

Culture of Science: Strange History of the Methodological Thinking in Psychology

Aaro Toomela

Published online: 7 July 2007

© Springer Science + Business Media, LLC 2007

Abstract In pre-World-War-II psychology, two directions in methodological thought—the German–Austrian and North American ways—could be differentiated. After the war, the German–Austrian methodological orientation has been largely abandoned. Compared to the pre-WWII German–Austrian psychology, modern mainstream psychology is more concerned with accumulation of facts than with general theory. Furthermore, the focus on qualitative data—in addition to quantitative data—is rarely visible. Only external—physical or statistical—rather than psychological controls are taken into account in empirical studies. Fragments—rather than wholes—and relationships are studied, and single cases that contradict group data are not analyzed. Instead of complex psychological types simple trait differences are studied, and prediction is not followed by thorough analysis of the whole situation. Last (but not least), data are not systematically related to complex theory. These limits have hindered the growth of knowledge in the behavioral sciences. A new return to an updated version of the German–Austrian methodological trajectory is suggested.

Keywords Methodology · Holism · Part/whole relations

In recent years an increasing dissatisfaction with methods and their corresponding theoretical thinking in psychology as a science can be observed (e.g., Essex and Smythe 1999; Grace 2001; Michell 2000, 2003b; Molenaar 2004; Molenaar and Valsiner 2005; Sohn 1999). Again the behavioral sciences are in a methodological crisis due to the limitations that currently used methods set for the given disciplines.

There are two ways for overcoming limitations of methods used in psychology. One is to invent new methods of research. That may entail the creation of more and

A. Toomela (✉)

Department of Special Education, University of Tartu, Salme 1a, Tartu 50103, Estonia
e-mail: aaro.toomela@ut.ee

increasingly sophisticated statistical methods of data analysis, different brain-imaging methods. The other way is to look back into the history of methodological thought and ask whether methodological principles applied in research long time ago and abandoned in the course of history were abandoned due to inherent limitations of them or did they disappear due to purely non-scientific cultural reasons. In this paper the historical approach is taken. Before World War II two general schools of psychology, North American and German–Austrian, could be distinguished. I suggest that German–Austrian psychological thinking was rich in its potential for building a unified science, and this exceeded in several ways the North American way of thinking in psychology. Recent mainstream psychology follows mostly pre-WWII North American way of thinking. Adoption of German–Austrian principles could substantially facilitate our understanding of the human mind.

Two Psychologies in the Pre-World War II Era

The differences between German–Austrian and North American methodological thought were described by Watson (1934). Before describing the differences between these two schools of thought it should be mentioned that German–Austrian school was actually more widespread, including Switzerland, Sweden, Norway, Hungary (Watson 1934), also Russia (cf. Vygotsky 1994; Vygotsky and Luria 1994), and other countries (Takasuna, this issue). Yet we want to transcend the geographical distribution of ways of thinking and discuss the general schools of thought that are visible in world psychology. Relationships to particular countries are of secondary importance. So, it is possible to find different schools of thought in all countries where scientific study of psychology was and is being conducted.

The discussion here is historical and modern *Zeitgeist* of psychology rather than a question as to whether some scholars who do not fit into certain school can be found in some specific country or not. According to Watson (1934) there were eight differences between German–Austrian and North American psychological thinking:

1. German psychology was more interested in qualitative descriptions than in objective scores. It did not mean that German science did not collect objective scores. American psychologists did not go further after finding people with similar objective scores, whereas German psychologists asked more questions.
2. German psychology was interested in psychological controls rather than in physical controls. The situation that was objectively constant in terms of the room used, apparatus, time of the day, and so on, was regarded as desirable by more American experimenters. For German researchers the equality of the situation was not guaranteed by physical characteristics. Psychologically the situation could be very different for participants of a study even in physically perfectly similar situation.
3. German psychology was predominantly concerned with wholes and relationships. It was understood that every aspect of mind studied is part of the complex whole. The analysis was conducted at different levels of analysis from *Gestalten* as small as a cell to as great as the universe.

4. German psychology showed a greater concern with understanding a single case than the probabilities in a group. It did not mean that group data were not taken into account. For German psychology, for example, a correlation of 0.80 was psychologically unsatisfactory, unless the exceptions were understood, analyzed, and accounted for. No detail was unimportant to German psychologists. All exceptions had to be taken into account. According to Watson, “Our typical American dissertation in which one single procedure is carried out with large numbers of cases and then treated statistically seems to them to fall short of any genuine psychological contribution.” (p. 766).
5. German psychology saw individual trait differences as consequences of more basic type differences. They were less concerned with how well the subject can do on intelligence or achievement tests, but more concerned with how these reflect his *Weltanschauung*. Instead of asking, “What are his special abilities?”, German psychologists asked “What kind of a person is he?” American psychologists talked about individual differences in terms of units of a measuring scale. For German psychologists such measures were only *intermediate* steps toward the goal of describing the individual as belonging to a type. Greater unity of an individual was also assumed by German psychologists creating typologies. In addition to describing a person with a few “objective” measures, German psychologists took into account many other sources of information about an individual as a whole personality—handwriting, gestures, various literary products, diaries, letters, questionnaire answers. For German psychologists typological classification of a person did not replace unique characteristics of a person as a whole; typology complemented the unique side.
6. German psychology was more concerned with insight than with prediction. The American aim of science seemed to predict from situation “a,” the succeeding state of affairs “b.” This aim did not characterize German work. German’s were seeking the other aim, they wanted to understand situation by seeing it in its necessary relationship to a larger whole.
7. German psychology was more systematic. In German universities and institutes research was usually guided within one framework; less attention was paid to other schools of thought. In other words, instead of trying to incorporate all research results into a theory German psychology acknowledged that some facts may be irrelevant or useless. Watson described: “Facts with us often become ends in themselves; in German laboratories they are stones in a clearly outlined structure” (p. 767).
8. German psychology was more interested in thinking than in the accumulation of facts. So, for example, when Watson presented facts of an American study to a German seminar, he was asked, “Yes, and what of it? What does it mean?”

Taken together, the differences described by Watson point to the fact that German approach included all considerations important for North American thinking. However, there was more to this difference than is visible at the first instance—German psychology asked qualitatively novel kinds of questions *in addition* to those asked by its North American counterpart.

One Mainstream Psychology after World War II

Situation in Germany was rapidly changing when Watson wrote his article. Watson (1934; see also Ash 1999, for changes in German science before and after WWII) mentioned that most distinguished leaders of thought in psychology (Wertheimer, Stern, Werner, Katz, Peters, Koffka, Kronfeld, Lewin, etc.) were forced to leave their home country because of the Nazi regime. On the other hand, psychology remaining in Germany became politicized. Hitler's *Mein Kampf* was used by some to teach "political psychology," and psychological typologies were used to represent enemies as inferior, for example. After WWII, the situation continued to change. German–Austrian way of thinking disappeared from the forefront of research in psychology (see Ash 1992, for reception of Kurt Lewin's ideas in the United States as an example).

The aim of this paper is neither the analysis of the reasons nor description of the history of changes of world psychology after WWII. Rather, our goal is to understand better the current mainstream psychology. It is hard, if not impossible, to describe the current state of psychology in all its details. Nevertheless, the *Zeitgeist* of contemporary mainstream Anglo-American psychology can be characterized as being very similar to the pre-WWII North American psychology; the cultural tradition of German–Austrian psychology is lost from the mainstream for no scientific reasons.

Obviously, a personal impression of the current state of affairs is not sufficient to make any serious claims about current state of psychology. I analyzed the papers published in the most recent numbers of six leading journals inside different fields of psychology in order to get some support to my claims. The journals I used were: *Developmental Science* (2005, Vol. 8, No. 4), *Child Development* (2005, Vol. 76, No. 3), *Cognitive Psychology* (2005, Vol. 50, No. 3), *Journal of Experimental Psychology: General* (2005, Vol. 134, No. 2), *Journal of Personality and Social Psychology* (2005, Vol. 88, No. 6) and *Neuropsychologia* (2005, Vol. 43, No. 8). Overall, 55 articles were published in these specific journal issues.

Watson reported eight noteworthy differences between pre-WWII German–Austrian and North American psychology. Last of them, according to which German psychology was more interested in thinking than in the accumulation of facts obviously is very subjective. Even accumulation of facts requires thinking. So, the articles were analyzed only in respect to the seven remaining characteristics. Qualitative description of data, in addition to quantitative data, was provided in two articles (Hansen and Markman 2005; Newman 2005). Attempt to discuss whole person and relationships was made in one article (Brown and Bigler 2005). Two neuropsychological case studies were reported (Fujinaga et al. 2005; Marangolo et al. 2005), but these case studies were not connected with group data. No other attempts to analyze cases in addition to group data were reported. Type differences were proposed in four articles (Bonanno et al. 2005; Newman 2005; Poehlmann 2005; Troop-Gordon and Asher 2005). With one exception, only one of the seven differentiating characteristics could be found in each of the analyzed articles. Two characteristics, qualitative description of data and typology, was found in only one article (Newman 2005).

There seems to be no reason to suspect that these journals' issues selected misrepresent the current mainstream psychology. The results of the analysis suggest a quite a straightforward conclusion—*modern mainstream psychology can be understood as a direct continuation of pre-WWII North American psychology*. It can also be concluded that in certain sense modern mainstream psychology is more concerned with accumulation of facts than with elaborations of the theoretical meanings of these facts. In order to provide qualitative description of data in addition to quantitative data, to take into account both physical and psychological controls, to pay also attention to relationships and wholes, to be concerned about single cases that contradict group data, to take into account complex psychological types, to go beyond prediction, and to systematically relate data to complex theory—in all these tasks, considerably more focus on theoretical thinking would be required than in the case of not following the principles outlined by Watson.

Taken together it can be concluded that the whole set of theory-oriented principles of pre-WWII German–Austrian psychology is not followed by modern mainstream psychology. Next I discuss limitations of mainstream psychology that, theoretically, follow from ignoring of pre-WWII German–Austrian principles of methodological thinking.

Limitations of Contemporary Mainstream Psychology

Every of the principles followed by pre-WWII German–Austrian psychology and ignored by North American—and contemporary mainstream—psychology allows to understand the *psyche* from a qualitative and complex angle. Fundamental limits of scientific knowledge follow when German–Austrian methodological principles are not followed. Limitations related to each of the eight principles are discussed next. In discussion some examples are given. These examples were selected without a full representativeness of the field, similar problems characterize the mainstream psychology as a whole, not only theories and studies used as examples.

Objective Scores Without Qualitative Descriptions

Limiting research results to “objective” scores is related to implicit assumption that the same “objective” score can result only from one kind of psychological processes. If actually different mechanisms underlie the same score, and these differences are not studied, large part of mind remains beyond scientific inquiry.

There is evidence that the same objective results can be obtained by different mechanisms. It has been found, for example, that the arithmetic operation of addition can be performed by many different ways by children: counting fingers, using fingers without counting them explicitly, counting without using fingers, and counting up from the larger addend the number of times indicated by the smaller addend (Siegler 1996). It is important that the same “objective” correct result can be found by all these strategies with numbers smaller than ten. At the same time, some strategies, such as counting fingers, are considerably more limited than the others (counting using a number system, for example). Thus, the “same” result may not be the same psychologically in addition tasks.

Consider another example from personality research. Perhaps the most popular personality inventory of today is the Revised NEO Personality Inventory (NEO-PI-R, Costa and McCrae 1992). This inventory-like most others in personality psychology—contains self-descriptive items with no relationship to specific context. Yet particulars of situations are powerful determinants of human conduct. For example, the answer to the NEO-PI-R item “*I often feel helpless and want someone else to solve my problems*” (Vulnerability facet of Neuroticism, Costa and McCrae 1992) depends on the problem and on the context where the problem occurs. A person usually would not expect any help with a decision which kind of socks to buy at the store. But the same person would not expect to solve his problems by himself when lying on the road with several broken bones after the car crash. Actually, very specific details become relevant here. Which bones are broken? Is the person lying on the road because he is drunk or because he is hurt? Is the road in the rural or city area? If in the city, then is it in the “good” or “bad” neighborhood? And so on and so forth. Buying socks may also be a problem where help is needed when, for example, socks need to fit some formal costume. Does “black tie” mean that white socks are not allowed? Am I helpless when I cannot make this decision by myself? Practically every of the items in the NEO-PI-R may be answered differently even by the same person if different contexts were related to the item. In addition, there is a possibility that some respondents may answer in “abstract” ways, without connecting their answers to specific situations. The inventory does not give any hint regarding whether respondents tie answer to certain situations or not. Some may and some may not. In that case, can we suggest that decontextualized items have the same meaning as contextualized items? “Objective scores” from the NEO-PI-R and all similar inventories may actually cover rich individual differences that can never be discovered unless studies explicitly go beyond numbers and start to ask questions whether there are different ways to reach the same “objective” score or not.

Limiting Description of Experimental Conditions to Physical Controls

Within the behaviorist school of thought it would have been entirely legitimate to assume that only external physical conditions are related to the products of mental activity. For over more than half a century, the behaviorist thinking was replaced by cognitive psychology—supposedly because behaviorist stimulus → reaction psychology missed individual cognitive mechanisms of thought. Despite that not only modern mainstream cognitive psychologists, but also researchers in all other fields of psychology behave like behaviorists when they do not take into account the possibility that individual psychological aspect of mind is as important as physical external situation. Externally identical context may have very different meanings depending on the individual, and also depending on the current state of an individual.

In clinical psychology, for example, self-report measures are very often used for describing a patient’s state of mind. Such quantitative measures are used as if a patient is an inert subject and the researcher/clinician is not interested in the internal state of the patient. Considerably more complex picture emerges when those who fill in or interpret the tests are asked about their personal opinions about the tests. In a very interesting study by Blount et al. (2002) the authors asked ex-residents of a

hospital as well as professionals in therapeutic community and laypersons to think about the content and structure of several quantitative self-report measures. Many respondents, especially ex-residents found that questionnaires are depressing. Also, many items in questionnaires were reported to be vague or confusing—questionnaires contained slang, Americanisms (e.g., a study was conducted in the United Kingdom, where North American tests are not usually adapted for differences in language) or language that was not transparent for ethnic minorities; several questions were irrelevant for in-patients (e.g., work or family life questions, especially when anchored to “last 2 weeks”); several items included more than one question. Respondents also mentioned a state-bias—the tendency to give more optimistic answers in a good mood and vice versa. Results of this study suggest that psychological condition of testing may have profound influence on test results.

Culturally conditioned individual differences should also be taken into account. So, for example, the Western concept of personality may not be universal. For Western thought, the concept of personality is related to the notion of the individual as a distinct and bounded entity. This concept may be irrelevant for some non-western cultures (Marsella et al. 2000). Correspondingly, the same test may have fundamentally different meanings, or even be meaningless, depending on the cultural background of the individual in the study.

Thus, on the one hand, limiting analysis of experimental conditions to physical controls does not allow a researcher to understand whether the same “objective” score has been gained by the same underlying psychological mechanism or have the mechanisms been different. On the other hand, psychological mechanisms of intraindividual variation in the results of the same test cannot be understood also.

Fragments Without Whole and Relationships

Theoretically, every component in the whole contributes something unique to the whole. The properties of the whole, in turn, change depending on the number and quality of its components (cf., Koffka 1935; Köhler 1947; Vygotsky 1994). When mind is studied as a pile of independent fragments, the fundamental problem arises—effects can be attributed to wrong causes. Research allows studying only relationships between measured components. So every fragment study may only report a relationship but never a causal explanation because important causal variables may have been left out of the study.

Also, it is not taken into account in the contemporary mainstream psychology that not only the person *but also the test materials constitute a whole*. It has been found that as subjects move from the beginning to the end of a test, they become more consistent in their answering of the items. They tend to accentuate their initial scoring tendency, and their answers become more accurate predictors of the total score (Knowles 1988; Knowles and Byers 1996). Item context in a test has been found important in determining test results also in other studies (Steinberg 1994, 2001). Thus, time dimension in the whole research situation is important too. If the results on some test depend on so many aspects of the whole psychological situation, it may be misleading, again, to assume that “objective” test scores can be interpreted unambiguously. In case of many tests, such as personality questionnaires, it is assumed that the test results reflect some preexisting underlying structure of mind.

This assumption may be wrong. Abovementioned studies where item context effects have found suggest that structure may be created in time of the testing situation rather than just reflected in the final scores.

Another example can be found in studies of cognitive ability. In interpretation of a cognitive ability or IQ test it is usually (and very often only implicitly) assumed that results of such tests reflect cognitive abilities alone. The possibility that human mind is whole and all its aspects may be reflected in test results is ignored. However, as demonstrated by Richardson (2002), variance in IQ scores can result from many non-cognitive factors such as the extent to which people of different social classes and cultures have acquired a specific form of intelligence, academic orientation and self-efficacy beliefs, test anxiety, self-confidence and other social–cognitive–affective factors. Again, interpreting IQ as characteristic of cognitive ability alone can be very misleading; IQ may reflect many additional sides of mind in addition to problem-solving ability per se. If so, then many lines of cognitive ability research—such as studies of genetic mechanisms of intelligence, where intelligence is practically always defined as a result of some cognitive ability test—may turn out to be ill-defined.

Probabilities in a Group Without Additional Case Studies

Studies where results are limited to only group differences and/or correlations between variables cannot go beyond probabilistic prediction. The nature of a relationship or a difference between groups cannot be understood unless all exceptions are understood. So, for example, all reports of sex differences, race differences, or differences between cultures should not be attributed to sex, race, or culture respectively, unless the exceptions are addressed.

In studies of sex differences in cognitive abilities very general mechanisms are usually proposed for “explaining” why men outperform women in mathematical reasoning ability, for example. There are possible hormonal mechanisms, brain mechanisms, and social mechanisms that may lead to sex differences in cognitive abilities (e.g., Benbow 1988, and commentaries on her target article; Kimura 1999). At the same time, there are data from very large studies on sex differences in mathematical abilities, which clearly demonstrate: sex differences are exceptions to overall similarities between men and women! The ratio of boys to girls scoring above the mean on the Mathematics Scholastic Aptitude Test is only about 1.5 to 1 whereas in top 1 in 10,000 group the ratio of boys to girls is about 13 boys for one girl (Benbow 1988; Benbow and Stanley 1983). Effect size for group difference in Benbow’s studies is only about a half of a standard deviation suggesting that there are at least some girls who perform better than almost all of the boys on SAT-M tests. Even smaller overall effect size (about one third of a standard deviation) with increasing proportion of men with increasing IQ has been found in a recent study by Irwing and Lynn (2005). Thus, the data suggest that what is actually needed is not a theory of sex differences in cognitive abilities in general but a theory for explaining why there are more exceptional men than women. General factors such as brain size or hormonal differences should apply to all men and women—these factors apparently are not the best candidates for explaining relatively rare exceptions.

Another problem of group data analysis is related to the fact that the same mean levels may result from different distributions of data. So, for example, in a study of short-term memory performance in patients with traumatic brain injury we found that healthy subjects outperform persons with brain damage in tests where words had to be recalled. There was no group difference in recall of non-words. Further case-by-case inspection of the data revealed, however, that in recall of non-words some brain-damaged persons actually had supra-normal recall level! In the control group only 2% of participants recalled more than 5 non-words out of 16; in a group of participants with brain damage 25% of a group recalled six, seven, or even eight non-words. At the same time, the number of patients with very low scores was also bigger in the brain injury group. Correspondingly, the mean level of performance was not different between groups (Toomela et al. 1999). Without studying individual scores, supra-normal performance in a substantial subgroup of brain-injured people would not have been discovered.

In addition to specific questions it is important that ignoring individuals in research has fundamental problems. Results of studies based on data about variation between cases—which is the focus of contemporary mainstream psychology—cannot be extended to individual cases because psychological processes are not ergodic. Psychology must study individuals changing in time in order to understand minds of individuals (Molenaar 2004; Molenaar and Valsiner 2005). Psychology's aim should be understanding the mind. If, as it turns out, contemporary mainstream psychology studies groups so that results of research cannot meaningfully be extended to understanding individuals then the question emerges, what we actually understand through this kind of research. Though socially culturally conditioned, mind still is always an individual phenomenon at the same time.

We have to conclude that contemporary mainstream psychology has left out of research the main object of study. The German–Austrian comment on a “typical American dissertation” made 70 years ago is still relevant:

our typical contemporary mainstream psychology group study in which one single procedure is carried out with large numbers of cases and then treated statistically seems to fall short of any genuine psychological contribution (cf. Watson 1934: 766).

Individual Trait Differences Isolated from More Basic (Weltanschauung) Type Differences

If individuals will turn into a focus of psychology, typology must become a necessary component of research and theory-building. The problem with idiographic (individual-focused) approach is that the number of possible models describing an individual is potentially very large. Every individual is in some sense unique; unique models may also be necessary for every different situation, setting, developmental phase, historical period, even a mood swing. It is plausible that such almost unlimited number of models can be meaningfully condensed into considerably smaller number of types and antitypes (Curran and Wirth 2004; von Eye 2004).

The idea that persons can be distributed into types implies that human psychological traits do not vary in isolation (Watson 1934). In contemporary

mainstream psychology the default factor analytic procedure for analyzing results of different tests and inventories is principal component analysis with orthogonal rotation of data. It is assumed that supposedly “independent” hidden dimensions in test performance are revealed through this kind of data analysis. This approach is based on analysis of linear relationships between variables. One of the founders of such approach, Thurstone, was aware of the limitations of linear orthogonal factor analysis. He mentioned,

There is an unlimited number of ways in which nature can be comprehended in terms of fundamental scientific concepts. *One of the simplest ways* in which a class of phenomena can be comprehended in terms of a limited number of concepts is probably that in which a linear attribute of an event is expressed as a linear function of primary causes (Thurstone 1935: 47, emphasis added).

He also claimed that orthogonality does not correspond to the real world: “In one sense, no two observations can ever be experimentally independent. The term can be used only with reference to the state of knowledge at the time the observations are made” (p. 50). Thurstone also warned that factor analytic approach is far from ideal science:

It would be unfortunate if some initial success with the analytical methods to be described here should lead us to commit ourselves to them with such force of habit as to retard the development of entirely different constructs that may be indicated by improvements in measurement and by inconsistencies between theory and experiment (p. 54).

This message from one of the founders of factor analysis seems to be forgotten now in mainstream psychology. The methodological way of thinking that relies on factor analysis with orthogonal rotation suppresses hypotheses about possible types of persons with nonorthogonal traits. Some typological research can be found in the mainstream as also suggested by the analysis of recent published articles in the leading psychology journals. Nevertheless, the human mind may be considerably more integrated than mainstream cognitive or personality psychology implies.

Prediction Without Insight

Prediction is a complex concept with different meanings. In research, experimenters conduct studies in order to establish whether data come out as theoretically predicted. In more specific sense, science predicts from situation “a” the succeeding state of affairs “b.” These two ways of prediction are interacting. Some researchers constrain the notion of “theory” to the pattern of probabilistic relationships between variables; theory predicts whether some situation “a” (test result, concentration of a neuromediator in the brain or in the blood, etc.) predicts (usually—is correlated) with some other situation “b.” For other researchers it is necessary to go beyond such limited understanding of theory and involve non-quantitative information into theories.

In contemporary mainstream psychology theories are very often descriptions of quantitative data-structures without qualitative perspectives, without “insight.” If quantitative prediction is the final goal of science, then this goal is very limiting.

Phenomena can be understood by far more deeply than any correlation-based prediction allows. Even more, “quantitative” theory may simply be inappropriate for understanding many phenomena in the behavioral sciences.

For example, students in their first years of studies of psychology have already been taught that any correlation cannot be taken as evidence for a causal relationship. Variable *A* can affect variable *B* or variable *B* can affect variable *A*. The relationship may also be caused by a third variable not taken into account in the study. Thorough analysis of the whole situation seems to be the only way in which it is possible to understand whether the third variable has been involved and if yes then what that third variable might be.

Routine utilization of factor analysis techniques in research can be an example here. In factor-analytic research it is assumed that mind is a collection (not a system, because a system would imply reciprocal influences between its components, cf. von Bertalanffy 1968) of “independent” components. According to such approach to research, these independent components of mind can be “discovered” by factor analysis. So, we can discover factor “*g*” or multiple factors in intelligence; 3, 5, or 16 factors of personality; four factors of aggression; and so on and so forth. In such research it is assumed that correlational ties between responses to certain items are caused by the third variable, the directly unobservable component or dimension of the mind. However, correlation, that underlies factor analysis, contains very little, if at all, information about the dimension—about whether there is one underlying cause that makes two variables correlated (Essex and Smythe 1999). Even more, very often it is questionable whether psychological attributes themselves are quantitative at all. It is possible that quantitative structure is imposed to qualitative psychological phenomena. In such cases quantification of data and the following utilization of quantitative data analysis methods is misleading and may end up with numbers that “are not just theory-free numbers, they are meaning-free numbers” (Essex and Smythe 1999: 746; see also Hoshmand 2003; Martin 2003; Michell 2000, 2003a,b, 2004, for thorough theoretical discussion of this issue).

For another example, consider the questions of effect. Contemporary mainstream psychology often asks the questions like “does the psychotherapy work?”, “does Head Start elevate scholastic performance?”, “does frustration cause aggression?” Answers to such questions are usually given as predictions—if a patient belonged to a group that participated in sessions of psychotherapy, if a child participated in Head Start classes, if a person belonged to a group with relatively high level of frustration then, respectively, the group where a patient, a child, a person belonged will have lower levels of depression according to some test or higher results in academic performance or higher level of aggression. Two problems are related to such predictions. One, already discussed above, is related to ignoring individual cases. The effects do not usually apply to all persons in groups that are compared by statistical data-analysis procedures (see also Sohn 1999). The other problem with probabilistic prediction can be understood as a lack of insight. Researchers do not realize that in causal relationships necessary is not sufficient. For example, there is evidence that a right drug can reduce the average level of depression. But “there is equally solid evidence that some patients respond better to one or another of the available therapies rather than drugs, or—hardly surprising—to a combination of drugs and psychotherapy” (Notterman 2004: 257). Again, quantitative prediction

turns out to be limited and potentially even misleading way for understanding psychological phenomena.

Nonsystematic Approach to Research and Theory-Building

In principle, it would be possible to build a systemic theory from fragments collected in relatively unrelated studies. The number of fragmentary facts collected in mainstream psychology over the last half of the century is huge. And still there is no systematic theory trying to unite these fragments into a theory of a person as a whole. This theory, of course, may turn out to be wrong. But without such theory psychology is doomed to endless trial-and-error search for more and more relationships between variables. There is no guiding principle that would direct attention to variables that should be taken into account in building a theory. Even more, it should be asked whether there are facts that should be ignored as irrelevant for understanding the phenomena under study. Several examples discussed above are also relevant here. We need to ask whether “objective” scores give us sufficient information or should we go beyond these numbers and ask whether there are different ways for ending up with similar numbers. If the latter is true, the “objective” scores turn out to be quite useless. We need to ask whether it is appropriate to use nomothetically obtained constructs for describing individuals (cf. Molenaar 2004; Molenaar and Valsiner 2005). We need to ask whether the phenomena we study are really quantitative, otherwise all quantitative data analysis results are just a bunch of meaningless numbers (Essex and Smythe 1999; Michell 2000, 2003b). As discussed above, we may have to conclude that we should actually ignore large amount of data collected over decades of psychological research because there is a lot of information that is incomplete, incoherent, misleading, or simply irrelevant.

One more specific example can be provided here. Contemporary mainstream psychology often approaches study of mind operationally; research procedures are used not to study a phenomenon but to create a phenomenon that is studied in its own rights (see for recent discussion of the issue Grace 2001, and comments to his article). Intelligence, for example, seems to be defined often as something that an intelligence test measures (cf. Mackintosh 1998). The situation is actually even worse. If intelligence is defined then it turns out that the defined phenomenon is a labeled collection of phenomena that cannot be quantified! Several other phenomena, such as personality, attitudes, achievement, motivation, learning, development, memory, cognition, are other examples of labeled collective phenomena that cannot also be quantified (Essex and Smythe 1999). Thus, in these cases the operationally defined phenomenon turns out to be a phenomenon that cannot be operationalized as something that can be measured with a quantitative test or a questionnaire. Theory-free operationism creates data that may be completely irrelevant for understanding mind.

Accumulation of Facts Without (Complex) Thinking

There is very little to add to the points already made. Obviously setting up even the simplest study to discover another fact requires thinking. The problem is that

modern mainstream psychology does very little to go beyond certain kinds of facts. It seems that one important question has not been asked. Are our facts good for understanding mind? The answer seems to be no. Facts that are isolated, out of the whole context, or facts that describe group differences but not exceptions are no good to understand mind. So, even if sometimes complex theories are built from facts modern mainstream psychology has collected, the quality of facts is not sufficient for developing a high-quality theory. Much more thinking is needed to improve the quality of the basic building blocks of a theory, the scientific facts.

Conclusions

Watson (1934) concluded his article where he compared two methodological ways of thinking:

The eight major trends previously noted continue. German psychology is still more thoughtful, more qualitative, more subjective, more concerned with wholes, more insistent on understanding the particular case, more apt to make typological and characterological studies, more interested in achieving insight, more concerned with schools and systems. Although many brilliant psychologists have left voluntarily or been forced to leave, this cultural tradition is likely to continue for many years (p. 771).

It turns out that Watson was wrong, the German–Austrian cultural tradition did not continue, at least not in the mainstream psychology. Contemporary mainstream psychology follows the traditions of pre-WWII North American psychology. Considerably more insightful methodological principles of German–Austrian psychology have moved into periphery of psychological thought.

Last 60 years in psychological research have given us thousands, perhaps even millions, of ways how to predict statistically one psychological variable by way of another. At the same time, many fundamental questions have even not been asked because of limited methodological thinking. We still find “objective” scores without knowing how many different psychological mechanisms may underlie the same score. We do not know how psychological aspect of experimental conditions may have contributed to study results. Study of fragments gives very little to understanding a human person as a whole. Group comparisons, almost with no exceptions, seem to have more exceptions than cases that follow the “rule.” Still these “rules” or “laws” are taken as evidence for “explanation,” as if belonging to a group would really explain individual psychological differences. Reliance on linear statistical data analysis procedures in too many cases rule out the possibility for discovering person types and other more complex aspects of mind. Statistical probabilistic prediction has become an end goal of studies even though most of the thinking and insight should begin where the science of mainstream psychology seems to end now. Fragmentary theories do not allow us to create complex hypotheses about relationships between interacting aspects of the human personality. We have too many facts published in “leading” journals that may simply be inappropriate or irrelevant for theory-building. Overall, modern mainstream

psychology, indeed, is satisfied with accumulation of facts instead of thinking. Last 60 years of research in psychology seems to have gone astray.

Acknowledgment This work was supported by Estonian Science Foundation Grant No. 5388.

References

- Ash, M. G. (1992). Cultural contexts and scientific change in psychology. Kurt Lewin in Iowa. *American Psychologist*, 47(2), 198–207.
- Ash, M. G. (1999). Scientific changes in Germany 1933, 1945, 1990: Towards a comparison. *Minerva*, 37, 329–354.
- Benbow, C. P. (1988). Sex differences in mathematical reasoning ability in intellectually talented preadolescents: Their nature, effects, and possible causes. *Behavioral and Brain Sciences*, 11, 169–232.
- Benbow, C. P., & Stanley, J. C. (1983). Sex differences in mathematical reasoning ability: More facts. *Science*, 222, 1029–1031.
- 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.
- Bonanno, G. A., Rennie, C., & Dekel, S. (2005). Self-enhancement among high-exposure survivors of the September 11th terrorist attack: Resilience or social maladjustment? *Journal of Personality and Social Psychology*, 88(6), 984–998.
- Brown, C. S., & Bigler, R. S. (2005). Children's perceptions of discrimination: A developmental model. *Child Development*, 75(3), 533–553.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five-factor inventory (NEO-FFI)*. Professional manual. Odessa, FL: Psychological Assessment Resources.
- Curran, P. J., & Wirth, R. J. (2004). Interindividual differences in intraindividual variation: Balancing internal and external validity. *Measurement*, 2(4), 219–227.
- Essex, C., & Smythe, W. E. (1999). Between numbers and notions. A critique of psychological measurement. *Theory and Psychology*, 9(6), 739–767.
- Fujinaga, N., Muramatsu, T., Ogano, M., & Kato, M. (2005). A 3-year follow-up study of "orientation agnosia." *Neuropsychologia*, 43, 1222–1226.
- Grace, R. C. (2001). On the failure of operationism. *Theory and Psychology*, 11(1), 5–33.
- Hansen, M. B., & Markman, E. M. (2005). Appearance questions can be misleading: A discourse based account of the appearance-reality problem. *Cognitive Psychology*, 50, 233–263.
- Hoshmand, L. T. (2003). Can lessons of history and logical analysis ensure progress in psychological science? *Theory and Psychology*, 13(1), 39–44.
- Irwing, P., & Lynn, R. (2005). Sex differences in means and variability on the progressive matrices in university students: A meta-analysis. *British Journal of Psychology*, 96(4), 505–524.
- Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: Bradford Book.
- Knowles, E. S. (1988). Item context effects on personality scales: Measuring changes the measure. *Journal of Personality and Social Psychology*, 55(2), 312–320.
- Knowles, E. S., & Byers, B. (1996). Reliability shifts in measurement reactivity: Driven by content engagement or self-engagement? *Journal of Personality and Social Psychology*, 70(5), 1080–1090.
- Koffka, K. (1935). *Principles of Gestalt psychology*. London: Routledge & Kegan Paul.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Mentor Books.
- Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: Oxford University Press.
- Marangolo, P., Piras, F., & Fias, W. (2005). I can write seven but I can't say it: A case of domain-specific phonological output deficit for numbers. *Neuropsychologia*, 43, 1177–1188.
- Marsella, A. J., Dubanoski, J., Hamada, W. C., & Morse, H. (2000). The measurement of personality across cultures. *American Behavioral Scientist*, 44(1), 41–62.
- Martin, J. (2003). Positivism, quantification and the phenomena of psychology. *Theory and Psychology*, 13(1), 33–38.
- Michell, J. (2000). Normal science, pathological science, and psychometrics. *Theory and Psychology*, 10(5), 639–667.

- Michell, J. (2003a). Pragmatism, positivism and the quantitative imperative. *Theory and Psychology*, 13(1), 45–52.
- Michell, J. (2003b). The quantitative imperative. Positivism, naive realism and the place of qualitative methods in psychology. *Theory and Psychology*, 13(1), 5–31.
- Michell, J. (2004). Item response models, pathological science and the shape of error. *Theory and Psychology*, 14(1), 121–129.
- 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>.
- Newman, D. L. (2005). Ego development and ethnic identity formation in rural American Indian adolescents. *Child Development*, 76(3), 734–746.
- Notterman, J. M. (2004). Persistent conceptual issues in psychology. A selective update. *Theory and Psychology*, 14(2), 239–260.
- Poehlmann, J. (2005). Representations of attachment relationships in children of incarcerated mothers. *Child Development*, 76(3), 679–696.
- Richardson, K. (2002). What IQ tests test. *Theory and Psychology*, 12(3), 283–314.
- Siegler, R. S. (1996). *Emerging minds. The process of change in children's thinking*. New York: Oxford University Press.
- Sohn, D. (1999). Experimental effects. Are they constant or variable across individuals? *Theory and Psychology*, 9(5), 625–638.
- Steinberg, L. (1994). Context and serial-order effects in personality measurement: Limits on the generality of measuring changes the measure. *Journal of Personality and Social Psychology*, 66(2), 341–349.
- Steinberg, L. (2001). The consequences of pairing questions: Context effects in personality measurement. *Journal of Personality and Social Psychology*, 81(2), 332–342.
- Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits*. Chicago, IL: University of Chicago Press.
- Toomela, A., Tomberg, T., Orasson, A., Tikk, A., & Nõmm, M. (1999). Paradoxical facilitation of a free recall of nonwords in persons with traumatic brain injury. *Brain and Cognition*, 39, 187–201.
- Troop-Gordon, W., & Asher, S. R. (2005). Modifications in children's goals when encountering obstacles to conflict resolution. *Child Development*, 76(3), 568–582.
- von Bertalanffy, L. (1968). *General systems theory. Foundations, development, applications*. New York: George Braziller.
- von Eye, A. (2004). The treasures of Pandora's box. *Measurement*, 2(4), 244–247.
- Vygotsky, L. S. (1994). The problem of the cultural development of the child. (Originally published in 1929). In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 57–72). Oxford: Blackwell.
- Vygotsky, L. S., & Luria, A. (1994). Tool and symbol in child development. (Originally written in 1930). In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 99–174). Oxford, UK: Blackwell.
- Watson, G. (1934). Psychology in Germany and Austria. *Psychological Bulletin*, 31(10), 755–776.

Aaro Toomela is a Professor of Special Education at the University of Tartu, Estonia. He received his MD from the University of Tartu in 1986. After spending seven 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. Last five years Aaro Toomela has been back to the academic world working in the field of (special) education in the University of Tartu and in the field of psychology in the University of Tallinn (Estonia).

Copyright of Integrative Psychological & Behavioral Science is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.