variable is that the meaning of the 'zero' does not imply the absence, and code 'one' does not imply presence, of the mental process that underlies the behavior. Quantification is related to all these qualitative aspects of variables as I discuss next.

Real Phenomenon, One Underlying Mechanism This is the only situation in which variable can meaningfully and unequivocally interpreted in a study of causal relationships *if and only if* three additional constraints are met at the same time. The quality encoded must be at the relevant for causal relationship level of analysis, the relevant for the causal relationship characteristic of an event must have been selected for encoding, and the meaning of 'zero' implies absence and the meaning of 'one' implies presence of the phenomenon encoded.

Real Phenomenon, More Than One Underlying Mechanism In this case variability has no unambiguous interpretation, qualitatively different phenomena (different underlying mental processes) are treated as the same.

Real Phenomenon, Wrong Level of Analysis The behavior of the elements of a system covaries with the behavior of the whole system the element belongs to. Therefore, it is not possible to decide at what level of analysis the variation is causally meaningful. Causally relevant characteristics of a whole event may inappropriately be attributed to the elements and vice versa.

Real Phenomenon, Wrong Characteristic of an Event This kind of a variable is open to different interpretations. First, selected for encoding characteristic is not causally relevant and does not covary with the causally relevant characteristic of the same event. In this case variable cannot show causal dependencies. And second, selected for encoding characteristic covaries with a causally relevant characteristic. In this situation causation will be attributed to a wrong characteristic.

Real Phenomenon, Meaning of 'Zero' is Ambiguous In this situation it is not clear whether causality should be attributed to the presence or absence of a phenomenon or to the presence of the phenomenon in a certain amount.

In case of *constructed phenomena*, another major problem of variable interpretation emerges. In this case variable does not contain information about the quantity of a phenomenon but rather how close or distant a particular observation is from the constructed prototype. No causal information can be encoded in this kind of a variable because prototypes—constructed phenomena—do not exist as causal events.

Constructed Quantitative Variability

Encoding of quantity brings also another problem. The problem is that variability can also be *real* or *constructed*. Perhaps one of the most common instances of potentially constructed variability emerges when a Likert scale is used. All questionnaires where Likert scale is used, force the informants to respond as if the quantity of the measured construct can vary at least at the number of levels determined by the used scale. There is no a priori law that determines how many possible values of the measured phenomenon actually exist. The phenomenon may exist qualitatively, yes or no, but the scale forces informants to introduce quantitative variability into a variable.

Variability Range

Depending on the studied sample and on the quality of the measurement tool a variable may not contain causally relevant information even when a causally relevant characteristic of the real phenomenon with one underlying mechanism is encoded at the appropriate level of analysis. Variable must encode information at the *causally relevant range of variability*. Let us assume, for instance, that certain size, say 4 units, of working memory is necessary for arithmetical operations with natural numbers from one to ten. If in a studied sample all participants have a working memory larger than sufficient four units, the variability in a working memory variable is in an irrelevant for arithmetical operations range. Relevant range must include cases with three units and cases with four units.

Variables and Statistical Data Analysis

Results of any kind of statistical data analysis can be theoretically meaningfully interpreted if and only if information encoded in variables is unambiguously defined. In psychology, meaning of variables seems to be, as a rule, ambiguous. In this situation statistical data analysis may reveal both misleading dependencies and misleading independencies. Therefore, it becomes problematic to interpret neither dependencies nor independencies in terms of mental processes supposedly underlying the variability in behavior encoded in variables.

It should be noted, that statistical data analysis can give misleading results even if only one out of all variables in the analysis contains misleading information. This is a sufficient condition for the emergence of misleading results.

Misleading Dependencies

Textbooks usually provide three possibilities why dependence between two variables cannot be causally interpreted. It is possible that a variable *A* causes *B*, *B* causes *A*, or both *A* and *B* are influenced by a third variable *X*. The fundamental problem, however, is that variables in psychology encode behaviors and not necessarily mental events that underlie a behavior. Behaviors are not in one-to-one correspondence with underlying mental processes. For this reason, as already discussed, the meaning of a variable becomes ambiguous. In this case dependence relations between variables cannot also unambiguously be extended to the mental events—there is no logical way back from patterns of dependencies to mental events, because there is already no way back from ambiguous variables to underlying mental events. Additional problem with the interpretation of dependencies is related to the difficulties in differentiating between two possible sources of dependence: causally relevant may be the presence of a quality or the presence of a quality in a sufficient quantity. As ‘zero’ does not imply the absence of a phenomenon and ‘one’ the presence of a phenomenon, dependence may wrongly be attributed to a quality or a quantity even if the variable correctly identifies the event.

There are several scholars who believe that causality can be revealed through statistical data analysis because certain patterns of probabilistic dependencies imply unambiguous causal dependencies (e.g., Pearl 2000; Sloman 2005; Spirtes et al. 2000). Variable relationships, however, cannot be unequivocally interpreted in terms of hidden mental events. Therefore, the belief that psychology can use statistical data analysis in order to reveal causal relations between mental events is not based on any scientifically acceptable evidence.

Misleading Independencies: Variable Encoding Problems

There is actually a more serious problem with interpreting variables statistically. All statistical inferences are based on the assumption that statistical independence implies the lack of causal relationships. This assumption, however, is not correct either. There are several possibilities for statistical independencies of variables when mental events that are supposed to be described by variables in the analyses are actually causally related.

One group of problems is related to the information encoded in variables. First, independence relations with variables that encode *constructed phenomena* can hardly tell anything about hidden from direct observation individual mental processes. For instance, according to the Five-Factor-Theory, the Big Five personality dimensions are not directly accessible either to observation or to introspection because they are so ‘deeply’ grounded in the organism (McCrae and Costa 1996 1999). Another possible interpretation of inaccessibility of the personality dimensions is, however, much simpler—these dimensions are constructed with the help of a factor analysis; these dimensions do not characterize the structure of mental processes of any individual (see also Molenaar 2004; Molenaar and Valsiner 2005, for a theory and evidence that entirely different personality factors emerge if data are analyzed for each individual separately). Therefore, there is nothing that could be accessed either by observation or introspection. If a variable encoding a constructed phenomenon does not correlate with other variables then no meaning can be attached to the lack of dependence.

Next, a variable may be independent from another if it contains information about *more than one mental event* that manifests in an externally similar behavior. It is possible, for example, that one mental mechanism in a variable is causally related to a variable *X*, and their correlation is positive; the other mechanism in the same variable to a variable *X* so that their correlation is negative. In this case the hidden mechanisms together in a variable may cancel each other out in a statistical dependence relation. For example, there are different strategies for solving simple arithmetical problems, some more advanced than the others (Siegler 1996). In this situation ‘mathematical intelligence’ may be independent of, let us say, ‘academic performance’ not because intelligence and academic performance are independent but because the same ‘correct’ answers are provided by both less intelligent children who rely on primitive calculation strategies and more intelligent children who rely on advanced calculation strategies.

Third, *confusion of levels of analysis* may exclude the variability in a critical range. For example, it is known that, as a group, men tend to outperform women in certain motor and perceptual-spatial tasks. More detailed studies, however, reveal that performance of women varies together with an activity of sex hormones. Depending on the hormone levels, the performance on the same task may differ more and less from men (Chiarello et al. 1989; Hampson and Kimura 1988). So, it is possible that sexual hormones—that vary together with many other sex and gender characteristics, such as personal space (Sanders 1978) or kinesthetic aftereffect (Baker et al. 1979)—if encoded as a collective variable ‘sex,’ reveal dependence or independence patterns that may mask the real causal relationships. Women as a group are more different from men than one woman in certain phases of menstruation cycle. The ‘cause’ of difference can wrongly be attributed to ‘sex’ whereas the real difference is caused by a hormonal process at another level of analysis. Theoretically—and there is no reason to believe that analogous situations cannot happen with some other constructs—if average female performance would be similar to men, then the real differences between sexes that appear only at a certain activity level of hormones, would not be discovered and no dependence between sex and test performance would be discovered.

Fourth, a *causally irrelevant characteristic of an event* may be encoded instead of causally relevant one. In that case, again, variables may be independent even though the dependence relations between events actually exist. For example, in a study on the cognitive consequences of the Traumatic Brain Injury (TBI) we found that the performance of a healthy control group did not statistically differ from the performance of the group of persons with TBI on a recall of a list of nonwords.

After recategorizing the study participants according to whether more or less than 6 nonwords was recalled, it turned out that there was a substantial subgroup of persons with TBI who *outperformed* healthy controls (Toomela et al. 1999). So, categorizing events of recall on the basis of the presence or absence of TBI revealed one pattern of in/dependencies and categorizing the same events on the basis of memory performance according to certain criterion revealed a different pattern of dependencies with theoretically completely different interpretation.

Fifth, independence between variables may emerge when a *measurement tool is not sufficiently sensitive*. It may turn out that persons who perform 'zero' on a short-term memory task (see, 'Zero' does not necessarily signify absence, 'One' does not necessarily signify presence, above), actually should be distinguished into two groups. One where short-term memory is truly missing and the other where the persons are able to remember the instruction but short-term memory capacity is exhausted with this. Another example can be found from clinical psychology. In some cases of personality disorders psychotherapy results in increase in the number of complaints because patients begin to recognize their destructive and depressive feelings (Blount et al. 2002). In this situation untreated persons would report similar lack of complaints as healthy persons and the number of personality disorder symptoms would be independent of the presence or absence of the disorder. Similarly, brain damage may be accompanied by supranormal optimism so that the number of complaints does not distinguish healthy persons from persons with some forms of brain damage (Tomberg et al. 2007; Tomberg et al. 2005; Toomela et al. 2004). Wrong measure—self-report—may hide serious problems that accompany brain damage.

Sixth, *some phenomenon may be absolutely necessary* for a complex process. It has been found, for instance, that transitivity judgments can be made by children operationally and nonoperationally (this is also another example of different mental mechanisms that lead to similar behavior of 'correct' answers). Operational solution requires invariable exact memory for the premises of the task (Chapman and Lindenberger 1992b). Scientists who confuse statistical theory with the theory about mental events may in this situation make a conclusion that thinking becomes 'memory-free' because invariable memory is stochastically independent of transitivity inference performance (Brainerd and Reyna 1992). Psychological approach to the question shows clearly that some forms of thinking require unconditionally verbatim remembering of the task. Statistical independence does not reflect independence of underlying mental processes (Chapman and Lindenberger 1992a).

The list of possible reasons can be extended to all problems related to *quantitative encoding of event information*. Statistically constructed events, for example, may have no relation to actually occurring mental events. In many cases results of the effect of obviously multiple mechanisms that manifest in a behavior are treated as one. For example, variability in IQ may be attributed to some (constructed?) general intelligence factor. But variability in IQ is also dependent on variability in social-cultural background related to a specific form of intelligence, academic orientation, self-efficacy beliefs, test anxiety, self-confidence, etc. (Richardson 2002). If in some sample variation is caused by one underlying mechanism and in another sample by another mechanism, then in the composite sample no dependencies may emerge even though both underlying mechanisms separately would reveal clear dependencies.

Finally, and obviously, if variability in a given sample is not in a *critical range*, a variable will be independent of another even when underlying mental processes are causally related. (See *Variability range*, above, for an example.)

Misleading Independencies: Statistical Paradoxes

Statistical data analysis can lead to ambiguous interpretations even if a variable encodes information about an underlying event unambiguously. The problem is that a pattern of relationships between variables can change fundamentally if new variables are added into an analysis or the information encoded in a same set of variables is analyzed in different ways. One commonly discussed paradoxical phenomenon in this context is a *reversal paradox* (also known as *Simpson's paradox* or *Yule–Simpson effect*) according to which any statistical relation between two variables may be reversed by including additional factors in the analysis. This situation is related to misleading dependencies—where it is not clear whether the dependence relation between variables is positive or negative. In both cases, nevertheless, some causal relationship between variables can be assumed.

Meehl's Paradox There are other situations, however, where events or variables are related so that there are no pair-wise relationships between variables and yet a *configuration of variables* may perfectly predict a criterion variable. One such paradoxical situation was described by Meehl (1950) (see also Pearl 2000; Sloman 2005; Spirtes et al. 2000 on a discussion of this phenomenon in the context of causal model theory). Meehl showed that two dichotomously scored test items used in predicting a dichotomous criterion may be individually independent from the criterion and yet predict the criterion configurally. If two items singly have zero validity then configural validity emerges when the interitem correlation is different in the two categories of the criterion. Configural validity may emerge both in situations where interitem correlations have either symmetrically or asymmetrically opposite signs in the two categories of the criterion and in situations where interitem correlations are with the same sign.

A theoretical example of Meehl's paradox is shown in Table 1. In Meehl's example, the dichotomous criterion (*X* in Table 1) is "schizophrenic" versus "normal." There are also two items (*A*, and *B*) in a test that can be answered "yes" or "no." If normals answer both items consistently (yes–yes or no–no) and schizophrenics answer the same items inconsistently (yes–no or no–yes) then both items alone will have zero validity for the criterion. Even more, *A* and *B* are also independent from each other. There is no statistical criterion showing that both of them should be included in the analysis if pair-wise dependencies are analyzed. Scoring of the same items configurally (if $A=B$, then "normal," if $A\neq B$, then "schizophrenic") the criterion can be predicted perfectly.

Some authors have suggested that this kind of paradox can be ignored because it does not arise very often in the world (e.g., Sloman 2005 assumed that the paradox arises only if probabilities for independent dichotomous variables are both equal to 0.5 and, therefore, such possibility should be ignored). There is no way in science to give a proof for absence of some possibility. Therefore, if some paradoxical

Table 1 Examples of statistical paradoxes with misleading independence effects

Meehl's paradox			Extended Meehl's paradox				Other effects				
A	*B*	*X*	*C*	*D*	Y1	Y2	*E*	*F*	|*E*−*F*|	*E*/*F*	*Z*
0	0	0	0	0	0	0	1	17	16	0.06	8
0	0	0	0	0	0	0	2	16	14	0.13	7
0	1	1	0	1	1	1	3	15	12	0.20	6
0	1	1	0	1	1	1	4	14	10	0.29	5
1	0	1	0	2	1	2	5	13	8	0.38	4
1	0	1	0	2	1	2	6	12	6	0.50	3
1	1	0	1	0	1	1	7	11	4	0.64	2
1	1	0	1	0	1	1	8	10	2	0.80	1
			1	1	0	0	9	9	0	1.00	0
			1	1	0	0	10	8	2	1.25	1
			1	2	1	1	11	7	4	1.57	2
			1	2	1	1	12	6	6	2.00	3
			2	0	1	2	13	5	8	2.60	4
			2	0	1	2	14	4	10	3.50	5
			2	1	1	1	15	3	12	5.00	6
			2	1	1	1	16	2	14	8.00	7
			2	2	0	0	17	1	16	17.00	8
			2	2	0	0					

phenomenon has been observed even once, it cannot be ignored in order to make problematic statistical models more plausible. There are such cases, which, even though not perfectly, correspond to the Meehl's paradox. For instance, it has been found that sexual physical characteristics are related to the level of cognitive test performance differently in men and in women. In men lower spatial ability is related to more masculine body characteristics whereas the opposite relationship characterizes women (Petersen 1976). If to encode a variable A so that 'male'=0 and 'female'= 1; variable *B* so that 'low masculinity'=0 and 'high masculinity'=1; and variable *X* so that 'low test performance'=0 and 'high test performance'=1, then the paradox emerges. Configurations $A=B=0$ and $A=B=1$ both predict relatively high cognitive performance and configurations $A \neq B$ both predict relatively low cognitive performance. Similar in men and women opposite relationship has been found between spatial abilities and levels of androgen (Shute et al. 1983). Therefore, independence between variables cannot always be interpreted as a lack of (causal) relationship.

Extension of Meehl's Paradox Meehl's paradox can be extended to other variable situations. So, for example, it is possible to find trichotomously encoded variables C and D that, as a configuration, predict a dichotomously encoded criterion variable Y1 even though C, D, and Y1 are all independent one from another (see Table 1, Extended Meehl's paradox). Even more, the variables do not need to be nominal. Similar paradox can arise with continuous variables as well. This paradox emerges with a set of variables *C*, *D*, and Y2. Variable Y2 can perfectly be predicted by variables *C* and *D*: $Y2=|C-D|$. It is possible to ask, again, whether such a situation may emerge in a real life. One example can be constructed from the studies of brain—emotion relationships. It is known that emotions are lateralized; the right hemisphere is more related to the negative emotions and the left hemisphere is more related to the positive emotions (Borod and Koff 1989; Lee et al. 1990). Next, results of studies of depression are in agreement with the idea that depression is not necessarily related only to the dysfunction of the right hemisphere. The problem may be related to dysfunctional relationship between hemispheres, depression may arise both because of the overactivation of the right hemisphere and underactivation of the left hemisphere (cf. Bruder et al. 1989, 2002). This view can further be supported by studies showing that not all patients respond in the same way to the treatment. Treatment responsiveness can be predicted by the pattern of hemispheric asymmetry. Also, treatments can have a lateralized effects affecting, depending on the specific treatment, the processes mostly in one or the other of the hemispheres (cf., Bruder et al. 1990, 1997, 2001, 2004, 2007; Chistyakov et al. 2005; Stewart et al. 1999). Altogether, it can be suggested that abnormal activity lateralization of the brain is related to affective disorders. Further, if to assume that abnormal activity pattern is continuously related to the increase in the severity of the affective disorder, a paradoxical relationship between variables *C*, *D*, and Y2 emerges. If C encodes the level of activation of the left hemisphere, *D* encodes the level of activation of the right hemisphere, and Y2 encodes the severity of an affective disorder, then neither *C* nor *D* individually predict Y2. Absolute difference between *C* and *D*, however, is perfectly related to Y2. Again, as *C* and *D* are also independent from each other, there is no clear statistical information that would suggest looking for configural dependence with the criterion variable.

Other Statistical Paradoxes There are other situations where predictor variables are not correlated with the criterion variable but some combination of them predicts the criterion significantly beyond chance, even perfectly. Such a situation emerges, for instance, when continuous independent variables (*E* and *F*, Table 1) are inversely correlated. Both of them are linearly independent of the criterion variable *Z*, however. One problem that may emerge in this situation is already discussed—a researcher may find that highly (perfectly) correlated variables are just redundant and to use in analyses only one or the other. They may assume, for example, that both represent equally well a behavior of a system as collective variables. It would be completely misleading, though, because *Z* can be perfectly predicted by *E* and *F* together. In this case the perfect predictor is an absolute value of the difference score ($|E-\mathrm{F}|$ in Table 1) that is perfectly correlated with the criterion variable *Z*.

A group of findings in psychology where such relationships between variables could be observed is related to so-called U-shaped changes in development. There is increasing evidence that in very different areas of development, previously performed behaviors may regress and appear again after some time (Siegler 2004). One possible reason for this situation to emerge in development is related to the possibility that more than one mechanism is responsible for externally the same behavior. One mechanism may become deactivated in time and another, more developed, would become more activated. To apply two mechanisms at the same time to solving a problem would lead to low level of performance because different mechanisms approach the same problem differently. High level performance follows if only one or the other of the mechanisms is used but not both together. If, for instance, there are two mechanisms possibly activated at the same time then equal activation of the mechanisms would lead to low performance and unequal level of activation to high level of performance. In the latter case, high level of performance, behaviorally the same, would result from different underlying mechanisms. If to represent activation level of one mechanism as a variable E (Table 1) and activation level of another mechanism with a variable F, then performance level would be linearly independent of activation of either of the mechanisms. Absolute difference score $|E-F|$, however, would predict the performance perfectly.

There are, yet, other possibilities where some combination of independent variables would predict the criterion even when both independent variables are not linearly correlated with the criterion. Multiplication of E by F would create a variable that almost perfectly would predict Z as well (in Table 1, the correlation would be 0.97). Summing and multiplying would give almost the same result but the question must emerge, what does it mean to sum or multiply? Without knowing exactly what mental events are represented in variables it would be impossible to answer this question.

Also, E divided by F would be a better predictor than E or F taken alone (see Table 1, correlation between E/F and Z would be 0.46). A version of the latter situation can also been found in some studies. For instance, there was a study where relationships between the level of testosterone (T) and dihydrotestosterone (DHT) with performance on spatial tasks were investigated (McKeever and Deyo 1990). It was found that neither the T nor DHT level was correlated with the spatial test performance. Correlation was found between the test performance and the absolute value of the DHT/T ratio. The more deviant from the average level in either direction was the DHT/T ratio, the better was the spatial test performance.

Taken together, conceptually misleading independencies emerge again and again in research. There are sometimes, of course, statistical procedures available to discover a dependence relationship between variables. But the problem is that the number of possible statistical tools needed, and different variable configurations to be tried, is not formally limited. Substantial theory must precede the statistical analysis. In that case, however, the question emerges, why statistical analysis would be needed because the understanding would have resulted already without these analyses.

Conceptual Problems with the Statistical Analysis of Probabilistic Dependencies

The list of potential problems related to statistical data analyses is not exhausted as yet. Another problem is conceptual—the problem is related to inappropriate translation of phenomena into the language of statistical data analysis, again. Statistical data analysis is based on only one kind of relationships between variables—variables can be only dependent or independent. Events and phenomena, represented as variables, however, may have qualitatively different kinds of relationships. The same phenomena in qualitatively different relationships compose qualitatively different higher-order systems. Inability to differentiate between qualitatively different phenomena means that the studied phenomena are not understood.

Let us take a simple example of water, utilized many times by Gestalt psychologists (e.g. Koffka 1935). Hydrogen and oxygen, the components of the molecule of water, can exist in different relationships. If the two gases are just mixed, a highly explosive gas is formed. The same elements in another kind of relationship compose water that does not explode. If we were just correlating the amount of hydrogen and oxygen in the environment with the amount of water, the correlation would be far from perfect even though the event of the chemical reaction between hydrogen and oxygen causes the emergence of water molecules.

Psychological examples would not be unusual as well. Take, for instance, dyslexia, a specific learning disability that manifests primarily as a difficulty with reading and spelling. Dyslexia is often viewed as a specific phonological deficit (Ramus 2004). Indeed, dyslexia is usually associated with phonological deficits that may underlie the reading difficulties. But what about deaf children who also can learn to read? They, obviously, have a very severe "phonological deficit." And they do read. Deaf children, however, are in a different relationship with the education, with social-cultural environment. The way deaf children learn to read is not alphabetical reading where letters are mapped to speech sounds. The same phenomenon—reading ability—emerges differently depending on the ways children are taught to read, that is, depending on the particular qualitative relationship between children and their educational environment. If we were studying children from a culture where reading education is not alphabetical, phonological deficits would not be related to dyslexias. Dependence or independence between phonological processes and reading ability would depend on qualitatively different kind of child–environment relationships. This is exactly the case when reading is studied in different cultures (Ardila 2003). Even more, the relationship between phonological abilities and reading level can, in different cultures, have all possible different forms: phonological ability is necessary prerequisite of reading in some cultures, in other cultures reading ability can develop independently of phonological skills and lead to increase in phonological processing, yet in other cultures the relationship can be truly bidirectional (cf., Ardila 2003).

Variable-Psychology—Useless for Theory Development?

Psychology studies mind that is not accessible for direct observation. Mind manifests through behavior. All knowledge about mind, therefore, is based on interpretation of behavior. In modern mainstream psychology majority of studies "translate" observed behaviors or consequences of behaviors, such as test results, into variables and base the psychological theories on statistical analysis of variables. Implicitly it is assumed that study of variables informs researchers about the underlying mental mechanisms. In this paper I argue that this "variable-approach" is extremely misleading. The problem is that variables do not represent directly mental phenomena but only behavior. And behavior is not in one-to-one correspondence with mental processes. Externally the same behavior may internally rely on different mental mechanisms and externally different behaviors may have the same underlying mental process.

So, any study of potentially causal relationships between variables can, at the best case, reveal only causal relationships between external behaviors. But not between the real objects of studies, hidden from direct observation mental processes. This problem leads to a paradoxical situation. Causality can be understood through the analysis of variables only if it is exactly known how mental processes cause behaviors that are represented as variables. But the (statistical) study of variables does not provide necessary information because variables represent only indirect effects of mental processes, mediated by behavior. Variables, therefore, do not give unambiguous information about how qualitative information is represented in variables, whether qualitative information in variables is meaningfully quantified, and whether quantitative variability range is appropriately represented. In addition, all interpretation of statistical data analyses is meaningful only if independence relations between variables truly reflect the lack of causal relations between events represented as variables. Because of ambiguity of information encoded in variables and also because of limitations of statistical data analysis procedures themselves, neither dependencies nor independencies can be interpreted in terms of causality.

Therefore, nonstatistical methods of research, such as used in certain schools of neuropsychology, for instance (e.g., Luria 1969, 1973), must be used instead for understanding the content of variables. If, however, we already understand how mental processes cause behaviors we represent in variables, then why to use statistical procedures at all?

There still remains a question, why statistical data analysis seems to give theoretically meaningful information about hidden mental processes? After all, millions of studies have been conducted and published, all based on different kinds of statistical data analysis. Is it really possible that all of them are theoretically misleading? The history of all other sciences—physics, chemistry, biology, geology, geography, etc.—clearly shows that the answer can be yes. In all sciences there have been phases where all scholars have been wrong. Nevertheless, if the statistical variable studies would be meaningless in all possible ways, there apparently would be no such amount of statistical studies conducted. If no unambiguous theory can be built about the hidden from direct observation mental processes with a variable-approach then the meaning must be somewhere else. One obvious perspective in which variable-psychology is meaningful is behavioral. Variables can be relatively unambiguously understood in terms of observed behaviors, behavior can be predicted with the help of variables. In other words, external behavior is easier to understand for exactly the very reason that it can be directly observed and directly related to variables. The fundamental problem is that prediction does not require understanding of underlying mechanisms.

Taken together, analysis of variables that represent behaviors can only lead to prediction of behaviors without understanding the causal mental processes that only manifest in behavior. Therefore, variable-psychology is an excellent tool for predicting behaviors until theoretical understanding of mind that manifests in behavior is developed. There are too many reasons to suggest, however, that variable-psychology cannot contribute to the development of this theory in principle. Variable-psychology has a very limited range of application for the development of a theory of mind—psychology.

I would not like to end with a pessimistic note. Negative critique can show the shortcomings of current practice, but without positive program the negative can only lead to confusion. If theory of psychology wins so little, if at all, from statistical variable-psychology, then there must be another methodological approach available. I believe this approach is actually not anything that needs to be created from the very beginning. On the contrary, there are reasons to suggest that pre-WWII Continental-

European psychology methodology can provide several ideas how to proceed (Toomela 2007a,b, in press; Watson 1934): Psychology should rely on qualitative studies that go beyond observation of quantitative relationships between variables; studies should take into account that the phenomenon under study, mind, only manifests in behavior; qualitative levels of analysis should be clearly distinguished, it should be taken into account that elements in a whole cannot be independent in principle; studies should focus on cases instead of groups; typology is the main methodological tool for generalization in psychology; prediction without insight, without a substantive theory, should not be acceptable; selection of facts should be systematically guided by theory; and scholars should not constrain themselves to interpretation of data provided by convenient methodological tools, such as statistical data analysis—psychologist should be more interested in thinking about the meaning of collected facts than in the accumulation of facts as such.

Acknowledgments This study was supported by the Estonian Science Foundation Grant No. 7490.

References

Ardila, A. (2003). Culture in our brains: Cross-cultural differences in the brain-behavior relationships. In A. Toomela (Ed.) *Cultural guidance in the development of the human mind* (pp. 63–86). Westport, CT: Ablex.

Baker, A. H., Mishara, B. L., Kostin, I. W., & Parker, L. (1979). Menstrual cycle affects kinesthetic aftereffect, an index of personality and perceptual style. *Journal of Personality and Social Psychology*, *37*(2), 234–246.

Blount, C., Evans, C., Birch, S., Warren, F., & Norton, K. (2002). The properties of self-report research measures: Beyond psychometrics. *Psychology and Psychotherapy: Theory, Research and Practice*, *75*, 151–164.

Blumer, H. (1956). Sociological analysis and the "Variable.". *American Sociological Review*, *26*(6), 683–690.

Borod, J. C., & Koff, E. (1989). The neuropsychology of emotion: Evidence from normal, neurological, and psychiatric populations. In E. Perecman (Ed.) *Integrating theory and practice in clinical neuropsychology* (pp. 175–215). Hillsdale, NJ: Erlbaum.

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071.

Brainerd, C. J., & Reyna, V. F. (1992). The memory independence effect: What do the data show? What do the theories claim. *Developmental Review*, *12*, 164–186.

Bruder, G. E., Quitkin, F. M., Stewart, J. W., Martin, C., Voglmaier, M. M., & Harrison, W. M. (1989). Cerebral laterality and depression: Differences in perceptual asymmetry among diagnostic subtypes. *Journal of Abnormal Psychology*, *98*(2), 177–186.

Bruder, G. E., Stewart, J. W., Ma, G. J., McGrath, P. J., Wexler, B. E., & Quitkin, F. M. (2002). Atypical depression: Enhanced right hemisphere dominance for perceiving emotional chimeric faces. *Journal of Abnormal Psychology*, *111*(3), 446–454.

Bruder, G. E., Stewart, J. W., McGrath, P. J., Deliyannides, D., & Quitkin, F. M. (2004). Dichotic listening tests of functional brain asymmetry predict response to fluoxetine in depressed women and men. *Neuropsychopharmacology*, *29*, 1752–1761.

Bruder, G. E., Stewart, J. W., Mercier, M. A., Agosti, V., Leite, P., Donovan, S., et al. (1997). Outcome of cognitive-behavioral therapy for depression: Relation to hemispheric dominance for verbal processing. *Journal of Abnormal Psychology*, *106*(1), 138–144.

Bruder, G. E., Stewart, J. W., Schaller, J. D., & McGrath, P. J. (2007). Predicting therapeutic response to secondary treatment with bupropion: Dichotic listening tests of functional brain asymmetry. *Psychiatry Research*, *153*(2), 137–143.

Bruder, G. E., Stewart, J. W., Tenke, C. E., McGrath, P. J., Leite, P., Bhattacharya, N., et al. (2001). Electroencephalographic and perceptual asymmetry differences between responders and nonresponders to the SSRI antidepressant. *Biological Psychiatry*, *49*, 416–425.
Bruder, G. E., Stewart, J. W., Voglmaier, M. M., Harrison, W. M., McGrath, P., Tricamo, E., et al. (1990). Cerebral laterality and depression: Relations of perceptual asymmetry to outcome of treatment with tricyclic antidepressants. *Neuropsychopharmacology*, *3*(1), 1–10.
Campbell, J. K., O'Rourke, M., & Silverstein, H. (2007). *Causation and Explanation*. Cambridge, MA: MIT.
Chapman, M., & Lindenberger, U. (1992a). How to detect reasoning-remembering dependence (And how not to). *Developmental Review*, *12*, 187–198.
Chapman, M., & Lindenberger, U. (1992b). Transitivity judgments, memory for premises, and models of children's reasoning. *Developmental Review*, *12*, 124–163.
Chiarello, C., McMahon, M. A., & Schaefer, K. (1989). Visual cerebral lateralization over phases of the menstrual cycle: A preliminary investigation. *Brain and Cognition*, *11*, 18–36.
Chistyakov, A. V., Kaplan, B., Rubichek, O., Kreinin, I., Koren, D., Hafner, H., et al. (2005). Effect of electroconvulsive therapy on cortical excitability in patients with major depression: A transcranial magnetic stimulation study. *Clinical Neurophysiology*, *116*, 386–392.
Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, *12*, 671–684.
Dunlap, W. P., Brody, C. J., & Greer, T. (2000). Canonical correlation and chi-square: relationship and interpretation. *Journal of General Psychology*, *127*(4), 341–353.
Essex, C., & Smythe, W. E. (1999). Between numbers and notions. A critique of psychological measurement. *Theory and Psychology*, *9*(6), 739–767.
Galton, F. (1878). Composite portraits. *Nature*, *18*, 97–100.
Hampson, E., & Kimura, D. (1988). Reciprocal effects of hormonal fluctuations on human motor and perceptual-spatial skills. *Behavioral Neuroscience*, *102*(3), 456–459.
Keppel, G. (1991). *Design and Analysis*. A Researcher's Handbook. Englewood Cliffs: Prentice Hall.
Koffka, K. (1935). *Principles of Gestalt psychology*. London: Routledge.
Lee, G. P., Loring, D. W., Meador, K. J., & Brooks, B. B. (1990). Hemispheric specialization for emotional expression: A reexamination of results from intracarotid administration of sodium amobarbital. *Brain and Cognition*, *12*, 267–280.
Lewin, K. (1935). *A dynamic theory of personality. Selected papers*. New York: McGraw-Hill.
Lewin, K. (1997a). Field theory and experiment in social psychology. (Originally published in 1939). In K. Lewin (Ed.) *Resolving social conflicts and field theory in social science* (pp. 262–278). Washington, DC: American Psychological Association.
Lewin, K. (1997b). Field theory and learning. (Originally published in 1942). In K. Lewin (Ed.) *Resolving social conflicts and field theory in social science* (pp. 212–230). Washington, DC: American Psychological Association.
Lewin, K. (1997c). Frontiers in group dynamics. (Originally published in 1947). In K. Lewin (Ed.) *Resolving social conflicts and field theory in social science.* (pp. 301–336). Washington, DC: American Psychological Association.
Luria, A. R. (1969). *Vyshije korkovyje funktsii tsheloveka i ikh narushenija pri lokal'nykh porazenijakh mozga*. *[Higher cortical functions in man and their disturbances in local brain lesions]*. Moscow: Izdatel'stvo Moskovskogo Universiteta.
Luria, A. R. (1973). *Osnovy neiropsikhologii*. Moscow: Izdatel'stvo MGU.
McCrae, R. R., & Costa, P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the Five-Factor Model. In J. S. Wiggins (Ed.) *The Five-Factor Model of personality* (pp. 51–87). New York: Guilford.
McCrae, R. R., & Costa, P. T. (1999). A Five-Factor Theory of personality. In A. Lawrence, & O. P. J. Pervin (Eds.) *Handbook of personality: Theory and Research.* (pp. 139–153). New York: Guilford.
McKeever, W. F., & Deyo, R. A. (1990). Testosterone, dihydrotestosterone, and spatial task performance of males. *Bulletin of the Psychonomic Society*, *28*(4), 305–308.
Meehl, P. E. (1950). Configural scoring. *Journal of Consulting Psychology*, *14*(3), 165–171.
Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory and Psychology*, *10* (5), 639–667.
Michell, J. (2003). The quantitative imperative. Positivism, naive realism and the place of qualitative methods in psychology. *Theory and Psychology*, *13*(1), 5–31.
Molenaar, P. C. M. (2004). A Manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218.

Molenaar, P. C. M., & Valsiner, J. (2005). How generalization works through the single case: A simple idiographic process analysis of an individual psychotherapy. *International Journal of Idiographic Science,* Article 1. Retrieved October 25, 2005 from http://www.valsiner.com/articles/molenvals.htm.
Munro, D. (1992). Process vs structure and levels of analysis in psychology. Towards integration rather than reduction of theories. *Theory and Psychology*, *2*, 109–127.
Pearl, J. (2000). *Causality. Models, reasoning, and inference*. Cambridge, UK: Cambridge University Press.
Petersen, A. C. (1976). Physical androgyny and cognitive functioning in adolescence. *Developmental Psychology*, *12*, 524–533.
Ramus, F. (2004). Neurobiology of dyslexia: A reinterpretation of the data. *Trends in Neurosciences*, *27* (12), 720–726.
Richardson, K. (2002). What IQ tests test. *Theory and Psychology*, *12*(3), 283–314.
Russo, F., & Williamson, J. (2007). *Causality and Probability in the Sciences*. London: College Publications.
Sanders, J. L. (1978). Relation of personal space to the human menstrual cycle. *Journal of Psychology*, *100*, 275–278.
Shute, V. J., Pellegrino, J. W., Hubert, L., & Reynolds, R. W. (1983). The relationship between androgen levels and human spatial abilities. *Bulletin of the Psychonomic Society*, *21*(6), 465–468.
Siegler, R. S. (1996). *Emerging minds*. *The process of change in children's thinking*. New York: Oxford University Press.
Siegler, R. S. (Ed). (2004). Special Issue: U-shaped changes in behavior and their implications for cognitive development. *Journal of Cognition and Development*, *5*(1), 1–151.
Sloman, S. (2005). *Causal models. How people think about the world and its alternatives*. Oxford: Oxford University press.
Smith, L. B., & Thelen, E. (1993). *A dynamic systems approach to development*. Applications. Cambridge, MA: MIT.
Sohn, D. (1999). Experimental effects. Are they constant or variable across individuals? *Theory and Psychology*, *9*(5), 625–638.
Sosa, E., & Tooley, M. (1993). *Causation*. Oxford: Oxford University Press.
Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search*. Cambridge, MA: MIT.
Stewart, J. W., Wuitkin, F. M., McGrath, P. J., & Bruder, G. E. (1999). Do tricyclic responders have a different brain laterality? *Journal of Abnormal Psychology*, *108*(4), 707–710.
Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT.
Tomberg, T., Toomela, A., Ennok, M., & Tikk, A. (2007). Changes in coping strategy, social support, optimism and health-related quality of life following traumatic brain injury: A longitudinal study. *Brain Injury*, *21*(5), 479–488.
Tomberg, T., Toomela, A., Pulver, A., & Tikk, A. (2005). Coping strategies, social support, life orientation and health-related quality of life following traumatic brain injury. *Brain Injury*, *19*(14), 1181–1190.
Toomela, A. (2007a). Culture of science: Strange history of the methodological thinking in psychology. *Integrative Psychological and Behavioral Science*, *41*(1), 6–20.
Toomela, A. (2007b). History of methodology in psychology: Starting point, not the goal. *Integrative Psychological and Behavioral Science*, *41*(1), 75–82.
Toomela, A. (2008). Kurt Lewin's contribution to the methodology of psychology: From past to future skipping the present. In J. Clegg (Ed.) *The observation of human systems: Lessons from the history of Anti-Reductionistic Empirical Psychology*. New Brunswick, N. J.: Transaction Publishers.
Toomela, A. (in press). Methodology of idiographic science: Limits of single-case studies and the role of typology. In S. Salvatore, J. Valsiner, S. Strout, & J. H. Clegg (Eds.) *YIS: Yearbook of idiographic science 2008*. Rome: Giorgio Firreira.
Toomela, A., Pulver, A., Tomberg, T., Orasson, A., Tikk, A., & Asser, T. (2004). Possible interpretation of subjective complaints in patients with spontaneous subarachnoid haemorrhage. *Journal of Rehabilitation Medicine*, *36*(2), 63–69.
Toomela, A., Tomberg, T., Orasson, A., Tikk, A., & Nômm, M. (1999). Paradoxical facilitation of a free recall of nonwords in persons with traumatic brain injury. *Brain and Cognition*, *39*, 187–201.
Watson, G. (1934). Psychology in Germany and Austria. *Psychological Bulletin*, *31*(10), 755–776.

Aaro Toomela is a Professor of Neuropsychology at the Tallinn University, Estonia. He received his MD from the University of Tartu in 1986. After spending 7 years as a child neurologist in the Department of Neurorehabilitation at the Tartu Children's Clinic, he moved into psychology. From 1993 to 2000 he was a lecturer specializing in neuropsychology in the Department of Psychology at the University of Tartu. Research from that time concerned the consequences of brain damage, development of memory and drawings, and the role of culture in the development of the human mind. He received his MSc in 1996 and his PhD in 2000 from the University of Tartu. To get acquainted with the applied aspects of psychology, Aaro Toomela moved to the military world, where he was a head of the Department of Social and Behavioral Sciences in the Academy of the Estonian Defence Forces. There he worked on the test battery for military personnel selection in the Estonian Defence Forces and on different aspects of the psychology of alcoholism and drug abuse. Now Aaro Toomela is back to the academic world. His current research interests concern, among other issues, methodology of research, philosophy of science, school effectiveness, semiotically mediated processes, and prediction of recidivism.