

Christian-Albrechts-Universität zu Kiel

Nr. 17

Monographien zur Prüfungsforschung

Herausgegeben vom
Projekt Prüfungsforschung
am Institut für Pädagogik

Projekt Prüfungsforschung
am Institut für Pädagogik
der Christian-Albrechts-Universität zu Kiel (Hg.)

Monographien
zur Prüfungsforschung
Heft 17

Peter Krope, Boleslaw Niemierko

Educational Measurement in Germany and Poland

For copies:

Peter Krope
D: 24118 Kiel
Olshausenstraße 75
Tel. (0431) 8801273
FAX (0431) 8801588
E-Mail: krope@ewf.uni-kiel.de

Kiel 2001

CONTENTS

Peter Krope	4
Educational Diagnostics in Germany. From Latent Dispositions to Manifest Behaviour: A Fundamental Naturalistic Mistake in Modern Test Theories?	
Boleslaw Niemierko	13
Educational Diagnostics in Poland. Diagnostic Studies as a Kind of Educational Applied Research	
The Authors	22

Peter Kroppe

EDUCATIONAL DIAGNOSTICS IN GERMANY

FROM LATENT DISPOSITIONS TO MANIFEST BEHAVIOUR: A FUNDAMENTAL NATURALISTIC MISTAKE IN MODERN TEST THEORIES?

1 INTRODUCTION

In his textbook INGENKAMP (1988) gives a detailed general view of educational measurement in Germany up to the eighties. His description can be completed with an important detail. Most German educationalists prefer the humanities (Geisteswissenschaften). That means they avoid to use explicit empirical methods. INGENKAMP (1989, pp. 153 – 179) speaks of an aversion of German intellectuals to test methods.

To simplify matters a quarrel can be stated between two scientific positions: The humanities and the natural sciences. In general the quarrel remains on the level of procedures. A theoretical debate is seldom. Thus educational measurement in Germany looks like a patchwork of divergent methods.

This private view is the background for my question in the following contribution¹.

2 THE PROBABILISTIC TEST THEORY

Educational measurement depends on a theory. Someone, who intends to produce empirically true statements about persons, has to rely on a theory for the application or development of tests, questionnaires, and similar procedures. At present, two relevant test theories are available. They are the classical and the probabilistic test theory. The problem dealt with in this presentation could be

¹ This lecture is the abridgement of an unpublished article (Peter Kroppe, Peter Petersen: Generic Acts. The Constructivistic Contribution to the Solution of the Body-and-Mind Problem).

demonstrated by either of them. Because of its clearer representation, the probabilistic test theory is chosen here.

The general starting-point of the probabilistic test theory reads as follows:

$$P(a_{vi}) = f(V_1, V_2, \dots).$$

This formula means that the probability P of a particular reaction a of the person v to the item i is a function of the latent variables V_1, V_2, \dots (ROST and SPADA 1978, p. 63/64). In the formula something manifest is mentioned indirectly, that is to say the reaction of a person. In the context of multiple-choice items, it is the behaviour of choice in ticking boxes. In the formula something latent is directly mentioned which is essentially identical with dispositions. A simplified explanation suggests that the probability of the (manifest) behaviour is explained by (latent) dispositions (ROST 1995).

Let me explain this theory with an example. If you want to measure anxiety, you can start with the observation of a physical manifestation of anxiety like pulse rate. According to a six step model of testing, you can make a statement on the basis of this not only concerning the manifestation itself but also the disposition to be anxious (see fig. 1).

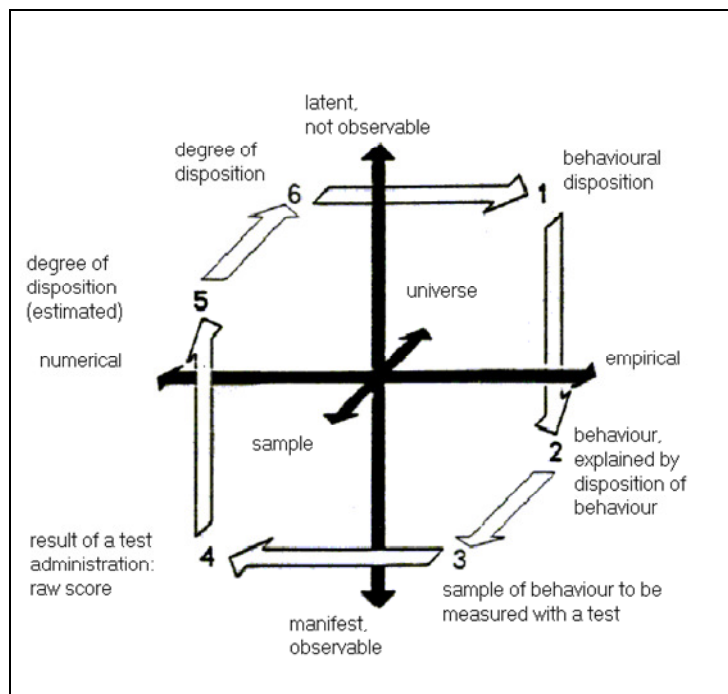


Fig 1.: Measuring a disposition in six steps. According to ROST & SPADA (1978).

Back to theory! Theoretically, we can assume that the probability of manifest facts may be explained entirely by identifying a finite number of latent variables. This is based on models, which

axiomatically describe the connection between the manifest and the latent parameters (SCHEUCH and ZEHNPFENNIG 1974, p. 134).

An axiomatic description is a problematic base for methods which are intended to give information about individuals. In fact, axiomatic theories have no empirical relevance at all, in so far as their basic symbols are not semantically interpreted (BUNGE 1967). But if the elements of an axiomatic system cannot be interpreted according to rules, the necessary condition for the validation of empirical statements about individuals is no longer applicable.

In analysing latent structures, probabilistic test theory may overcome the instructions of extreme behaviourism, which prescribes the observation of external behaviour as opposed to internal motives (WATSON 1966). These models connect external with internal dimensions in those tested. But the answers to the question regarding the type of relation between the manifest and the latent dimension, are nothing less than the correctly formed logical sentences in an axiomatic theory. The question of their validity may only be answered within the framework of the mathematical description of the test method.

3 THE BODY-AND-MIND PROBLEM

This test theory divides its subject into two entities, and this has a parallel in the body-and-mind debate of philosophical tradition. Modern philosophy starts in the 17th century with the theory of DESCARTES. He separates the world into an area of the mind (*res cogitans*) and one of the matter (*res extensa*) (compare figure 2).

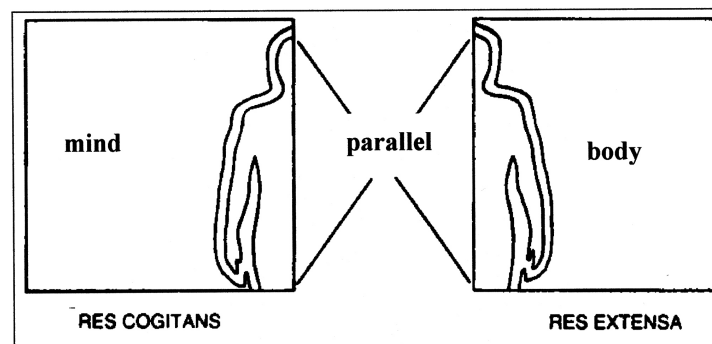


Fig.2: The body-and-mind theory in dualism of DESCARTES. According to KUNZMANN et al. (1992, p. 106).

The dualism of DESCARTES is contradicted by the monism of SPINOZA. According to SPINOZA mind and body are not two separate entities, but two sides of the same substance, which is God (compare figure 3).

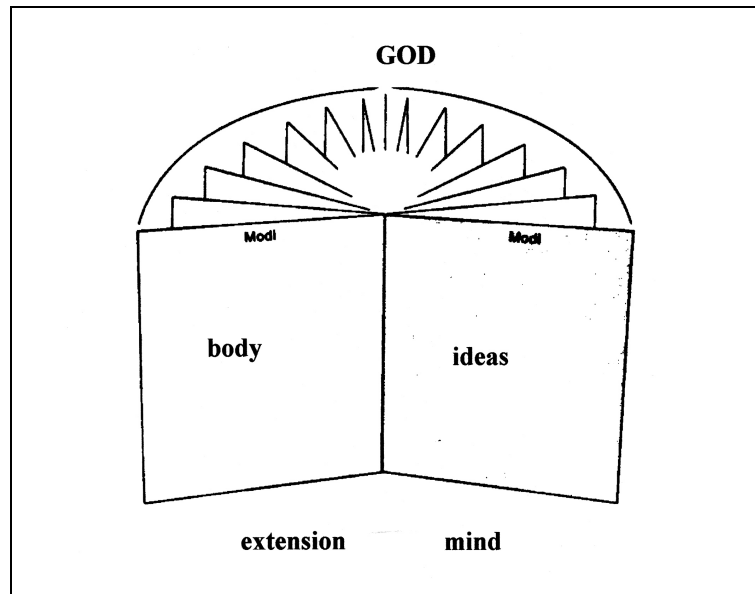


Fig.3: The body-and-mind theory in monism of SPINOZA. According to KUNZMANN et al. (1992, p. 108).

Finally the theory of LEIBNIZ should be mentioned. LEIBNIZ describes the meeting of certain physiological with certain psychological problems as a previously arranged (prestabilised) harmony, containing all event occurring in the world (Compare figure 4). He himself illustrates his theory with the image of two previously arranged clocks.

The question, as to how mind and body – for example the feeling of anxiety and its physical effects - are connected would be answered by each philosopher differently. According to DESCARTES they would be clearly separated, to SPINOZA different aspects of the same substance, and to LEIBNIZ parallel proceedings. But according to POPPER and ECCLES no attempt to solve the problem caused by the search for a relation between body and mind can be proven empirically (POPPER and ECCLES 1982).

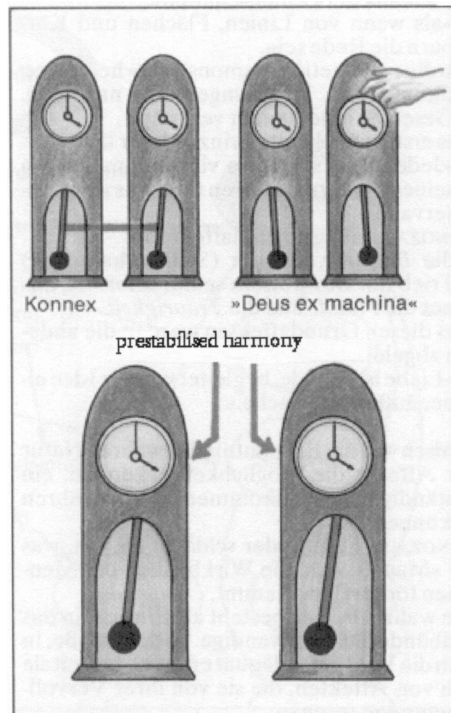


Fig.4: The body-and-mind theory in the philosophy of LEIBNIZ. According to KUNZMANN et al. (1992, p. 112).

4 A CONSTRUCTIVISTIC ANALYSIS

In the following an analysis of the body-and-mind problem is introduced. Its basis is the Methodical Constructivism (the Constructivism of the Erlangen School, cf. HARTMANN 1998, JANICH 1996, KAMLAH and LORENZEN 1973, KROPE 2000, KROPE et al. 2000, LORENZEN 1987).

In order to carry out the analysis the distinction between acts (Handlungen) and generic acts (Handlungsschemata) is explained. Examples of acts are to go for a walk, to close a window, to write a letter, to ride a bike. Opposite examples are to stumble, to be startled, to turn white. An act is described in constructivism as an action, which may be invited in which case it is irrelevant whether another person has elicited it or whether the subject elicits it himself. You can ask someone to go for a walk but not to give themselves a fright. Being startled is something unintended, something which may happen to a person, an unintended experience (Widerfahrnis).

In order to introduce the term “generic act” we will look at the following example. After learner drivers have collided several times while turning off on a traffic training ground, they may make an agreement after an intervention by the driving instructor. They say: „From now on when we want to turn, we will indicate the direction in which we want to turn by stretching out an arm in that

direction“. If the youngsters keep to their agreement, an observer will be able to determine individual differences in their acts. Some students will stretch their arms out straight, others in a bent fashion, some may move their arms up and down, some are brash and others are hesitant. But all these acts are interchangeable and cause no change in the agreed meaning of the gesture. In looking at to examples like these we may say: If only such statements can be made about individual acts which are also valid for all acts (in all relevant respects) which are described in the same way, then we are speaking about a generic act. We can generalise this in the following way: If we make statements about acts, which do not vary in relation to the incidental characteristics of these acts, we can speak of these acts as generic acts. The transition from an actual act to its generic act is an abstraction.

We distinguish the actual act (Frerk stretches out his left arm now on the traffic training ground) from the potential act (which is the generic act agreed upon in order to avoid collisions). Accordingly, we may say if a person answers the question “What do you think about commented educational reports?“ in the questionnaire with “I approve of them“ and “I appreciate them“, an actual act is observed. If we assign the predicator “agreement“ to this answer we are talking about a generic act. The statement that this person has answered “I approve of them“ and “I appreciate them“ is an empirical statement. And the statement that this represents an agreement, reminds us of a predicator rule, which has been explicitly agreed upon previously.

To note that it might be possible to regard the update of an act and a generic act as the empirical or the rational part of an act, may be helpful to our understanding. However, it may cause an epistemological dilemma if it leads to discussion of two independent entities in a tradition marked by the probabilistic test theory. This is then called the body-and-mind problem.

An analysis supplied by the philosopher DIRK HARTMANN (1998) draws our attention to such a danger. HARTMANN starts out from the program laid down by the Constructivism of the Erlangen School. In it the living world (*lebensweltlicher Alltag*) is the systematic starting point of science on the one hand. On the other hand he sees science in its function of supporting the practice of the living world, and thus returning to it. For physics HARTMANN shows how a realistic interpretation of science during its history has lead to a reification (*Vergegenständlichung*) of scientific constructs. He stresses that “the interplay of theoretical differentiation and integration has lead to a continuing detachment of the natural sciences from the specific practices and purposes of the historical beginnings. All of this resulted very early on in overlooking the fact that it is the

epistemological interest of the practice support which supplies the criteria for the theories of the natural sciences. This way the realistic opinion of correspondence of theory and reality could prevail“ (HARTMANN 1998, p. 325). Via ontological hypostasis – according to HARTMANN – the body-and-mind problem develops in the natural sciences.

A reification of the theory as described by HARTMANN for the natural sciences, may be observed for probabilistic test theory as well. As showed in the beginning, its general approach claims that the probability of a particular reaction of a person to an item is a function of latent variables. Now, a function is defined as “a unique representation of two sets, i.e. a certain set of organised pairs $[x; y]$ with the characteristic that for every x an y has been assigned” (GELLERT et al. 1972, p. 126). The graphical representation of a function (comp. figure 5) makes the use of the two separate areas clear - a definition area and a value area. For probabilistic test theory these are the areas of behaviour patterns and of the dispositions to behave which explain these behaviour patterns. This is the theoretical construct for which universal validity has been claimed since the establishment of latent structure analysis approach by LAZARSELD (for example LAZARSELD 1950).

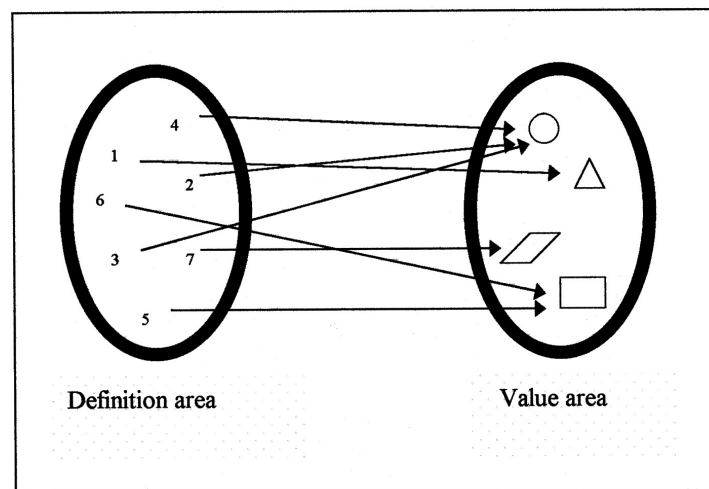


Fig.5: Graphic representation of a function. According to GELLERT et al. (1972, p. 126).

In constructive educational measurement the epistemological dilemma described as the body-and-mind problem is avoided by introducing the term generic act, which prevents an equation with the term behavioural disposition. Behavioural dispositions are characteristics in a wider sense whereas a generic act may be understood in the diagnostic context as a model for the explanation which an observer constructs for himself. A generic act represents in empirical surveys the anticipation of an observer of how a person will react to a given stimulus. In the Constructivism of the Erlangen

School the objects of science are “constructs, i.e. products of intentional human acts“ (GETHMANN 1996, p. 746).

5 References

- BUNGE, M.: Scientific Research. Vol. I. Berlin etc. 1967.
- GELLERT, W., et al.: Kleine Enzyklopädie Mathematik. Frankfurt/M. and Zürich 1972.
- GETHMANN, C. F.: Konstruktive Wissenschaftstheorie. In: J. MITTELSTRAß (ed.): Enzyklopädie Philosophie und Wissenschaftstheorie. Volume 4. Stuttgart/Weimar 1996, p. 746 - 758.
- HARTMANN, D.: Philosophische Grundlagen der Psychologie. Darmstadt 1998.
- INGENKAMP, K.: Lehrbuch der pädagogischen Diagnostik. Weinheim/Basel 1988.
- INGENKAMP, K.: Diagnostik in der Schule. Beiträge zu Schlüsselfragen der Schülerbeurteilung. Weinheim/Basel 1989.
- JANICH, P.: Konstruktivismus und Naturerkenntnis. Frankfurt/M. 1996.
- KAMLAH, W., LORENZEN, P.: Logische Propädeutik. Vorschule des vernünftigen Redens. Mannheim etc. ²1973.
- KROPE, P.: Konstruktive Pädagogische Diagnostik. Münster etc. ²2000.
- KROPE, P., NIEMIERKO, B.: (ed.): Proceedings of the 1992 Workshop on School Examinations. In: Projekt Prüfungsforschung (ed.): Monographien zur Prüfungsforschung. 15. Kiel 1992.
- KROPE, P., with the co-operation of O. ALBERS, B. FRIEDRICH, A. KOHRS, P. PETERSEN, M. VESPER: In acht Schritten zum Modernen Berichtszeugnis. Münster etc. 2000.
- KUNZMANN, P., BURKARD, F.-P., WIEDMANN, F.: dtv-Atlas zur Philosophie. Tafeln und Texte. München ²1992.
- LAZARSFELD, P.: The Logical and Mathematical Foundation of Latent Structure Analysis. In: S. A. STOUFFER et al. (ed.): Studies in Social Psychology in World War II. Vol. IV. Princeton/N.J. 1950, pp. 362 - 412.
- LORENZ, K.: Handlungsschema. In: J. MITTELSTRAß (ed.): Enzyklopädie Philosophie und Wissenschaftstheorie. Volume 2. Mannheim etc. 1984, p. 37.
- LORENZEN, P.: Lehrbuch der konstruktiven Wissenschaftstheorie. Mannheim etc. 1987.
- POPPER, K. R., ECCLES, J. C.: Das Ich und sein Gehirn. München/Zürich 1982.
- ROST, J.: Lehrbuch Testtheorie Testkonstruktion. Bern etc. 1995.
- ROST, J., SPADA, H.: probabilistische Testtheorie. In: K. J. KLAUER (ed.): Handbuch der Pädagogischen Diagnostik. Volume 1. Düsseldorf 1978, pp. 59 - 97.

SCHEUCH, E. K., ZEHNPFENNIG, H.: Skalierungsverfahren in der Sozialforschung. In: R. KÖNIG (ed.): Handbuch der empirischen Sozialforschung. Taschenbuchausgabe. Volume 3a. Stuttgart ³1974, pp. 97 -203.

WATSON, J. B.: Behaviorism. Sixth Impression. Chicago/London 1966.

Boleslaw Niemierko

EDUCATIONAL DIAGNOSTICS IN POLAND

Diagnostic Studies as a Kind of Educational Applied Research²

1 Introduction

Let me begin with an attempt to give you my hopefully exact explanation of what I mean by the term *applied research*. Borrowing partly from the worldwide known *Scientific Method: Optimizing Applied Research Decisions* by Russell Ackoff (1962) I arrived at the following comprehensive statement of the important characteristics of the two main types of research:

Table 1. A Comparison Between Basic and Applied Research on Education

Trait	Basic Research	Applied Research
Main purpose	educational theory building	educational problem solving
Main subject	educational theory	people, curricula, facilities, methods
Expected result	hypotheses confirmed	educational progress made
Strategy	finding reasons	inspecting outcomes
Method	best method	various methods
Initiator	research institute, the researcher himself	educational administration, advisory & methodological centers
Research themes	acc. to the researcher's interest	acc. to articulated social needs
Type of researcher	narrow specialist	many-sided expert
Researcher's aim	increasing in scientific theory	relieving actual needs in education
Setting	strictly controlled	natural, authentic
Time	acc. to the research progress	strictly determined
Budget	estimated by stages	carefully determined

Thus, on the strength of the distinction made in Table 1, when we call educational diagnostics a kind of *applied research*, we shall assume it to be problem-oriented, multimethod, maximally

² A paper prepared for the *Autumn Workshop 2000* in the Institute of Education, University of Kiel, Germany.

authentic and making achievement progress. Such research may be interpreted in terms of *systems approach* to educational programs, *i.e.* according to the following simplified model (based on Astin & Panos, 1971):

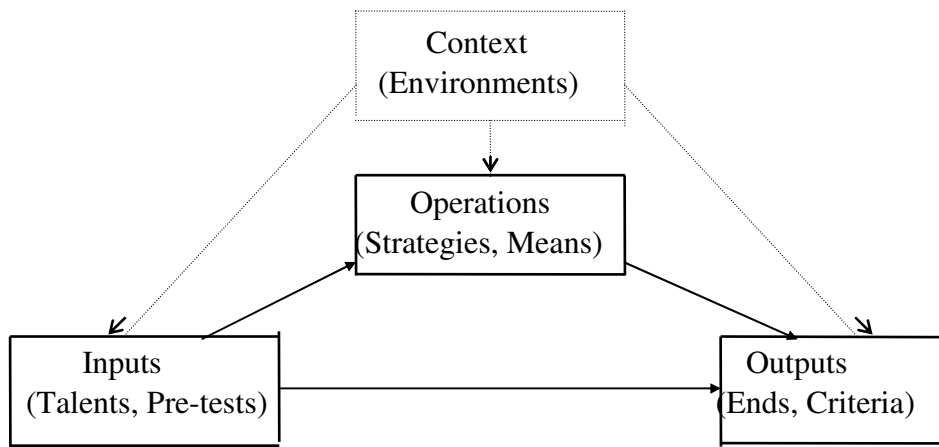


Figure 1. Components of an Educational Program

The context variables, marked in the Figure 1 with the broken lines as external components of educational system, draw more and more attention of educational theorists and practitioners every decade. In my opinion, *educational diagnostics*, as in-depth recognition of educational processes, involves all the four groups of variables shown in the figure with special consideration given to the context variables. Consequently, the principal purpose of diagnostic studies is to provide information concerning various environmental factors which could affect the students' learning motivation and their cognitive achievement. Furthermore, educational diagnosis supplements curriculum studies and underlies evaluation and implementation of educational phenomena. The sequence is presented in Figure 2:

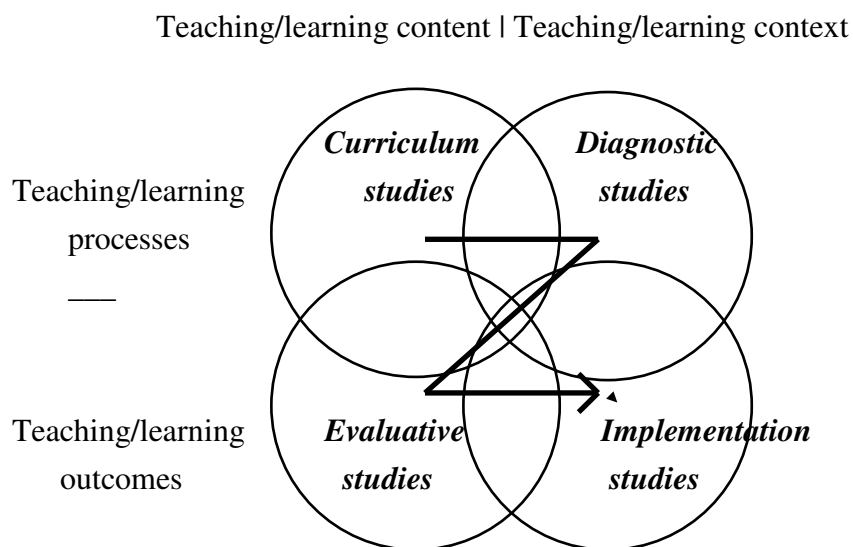


Figure 2. A Simplified Typology of Educational Applied Research

The positive position on educational diagnostics is by no means common to the Polish educational theorists. Prof. Krzysztof Konarzewski (1999), a leading educational psychologist, says that "to diagnose means to draw conclusions on the non-observable variables from the profiles of observable variables", *i.e.* to examine the variables *inside* but not *outside* a student, *e.g.* the hidden sources of disturbance in his/her learning processes. Since „in the human world the cost of false diagnoses could be enormous” the whole task should be entrusted to specialist, that is to psychologists. Prof. Konarzewski warns us of the danger of creating a quasi-scientific discipline called "educational diagnostics". "The harm would be serious" - he writes - "because allegations grow up from interpersonal conflicts, they are notoriously uncertain, and set going mechanism of self-fulfilling prophecy".

Aleksander Nalaskowski (1999), one of the youngest and most brilliant Polish professors of education, reminds that such terms as "standard" and "diagnosis" are unavoidably ideology-laden. Their meaning is specific in the „romantic”, „cultural transmission”, and „progressive” educational ideologies. There is a "countless variety of interpretations of the right a school has to decide on people who had neither choice nor awareness of entering an educational system which had been established by others". And standards give reason to diagnoses, since "they are grown together as Flip and Flap or Marx and Engels".

2 A Short History of Post-War Developments in Educational Diagnostics in Poland

As early as in 1950, a national survey of students' achievement in primary and secondary schools was conducted by Prof. Wincenty Okoń (1951), my university teacher and career-long patron. A few independent variables, such as school size and location, teacher qualifications, and student gender were also recorded in the survey. The finding, repeated in a number of similar studies accomplished later by local advisory and methodological centers, was that *students' achievement depends mostly or even solely upon their teachers' skill and effort*. This statement remained in perfect agreement with the socialist ideology of maximizing everybody's involvement in authoritatively designed social change.

When I started my interest in educational diagnostics, I obviously read some books on this subject, including Ingenkamp's *Paedagogische Diagnostik* (1975) and, for the ideological balance, Helmut Weck's *Leistungermittlung und Leistungsbewertung im Unterricht* (1976). However, momentum to the subdiscipline was rather gained not from readings but from the Polish participation in the IEA (*International Association for the Evaluation of Educational Achievement*) *Six Subject Study, 1968 - 1972*. Probability sampling, test construction, item analysis, and regression analysis led the group of my coworkers into a different world which was less reminiscent of a high-flown philosophy and more like social engineering. Soon did it appear that the IEA

research projects constituted a methodological revelation to many "technical officers" coming from the "second" and the "third" world to the IEA headquarters in Hamburg and Stockholm (Phillipps, 2000). As far as independent variables are concerned, IEA was fairly successful in gathering questionnaire data on national educational system structure, centralization and flexibility; school size, personnel, environment and administration; teacher education, experience, interest, style, and attitude; student background, work, verbal ability, aspiration, and motivation; and many other useful pieces of information.

The Ministry of Education was not quite pleased with the IEA studies and Poland's participation in them is not being continued. "Nationwide research on educational achievement in Poland has been carried out a couple of times, usually within international surveys, with active, yet inconsistent, resistance on the part of authorities. This accounts for the tiny amount of research, the lack of recourses needed, and limited interest on the part of pedagogical theorists, which in turn accounts for its little resonance" - wrote Professor Krzysztof Kruszewski (1998), my good friend, who in the eighties as Minister of Education withdrew the ban on sending IEA data for between-country analyses abroad.

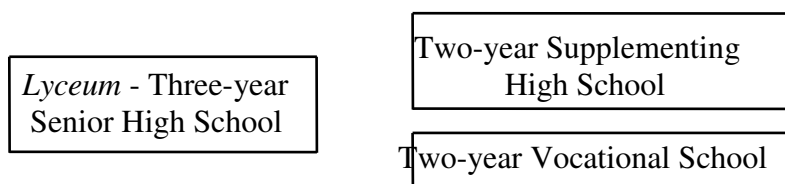
Another distinguished theorist, as well as a friend of mine, Professor Heliodor Muszyński (1998) acknowledged that "an educational policy-maker has got many reasons for regarding the outcomes of diagnosis as dangerous, that is undesirable, or for being interested in the outcomes solely for himself, for treating them with distrust and distance, and, finally, for handling them in a simplified manner that only claims to take them into account in the decisions taken."

In the nineties, when working with a large group of post-IEA measurers, I initiated a series of conferences on educational diagnostic, kindly assisted by Professor Peter Kroepe with his papers on *Dogmatic and Authoritarian Education Styles in Germany* (1994), *Commented Educational Reports: The Source of Misunderstanding* (1994), and *How to Determine the Levels of Measurement?* (1998). Besides, Prof. Kroepe delivered several presentations on examination theory and practice in Post-graduated Studies on Educational Evaluation in Gdansk last years. Four conferences - in Gdansk, Elblag, Legnica, and Szczecin - were successfully carried out in 1993 - 2000 and two others are now scheduled for 2000 - 2001 in Walbrzych and Krakow. The number of participants raised from about 50 people in Gdansk to about 300 in Walbrzych, constituting educational diagnostics a prolific research area.

3 A New Examination System in Poland

In the last few years the main challenge to be taken by Polish educational measurers is designing and implementing diagnostic examination system embedded in the forthcoming school reform in Poland. In the middle of 1999 Poland introduced a new structure of elementary and secondary education (*Ministry...* 1998). It contains three school levels and three adjacent external examinations.

Examination 3. *Matura* - The High School Finals



Examination 2. *Pre-orientation Test*

Gymnasium - Three-year
Junior High School

Examination 1. *Competency Test*

Six-year
Elementary School

Figure 3. The Polish School System after 1999 Education Reform

The examinations are supposed to assure quality control of extensively decentralized school systems and all of them therefore will be state-organized. A network of the Central Examination Committee with eight Regional Examination Committees has been already established. Now the committee leaders and their "edumetric" advisers have to select theoretical models and working procedures for the examinations.

Although the examinations conclude respective education levels and will be held in schools the students graduate from, they may also be considered as the entrance examinations to higher level schools or merely a non-committal pieces of information on a student's strengths and weaknesses. Each of the alternatives has its advocates and some evidence in support of its educational merits.

Educational measurers argue that any national examinations should be „seamlessly” linked to the national core curricula, otherwise a serious damage to the teaching-learning processes will be unavoidable (Nitko, 1998). It means, however, that insufficient achievement should also be recognized and stated. Higher level schools want to select the best candidates either by national examinations or, better, by self-organized procedures. Students, their parents, some Ministry of National Education officials, and some Examination Committee members - they all prefer students' smooth learning progress assisted by non-directive ability diagnosis or, at most, placement suggestions. That is why two years ago the tests were officially called *competency* and *pre-orientation*, and why letter-grading of the test scores was not intended.

Poland, as well as the other Central European countries, has limited experience in standardized *norm-referenced* testing. All the more sensational was the appearance of a well organized objective testing of every elementary school graduate in Walbrzych region (Sokolowska & Sroka, 1997) and then, in about one third of the country. The competencies measured in

”Walbrzych experiment” by two sets of multiple choice items (plus a short essay) are listed in Table 2.

Table 2. Competencies Measured on the Experimental High School Entrance Examination

Subject	Competency/Skill
Polish	1. Reading comprehension 2. Understanding literature and terminology 3. Use of language knowledge 4. Study skills 5. Written composition (essay item)
Mathematics	1. Number use 2. Use of mathematics language 3. Recognition of shapes and relations 4. Information processing 5. Application of mathematics

School principals and superintendents are mostly in favor of objective testing and optical scanning but subject-matter teachers are rather reserved about it. Some weaker students find the possibility of guessing and cheating on the tests attractive (Niemierko, 1998a).

The *New Matura* Experimental Program (1995-1998) brought in quite opposite experience. Based on well-defined achievement standards, as well as essay items, and grade-leveled scoring schemes, it was easier accepted by teachers, although the coding, scoring, and score proving procedures were very laborious (Niemierko 1998b).

4 Perspectives of Educational Diagnostics

One of the consequences of authoritarian educational system in Poland is the students' reluctance to be formally assessed, especially with the letter-grade scale. Some Polish educators believe that using raw score scales in educational examinations is more humane because ”any non-zero score is positive”. Now we are trying to encourage a broader use of stanine scale and profile charts in competency testing in Poland.

Some educators, including the author of the paper, hold the opinion that the (six-point) letter-grade scale provided with achievement standards and properly instrumented (Niemierko, 1990; Niemierko, 1999) will appear to be the most valid medium of communication for students, teachers, examiners, and measurement theorists in Poland. This conviction is supported by the experience of some West European countries, especially Great Britain (Good & Cresswell 1988).

But what about the measurement of context variables regarded as crucial for educational diagnostics in the previous sections of my paper? Now I believe that it should be anchored to the large-scale surveys and external examinations and it should contain all the components of

instructional situations in which students' behaviors tested have been learned. A short specification of independent variables to be included into diagnostic studies is presented in Table 3.

Table 3. An Outline of Learning Context Diagnosis in Large-Scale Surveys

Situation Component	Selected Independent Variable	Exemplary Indicator
Student	Family socioeconomic status <i>Student's learning effort</i>	Parents formal education Student's homework quality (portfolio)
Teacher	Teacher qualifications <i>Teacher involvement</i>	Graduate and postgraduate diplomas Hours spent in school (weekly)
Learning Content	Curricula <i>Teaching strategies</i>	Exceeding core curriculum (percentage) Classroom activities
Facilities	Possession <i>Effective use</i>	Students/computer ratio Facilities used during a week
Management	School administration <i>Management style</i>	Classroom size Decisions left to teaching staff

Italics in Table 3 mark the selected variables harder to objective assessment but all the exemplary indicators seem likely to adapt both to a small-scale and to a large-scale school achievement survey.

There is not much doubt that independent research centers like the Institute for Competence Assessment in Walbrzych are in a better position to perform diagnostic studies than the governmental examination commissions. Examination system in Poland, even if maximally progressive and promising, appears too centralized, too rigid, and too formal to take the risk of assisting schools, teachers, and students in recognizing their educational environment.

5 Conclusions

In my paper, educational diagnostics was defined as a kind of applied educational research aimed at accurate recognition of teaching/learning context. Such studies should be attached to students' cognitive achievement surveys and examinations, *i.e.* to evaluative studies which are supposed to gain a valid measure of dependent variable. An open competition between the new-established regional examination commissions and some private educational assessment centers would create optimal conditions for successful development of educational diagnostics in Poland.

6 References

Ackoff R.L. *Scientific Method: Optimizing Applied Research Decisions*, New York 1962,
John Wiley & Sons, Inc.

- Astin W.A., R.J. Panos *The Evaluation of Educational Programs*
(In:) R.L. Thorndike (Ed.) *Educational Measurement. Second Edition*,
Washington 1971, American Council on Education.
- Good M., F. Cresswell *Grading the CGSE*. London 1988, Secondary Examinations Council.
- Ingenkamp K. *Paedagogische Diagnostik*, Weinheim 1975, Beltz.
- Konarzewski K. *Komu jest potrzebna diagnostyka edukacyjna?* [Who Needs Educational Diagnostics?] (In:) B. Niemierko, B. Machowska (Eds.) *Diagnoza edukacyjna. Oczekiwania, problemy, przykłady*, Legnica 1999, ODESiP (in Polish).
- Krope P. *Dogmatyczny i autorytarny styl edukacyjny w Niemczech* [Dogmatic and Authoritarian Education Styles in Germany] (In: B. Niemierko (Ed.) *Diagnostyka edukacyjna*, Gdańsk 1994, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).
- Krope P. *Ocena opisowa w pedagogice jako Źródło nieporozumień* [Commented Educational Reports: The Source of Misunderstanding] (In: B. Niemierko (Ed.) *Diagnostyka edukacyjna*, Gdańsk 1994, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).
- Krope P. *Jak określać szczeble pomiaru?* [How to Determine the Levels of Measurement?] (In:) B. Niemierko, E. Kowalik (Eds.) *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).
- Kruszewski K. *Test wiadomości dobrego i złego* [Test of the Knowledge of Good and Evil] (In:) B. Niemierko, E. Kowalik (Eds.) *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).
- Ministerstwo Edukacji Narodowej *Reforma systemu edukacji w Polsce. Projekt* [Education System Reform: A Proposal], Warszawa 1998: WSiP (in Polish).
- Muszyński H. *Diagnoza edukacyjna i jej miejsce w praktyce i teorii pedagogicznej* [Educational Diagnostics and its Place in Educational Theory and Practice] (In:) B. Niemierko, E. Kowalik (Eds.) *Perspektywy diagnostyki edukacyjnej*, Gdańsk 1998, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).
- Niemierko B. *Pomiar sprawdzający w dydaktyce* [Criterion-Referenced Measurement in Education: Theory and Application], Warszawa 1990, PWN (in Polish).
- Niemierko B. Educational Testing in Poland. *Educational Measurement: Issues and Practice*, 1991, 10(3), 27-32.
- Niemierko, B. *Ewaluacja sprawdzianu kompetencji uczniów szkół podstawowych* [Evaluation of the Competency Testing in Elementary Schools]
Unpublished Report for the Ministry of National Education, 1998a (in Polish).
- Niemierko, B. *Probna matura 1998* [The Mock Matura 1998].
Unpublished Report for the Ministry of National Education, 1998b (in Polish).
- Niemierko B. *Pomiar wyników kształcenia* [The Measurement of Educational Achievement], Warszawa 1999, PWN (in Polish).
- Nitko A. J. *Model egzaminów państwowych opartych na programie nauczania, sprawdzających i różnicujących, przeznaczonych do dyplomowania i selekcji*

uczniów [A Model for Curriculum-Driven Criterion-Referenced National Examinations for Certification and Selection of Students. (In:) B. Niemierko, E. Kowalik (Eds.) Perspektywy diagnostyki edukacyjnej, Gdańsk 1998, Wydawnictwo Uniwersytetu Gdańskiego (in Polish).

Phillipps R.W. *30 Years On. Alumni Register of the National Technical Officers from the IEA Six Subject Study*, The Hague 2000, IEA Headquarters.

Sokolowska J., D. Sroka *Raport o eksperymencie walbrzyskim [Report on the Walbrzych Experiment]*. Wa³brzych 1997, Kuratorium Oświaty w Walbrzychu (in Polish).

Weck H. *Leistungermittlung und Leistungsbewertung im Unterricht*, Berlin 1976, Volk und Wissen.

THE AUTHORS

Peter Krobe

Dr. phil. habil., Professor of Education, University of Kiel

Boleslaw Niemierko

Dr. phil. habil., Professor of Education, University of Gdansk

Folgende Monographien zur Prüfungsforschung sind bisher erschienen:

- | | |
|--------|---|
| Nr. 1 | Peter Krobe, Hans-Werner Zumach
Leistungsmessung als Deformation der Wirklichkeit |
| Nr. 2 | Peter Krobe, Lothar Viehöfer
Prüfungsangst und Aufgabenkomplexität |
| Nr. 3 | Arne Kohrs
Bibliographie zur Angst unter besonderer Berücksichtigung der Prüfungsangst |
| Nr. 4 | Lothar Viehöfer
Entwicklung eines Instruments zur Messung motorischer Reaktionen |
| Nr. 5 | Peter Krobe
Die Kieler Affekt-Adjektiv-Liste (KAAL) zur Messung situativer Prüfungsangst |
| Nr. 6 | Jochen Hestermann
Kooperative Gruppenprüfungen als Alternative zu Einzelprüfungen an Hochschulen |
| Nr. 7 | Olaf Albers
Das Problem der Ausgangswerte bei der Physiologischen Messung von Prüfungsangst |
| Nr. 8 | Peter Krobe, Arne Kohrs
Prüfungsangst und kooperative Gruppenprüfung |
| Nr. 9 | Cornelia Klink
Mündliche Hochschulprüfungen und Labeling Approach |
| Nr. 10 | Peter Krobe, Arne Kohrs
Zur medikamentösen Behandlung von Prüfungsangst bei Studenten - eine Methodenstudie |
| Nr. 11 | Burkhard Seidler
Inhaltsanalyse und Leistungsmessung - eine Methodenstudie |
| Nr. 12 | Peter Krobe, Arne Kohrs, Karin Recklies, Maria Peters, Jutta Schneider
Über kriterienorientierte, nicht-quantitative Leistungsmessung - Ein Untersuchungsbericht |
| Nr. 13 | Olaf Albers, Beate Conzen, Axel Klimek, Cornelia Klink, Wolfgang König, Peter Krobe, Annette Krüger, Karin Lüdders (Hg.)
Das Mediziner Auswahlgespräch im Dialog |

- Nr. 14 Peter Kroepe (Hg.)
 Prüfungs-forschung in der Bundesrepublik - Zwischen-
 bilanz und Perspektiven
- Nr. 15 Peter Kroepe, Boleslaw Niemierko (Ed.)
 Proceedings of the 1992 Workshop on School Examinations
- Nr. 16 Peter Kroepe, Berichtszeugnisse: Die Quelle der Mißverständnisse

Die Monographien zur Prüfungs-forschung sind, soweit nicht vergriffen, zu beziehen direkt über die Verfasser oder über das Projekt Prüfungs-forschung am Institut für Pädagogik der Universität, 24118 Kiel, Olshausenstraße 75.

Die Monographien zur Prüfungs-forschung werden ab Heft 16, 3. Auflage, unter der folgenden Adresse als Word-Dateien elektronisch publiziert und kostenlos zur Verfügung gestellt:

<http://www.uni-kiel.de/Paedagogik/Kroepe>

Die Monographien werden inhaltsgleich in geringer Auflage in traditioneller Weise gedruckt und vorrätig gehalten. Auf diese Weise soll in Zweifelsfällen die Autorenschaft gesichert und der Originalzustand der Textdateien überprüft werden können.