

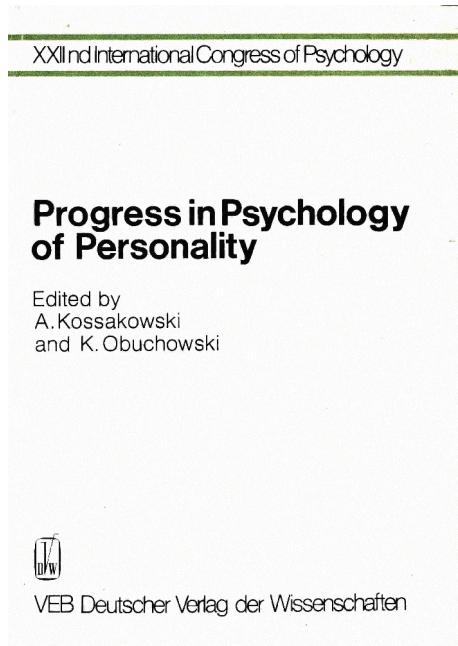
Georg Lind

Experimental Questionnaires:
A new Approach to Personality Research

1980

erschienen in

Adolf Kossakowski/Kazimierz Obuchowski (Ed.)



Deutscher Verlag der Wissenschaften 1982

Kontakt:

Prof. Georg Lind
Universität Konstanz
FB Psychologie
78457 Konstanz
E-Mail: Georg.Lind@uni-konstanz.de

Weitere Informationen und Publikationen des Autors zu diesem Thema finden sie im Internet unter www.uni-konstanz.de/ag-moral/b-publik.htm

The original was published with the following title: Progress in Psychology of Personality, Adolf Kossakowski/Kazimierz Obuchowski (Ed.), Deutscher Verlag, Berlin 1982.

EXPERIMENTAL QUESTIONNAIRES: A NEW APPROACH TO PERSONALITY RESEARCH¹

Georg Lind

The concept of Experimental Questionnaires is an attempt to facilitate new, adequate means for assessing dynamic-structural traits. In this paper, I will give, as briefly as possible, an account of the rationale, the design and the utility of this concept.

1. Introduction

Traits and dispositions of individual persons are the genuine objects of personality research. Contrary to Differential Psychology which embarks upon the assessment of any difference between persons, for personality research the assumption of behavior organizing traits is constitutive (Mischel 1977, 248).

But, in spite of the simplicity of this insight, it remains a rather difficult task to translate the notion of traits into a unanimously accepted scientific concept. Hitherto, the empirical evidence pertaining to this issue has been ambiguous and has served both, those who believe in the reality of traits and those who deny it. While in some fields of personality research, people's behavior is strikingly consistent, in others consistency is disappointingly low or absent in others.

How can this persisting crisis be overcome? How can we cope with the stagnation in personality research? For a considerable time we believed, maybe with a few exceptions, that the crisis was due to a lack of either theoretical or methodological sophistication. We have witnessed a flood of new theoretical speculations and of new research strategies that are triggered by the failure of personality psycholo-

gy to account for its perplexing findings. Doubtless, these endeavors have enriched personality research, yet we must admit that the “slow progress of soft psychology” (Meehl 1978) is still among the few facts on which psychologists would most easily agree.

If the stagnation of personality research cannot be overcome by the improvement of either theory or method, the problem may be more fundamental. In fact it may be located in the relation between theory and method. As early as 1951, Travers pointed to the importance of theory-guided construction of personality tests for the enhancement of psychological knowledge. Recently, Fiske (1970) mentioned again the “problem of coordinating the measure and the concept” (51). The programs of “construct validation” (Cronbach & Meehl 1955) and of “substantial validation” (Loevinger 1957) have also this problem in mind, but unfortunately reduce it to the problem of optimizing the correlation between different measures.

The cleavage between theory and method in personality research has led to two major 'solutions'. Some eminent researchers, like Freud and Piaget, have turned completely away from psychometrics and formed methods of their own, the so-called clinical methods. The clinical interview method, if theory-related, deserves consideration. Yet its lack of transparency and testability compels us to search for more 'objective' methods. (The exceptional method by G. A. Kelly is somewhere in the middle and would need to be treated in a larger paper.)

Most others apparently have adopted a dualistic view that conceives of theory and method as two somehow related yet not strictly connected realms of research. Some believe that it is possible to escape the reproach of positivism and still take advantage of “objective” methods. But this belief is not tenable either because, as Allport (1951), who once himself advocated this point of view, has noted, this would imply that “we must rest content with mere approximations to the structure of personality” (p. 422). As Block (1968) sharply stated, “if a science of personality is to be formed, the responsibility of coupling concept and measure must be met” (p. 30). For that we cannot borrow methods from elsewhere: The task itself must create the method that is appropriate for it (Dilthey).

Yet, the problem of coordinating theory and method is difficult to achieve par-

ticularly in the field of personality research. Koffka (1934) already noted, “no-where is it easier to miss the point, and run either into the Scylla of blind statistical investigations of trait, or the Charybdis of ultimately unscientific abstract discussions” (p. 677). Moreover, the solution of this problem cannot be confined to the construction of a 'valid' and 'reliable' assessment method but must also comprise a rethinking of the methodological criteria for the test construction. We must be aware of the fact that concepts like 'validity' and 'reliability' themselves may be inadequate.

Validity: Given that research methods are not ends in themselves but are always means for executing and testing a theoretical proposition, the validity problem is twofold:

- (a) How empirically valid is the theoretical proposition?
- (b) And how theoretically valid (adequate) is the assessment method?

In 1957 Loevinger pointed out the necessity of making this distinction, but unfortunately did not recognize the profundity of its implications for her cause and for psychometric methods in general. If one accepts it, speaking of 'the empirical validity of tests' becomes meaningless. A psychometric test cannot, at least not directly, empirically validated. One may infer that, if a test has been successfully employed for validating a theory in empirical research, it must have possessed theoretical validity. Note, however, that this conclusion is asymmetric: A failure to produce supporting findings may be either due to a lack of empirical validity of the theory or to a lack of theoretical validity of the method or to both. Hence theories are falsified by empirical findings if and only if the method was theoretically valid (Lakatos' objection to crude falsificationism). So, in order to prevent confusion, I rather prefer to speak of the test's utility than of its (empirical) validity, when the test helped to produce 'good' results.

Reliability: Lumsden (1974) has made a strong case for regarding the reliability problem also as a problem of a test's validity. In fact the necessity of doing so is obvious since (a) classical test theory has advanced a substantial theory of inconsistencies of test behavior, namely error theory, and (b) this theory is testable, for it

entails the assumption that this error is solely due to the test (and not to the property of the individuals tested). If this objection is disregarded, psychology runs the risk of carrying along hidden “anthropological assumptions” (Hilke 1980) which are at odds with the very theory involved. So “the study of measurement must begin within discussion and analysis of such matters as the nature of traits and the mode of representation of the total personality” (Cattell 1946, 4).

Notes

1. Paper presented at the XXII International Congress of Psychology in Leipzig, GDR, 1980. [Notes for the edition in 2000: More recent information on the MJT can be found in G. Lind (2000). Review and appraisal of the Moral Judgment Test. Invited presentation, Annual conference of *the Association for Moral Education*, in Glasgow, July, 2000.] A revised version of this paper is published as G. Lind (1982). Experimental Questionnaires: A New Approach to Personality Research. In: A. Kossakowski & K. Obuchowski, Hg., *Progress in Psychology of Personality*, pp. 132-144. Amsterdam, NL: North-Holland.

2. Problems of translating personality theory into an adequate research methodology (Theory of trait assessment)

Experimental Questionnaires (EQs) are designed to fit the intentions of dynamic-structural theories of personality (e.g., Warren & Carmichael 1930; Lewin 1961; Allport 1961; Kohlberg 1969; but also many of the theoretical propositions of Stern 1911; Hartshorne & May 1928; Cattell 1946). According to these theories, *personality is the dynamic organization of the individual's mind that determines his or her interaction with the social environment at a particular level of development.*

This is not the place to discuss in depth the problems that result from translating this definition of personality into an adequate research methodology. But a condensed account of ten recurring problems of personality research may help to understand the rationale of the concept of Experimental Questionnaires and some of its implications. As far as I can see, these problems are, by and large, those

which other researchers in this field also encounter frequently: (1) the problem of conceptualizing traits, (2) the problem of inferring traits from behavioral acts, (3) the problem of functional ambiguity of acts, (4) the problem of separating the dynamic and structural components of a trait, (5) the problem of conceptualizing person situation interaction, (6) the problem of assessing meaningful behavior, (7) the problem of determining the experiential unit, (8) the problem of defining behavioral consistency, (9) the problem of locating the source of inconsistency, and (10) the problem of representativeness. For the purpose of this paper these problems need not, and cannot, be extensively treated. It will suffice to make the set of tentative solutions explicit which constitute the rationale of EQs.

(1) Graumann (1960) juxtaposed two scientific concepts of traits: One is the “statistical,” which “is nothing else but a labeling of any frequent coincidence of particular, consistent, (behavioral) signs for the sake of better communication.” The other is the “hypothetical construct” concept of traits, which comprises only those correlations among signs that “can be interpreted through recourse to a hypothesized common instance, i.e., through reference to a substruction, as meaningful associations” (p. 147; my translation). Resorting to a dynamic (teleological, organismic) definition of personality brings us, of course, closer to the latter, internal structure concept (Allport), rather than to the former, positivistic concept of a trait. Since the latter definition seems to contain a moment of unwarranted speculation, we prefer not to distinguish between “real” behavior and “hypothetical” traits. Surely, traits are different and need to be assessed in a different way. But they are not real, or more hypothetical, than single acts (cf. the critique of Cronbach and Meehl’s understanding of “hypothetical constructs” by Loevinger, 1967, p. 47; also Hempel’s, 1958, discussion of “theoretical entities”). According to the internal structure concept, traits do not exist ‘behind’ behavioral acts, nor do they ‘cause’ acts; rather they are regarded as the combined whole of (a) dynamic goals (values, orientations), (b) instrumental acts, and (a) the relational structure between acts and goals (cf. Lind, Sandberger & Bargel, 1980). The formal ground-structure of dynamic-structural traits is similar to, but more complex than, Miller et al’s (1960) T-O-T-E model in so far as we regard the multiple determination of acts as indispensable for theory and research of personality (e.g., for conceptualizing ‘con-

flict'). But it is still so elementary that the difference to other trait concepts can easily be seen (see the following graph):

(2) This structure model of personality has a direct bearing upon the problem of inferring traits from behavioral acts. First, the relation between acts and dispositions should be conceived of neither as unknown, nor as stochastic nor as mediated by an intervening variable but as a logical inclusion: Acts are constitutive parts of a trait. Second, the relation of acts to the two components of a trait, the dynamic and the structural, is well defined and can be reconstructed empirically if that particular trait really exists.

(3) To view acts as determined by more than one disposition, implies that in interpreting a person's solitary act we have to cope with its functional ambiguity (Kempf 1978, p. 12; Nunner-Winkler 1978, p. 352). Since there is no one-to-one relationship, the problem of making inferences about a trait from a single observation is a substantial one. For example (cf. Lind 1978a), a person may agree to an argument of mine in defense of someone's decision . . .

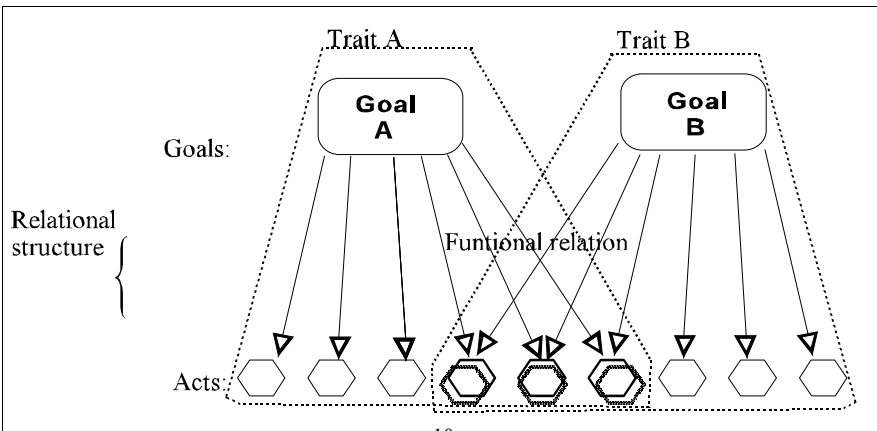
- because it fits nicely into his opinion about the matter which he, or she, already held before (opinion agreement)
- because the person feels, the argument is referring to a good reason (quality or level of reasoning)
- or because it is I who has forwarded the argument (acquiescence).

The index in parentheses alluded to the diversity of traits that may be involved, either alone or in combination, in producing a single act and that is of so great annoyance to the psychometrician who attempts to design a one-dimensional test. This problem seems to be ubiquitous, as it is encountered in the research into moral attitudes (Nunner-Winkler 1978) as well as in intelligence testing (Kempf 1978; Loevinger 1957, p. 647).¹

(4) A further problem pertains to the conceptual relation between the dynamic (valuing, motivational) component and the structural (cognitive) component of a trait, i.e., between the "Richtungs-" and "Rüstungs-" component (W. Stern) of a disposition. According to the concept adopted both components are clearly distin-

guishable but not separable. We, therefore, do not speak of different ‘dispositions’ or ‘faculties of mind’ but of different components of one and the same trait. The theorem of *inseparability* also implies that neither component can be defined without reference to the other. While purely formal relations are empty, unrelated acts and values remain meaningless. This problem is, in my opinion, of great relevance for the prevailing discussions about (a) the ‘content-specificity’ of cognitive structure (complexity, style, etc.) and (b) the ‘cognitive turn’ of motivation and attitude research. In both instances, the superficial separation of the two components of traits apparently has resulted in severe difficulties with regard to conceptualization and measurement of personality variables.

(5) Another perplexing conceptual problem arises from the notion of person-situation interaction. In fact it seems as if we are even moving away from some promising attempts of early personality research. When we refer to the well-known formula of Lewin $V = f(P \& U)$ (read: behavior is a function of person and environment interaction), we are rarely aware of the fact that this formula does *not* imply that, as ‘modern’ interactionism surmises, person and environment can be empirically separated. Lewin (1961) did not mean that we should conceive of separate entities but that we should focus on their interaction, i.e., ‘life-space.’ Accordingly, personality is the way a person perceives, evaluates, and reacts to a situation (Helm 1960). Hence “in concrete experiments both cannot be separated” (Olweus 1976, 65). Maybe, that our language is responsible for some of the prevailing confusion



as it does not distinguish between a “separating and,” and a “combining and.””²

(6) Personality research is concerned not with the mere physical properties of human behavior but with the meaning of behavior. Verbal behavior is meaningless unless we ascertain its functional and structural properties with regard to the ‘totality’ of a trait.

(7) Nothing seems to be more obvious than that the unit of personality research is the individual. In theory, we almost exclusively refer to the individual personality. There seems to be no meaningful way of setting the ‘personality of a sample of persons’ equal to the individual personality structure. In practice, however, it is often implicitly or explicitly assumed that traits are commonly organized. Since such a notion does not take into account the individuality of traits, Allport (1961) has called this the External-Effect definition of traits. Persons are judged merely according to some external, socially defined standards of conduct and achievement. This is an approach in its own right but it does not accord with our concept of personality.

(8) In many textbooks one can find the statement that personality development means both integration and differentiation. In empirical research this often is translated as: Behaviors become more and also less consistent. An obvious contradiction! In fact research has produced supporting evidence for both propositions. A closer look shows that at the heart of this problem there is an insufficient definition of ‘consistency’. If ‘consistency’ means “holding to the same principle or practice” (Webster’s New World Dictionary), a proposition which does not specify the ‘principle’ or ‘practice’ is incomplete. In this context ‘consistency’ is a relational concept. Behaviors cannot be consistent per se, but only ‘consistent with regard to something’; in our case ‘with regard to the function it is to serve’. So behavior may become more consistent with regard to function A and, at the same time, become less consistent with regard to function B. Besides such simple processes of integration and disintegration there may also be a superposition of traits, whereby the superposed trait does not become disintegrated but ‘differentiated’ with regard to another ‘qualifying’ trait.

The problem of defining consistency also has a bearing upon the distinction between the External-Effect and the Internal-Structure model of traits (Allport).

While the External-Effect model implies that behavior must be, if at all, consistent with regard to some external standard, the Internal-Structure model implies the existence of internal standards to which behavior may, or may not, accord.

(9) One way to save the External Effect (a common trait) model in the face of an overwhelming number of discrepant facts is to assume that the measurement process introduces random error, i.e., that inconsistency of behavior with regard to a hypothesized common trait is due to a property of the assessment method. This assumption, advanced by classical psychometrics (Spearman, Gulliksen, Guilford etc.), raises the problem of attributing inconsistency. While 'error theory' attributes inconsistency to the 'reliability' of the method, the theory of dynamic-structural personality conceives of it as an attribute of trait organization. This is an empirically testable question! In fact, research into personality and personality development has produced ample evidence that inconsistency of behavior varies systematically with the kind of trait measured and with the phase or stage of personality development (see, for example, Hartshorne May 1928, Kohlberg 1969, Loevinger 1976, Lind 1978 a). In a sum, I think, we should not rule out the assumption of random error (unreliability). This precisely defined concept (Kempf 1978), which is to be distinguished clearly from the trivial explanation of 'unknown factors', may, however, come in as the last, and not as the first, explanation for observed inconsistencies.

(10) Finally we need to provide a tentative solution for the problem of representativity. Brunswik (1955), as many psychologists after him, has invoked the concept of ecological validity which means that the experimental stimuli should be distributively representative. But what population of stimuli should we refer to when designing an experiment? There are uncountable possibilities; some of them are not observable in principle (e.g., the distribution of events over time). So, which should we choose? This problem is present in theory-guided research. Generalizations can never be based merely on a single experiment but are always based on the comprehensive experience comprised in a scientific theory. In order to test the empirical validity of a hypothesis, we need to ascertain categorical representativity. Only this guarantees that the research instrument tests what it is supposed to test. I, therefore, believe that many objections against systematic experimental de-

sign are truly directed at the positivistic spirit in which they are often used. Like Lewin (1963), I think that “experiments become superficial only if just one or the other condition is operationally defined, but not the essential structure” (p. 200). Of course, what is “essential” must be defined by a theory. It is noteworthy that Brunswik (1955) has not only introduced the concept of ‘ecological validity’ but also shown very convincing causes for the epistemological utility of systematic research designs (the “diacritical method”).

For the choice of a theoretically valid research strategy our previous assertions have three basic implications. First we must conceive of the research situation as part of the personality investigated. Adequate research methods need not eschew interaction with the subjects but may rather profitably make use of “those individual and concrete person-environment constellations which conspicuously expose their dynamic characteristics” (Helm 1960, 374).

Second, according to the dynamic-structural trait concept “the task of psychometrics is to isolate, to identify and, so far as possible, to measure separately (though not as separate objects!) the important components of (behavioral) variance” (Loevinger 1957, 649; parentheses added). We would add: The task is to translate a particular personality theory into an adequate research situation which allows the researcher to infer the structural and the dynamic component of the hypothesized traits, i.e., their behavioral relevance (‘cognitive anchoring’) and their directional value (attitude, orientation).

Third, the problem of multiple determination of acts, and the resulting functional ambiguity of acts, should not be viewed as insolvable in principle (‘indeterminism’). It can be coped with through applying the ‘diacritical method’ (Brunswik). I believe this hermeneutic device is well suited to sort out the functional unities to which an act belongs. It requires that one designs a pattern of systematic probing questions to elicit a telling pattern of responses. By such an observational design, it is possible to find out which trait, or set of traits, determines this particular interaction between a person and a situation.

With regard to these three points, I think, the question of whether trait assessment is executed by a skillful clinician through a personal interview or by a survey researcher through carefully prepared questionnaires is of minor importance (Lind

& Wakenhut 1980). The variation of theoretical validity within each approach seems to be greater than between them.

There are good reasons why personality research cannot rest solely on highly elaborate and extensive clinical interviews. The lack of transparency and systematic design makes it all too easy for uninformative circularity and trivial explanations to creep in. Hence it is important to try out a new, complementary, approach to personality research.

3. The Design of Experimental Questionnaires

At the first sight this method appears to be a contradiction in itself. Experimental Questionnaires (EQs) neither fit into usual methodological categories, e.g., Cattell's distinction of L-, Q-, and T-techniques, nor are they a new way of utilizing questionnaire methods within experiments. This method rather combines, maybe in a surprising way, the meaningfulness and economy of the questionnaire technique with the epistemological advantages of systematic, multi-factorial design.

Unlike self-report questionnaires, EQs are conceptualized as "objective tests" in Loevinger's (1957) sense, namely in the sense "of structures tests viewed behavioristically" (648). The behaviorism of EQs, however, is not a Watsonian or a Skinnerian one but more like the "subjective behaviorism" of Miller et al (1960) and the "social behaviorism" of G.H. Mead (1968/1934). EQs are viewed "subjectively" because they are designed to penetrate through the surface of unrelated behavioral acts and to assess the dynamic trait structure of individuals. EQs are, nevertheless, also social in so far as their scientifically construed categories are bound to the sociality of communication. EQs are used nomothetically (i.e., they seek systematic regularities of human behavior), but, at the same time, they possess idiographic sensibility.³ EQs are designed to answer two questions of measurement in this order:

- First, to what degree does a hypothesized trait account for an observed pattern of responses of an individual?

Independent variables (factors)			Dependent variable																			
1. Story (dilemma)	2. Stage of reasoning	3. Pro-Con (opinion- agreement)	Judgment of acceptability																			
- A	-I	Pro Con	completely	completely																		
	-II		unacceptable	acceptable																		
	-III	.	-4	-3	-2	-1	0	+1	+2	+3	+4											
	-IV	.	<table border="1" style="width: 100%; height: 20px;"> <tr> <td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td> </tr> </table>																			
	-V	.																				
-VI	.																					
	.																					
- B	-I	Pro Con																				
	-II																					
	-III	.																				
	-IV	.																				
	-V	.																				
	-VI	.																				

Tab. 1 Experimental questionnaires: The design of the “Moralisches-Urteil-Test” (m-u-t)

- Second, in which direction does this factor influence the test response?

While the first question aims at the assessment of the cognitive-structural component of a particular trait, the second question aims at a trait’s dynamic-motivational component. In addition, EQs may be designed as a multi-factorial experiment so that more than one dynamic trait can be incorporated in order to study a more complex system of traits. This way EQs also facilitate means to cope with the problem of functional ambiguity. Through multi-factorially designed EQs it becomes possible to find out which of several traits functions alone, or in combination with other traits, as a frame of reference for a particular person-situation interaction. In principle, an unlimited number of traits can be analyzed diacritically by EQs, although in practice the rule of parsimony (Ockham’s razor) may be applied which entails starting always with as few hypothesized traits as possible. There are a number of applications possible and some have been realized indeed.

As an example for EQs, the three-factorial design of the Moral Judgment Test (MJT; German: “Moralisches-Urteil-Test”, MUT) is depicted in Table 1. The aim of this test is to assess the cognitive and the value component of an individual’s moral consciousness. The MJT has been developed on the basis of Kohlberg’s theory of moral-cognitive development (Kohlberg 1969; Rest 1978; Lind 1978a, 1980e), and on a theory of response behavior which contains three tentative assumptions derived from recent moral judgment research (Keasey 1974, Lind 1977):

1. An individual’s evaluation of moral arguments for and against a particular moral decision is determined by the quality of the moral argument, i.e., by the individual’s attention to the Stage-type of moral reasoning exhibited by the arguments (according to Kohlberg 1969);
2. that the statements may be also evaluated with regard to their agreement or disagreement with one’s own opinion about the moral dilemma;
3. This inconsistency of judgment behavior may not just indicate a lack of moral development but a greater differentiation; it may indicate (a) the successful coordination of the moral consciousness factor and the opinion agreement factor (‘mature moral commitment’; Perry 1970); and (b) the differentiation of the preferences for a particular level of moral reasoning according to the situation context in which the moral decision is made (‘contextual relativism’).

The design of the MJT is depicted in the following table 1 contains two sub tests (Theft and Mercy Killing). Each is introduced by a brief description of a dilemma of action and by a subsequent question whether the respondent judges the presented solution of the dilemma as “wrong” or “right.” Thereafter six pro-arguments and six con-arguments are presented, each at random order. Thus, the MJT forms a 2 x 6 x 2 idiographic experiment. For scoring the MJT we have surveyed several computational procedures (Lind et al 1976; Lind 1980 g). Hitherto the analysis of variance components proved to be one of the most adequate and fruitful methods. It enables the psychologist to quantify the degree to which a certain hypothesized trait, alone or in conjunction with others, determines an individual’s pattern of response behavior (the structural-cognitive component). The value component is measured as usual by the sum of the ratings (average response).

4. Conclusion

We have argued that the prevailing standards of test evaluation do not apply to assessment methods like Experimental Questionnaires. But this does not mean that there are no other ways to ascertain the goodness of a research method. I can see at least two possibilities:

a) One possibility is concerned with the theoretical validity of the assessment technique. Since we have ‘deduced’ the concept of Experimental Questionnaires from a theoretical proposition this method may legitimately be evaluated with regard to that claim.

b) The other criterion involves both theory and method. The choice of the dynamic-structural definition of traits entails that a test based on this definition is an empirically useful device for personality research in so far as it widens the ‘categorical overlap’ (Campbell 1963) between the researcher and his, or her, subjects. The greater this categorical overlap is, the better, we can say, we understand a particular person-situation interaction.

I believe that both criteria provide good means for criticizing the method of Experimental Questionnaires, and thus foster the progress of personality research. Theoretical validity cannot be measured but must be established on the grounds of conceptual analysis and theoretical reflection. A profound explication of the respective theory as well as some kind of expert-rating of the resulting test are the major devices for securing theoretical validity. This has been done in the case of the MJT. So for the time being we shall assume that it is theoretical valid.⁴ With respect to the categorical overlap of Experimental Questionnaires there is some quantitative evidence available. The MJT which is based on the EQ-methodology has been employed recently in a number of studies comprising several hundred of subjects. One of the most important results is that some propositions of Kohlberg’s (1969) theory of cognitive-moral development were rather well corroborated (Lind 1980 d). A second important result is that it may suffice to take into account only a few dynamic-structural traits in order to understand a person’s judgment behavior in a moral situation. This is of relevance for the utility criterion.

Ideally the hypothesized trait-structure of moral consciousness should account

completely for the variance of the individual response behavior in the experimental situation. This does not necessarily imply that behavior is completely predictable, but that we can comprehend it as meaningful given the knowledge of a particular trait structure. Past personality research has shown that such a complete overlap between theory and reality is far from being reached. Hartshorne and May's (1928) *Study in the Nature of Character* revealed only minimal categorical overlap; the squared correlations between children's behavior across different 'moral' situations, which were used as an indicator, was as low as $r^2 = .05$. This is even below the .10 which Mischel (1968) has identified as a barrier to personality research.

Gordon Allport criticized that study for being a typical representative of the external-effect definition of traits. Children may not behave consistently with regard to some socially defined criteria and still be consistent with regard to their own motives. Bem and Allen (1974) showed, though with regard to another topic, that consistency increases up to .20 when the individual's awareness of trait consistency is taken into account. Since EQs represent a new approach which claims greater adequacy with regard to structural trait concept we should expect a better agreement between the hypothetical model and the reality of an individual's reasoning in a moral situation. In fact, in a study of German high school graduates (N= 516), one trait (moral consciousness) already accounts on average (median) for .27 of the individuals' behavior variance. There is also evidence that this number varies with the level of moral development; the less people are developed the less they apply moral dimensions in their judgment behavior (Keasey 1974).

But this can also be looked at from another angle. Usually a trait is assessed, regardless of whether there is a large, medium or no categorical overlap between the researcher and the particular subject. If the consistency of test-behavior across a sample of persons is not perfect, according to most textbooks, this can be coped with by enlarging the number of items, reliability is zero.

Besides the frequent experience that more items often add more 'heterogeneity' to the assessment device and thus, in practice, often do not improve it, this strategy is basically problematic as noted above.

Approach (- Example)	Categorical Overlap (r^2)*
1. Common Trait, External-Effect definition	
- Hartshorne & May (1928)	0,05
- Mischel (1968)	0,10
2. Restricted Common Trait definition	
- Bem & Allen (1974)	0,2
3. Internal-Structure definition	
- Lind et al (1980)	0,27

* squared correlational coefficient, i.e. proportion of variance of judgment accounted for by moral consciousness and personality traits, respectively.

Tab. 2 ‘*Categorical overlap*’ in three approaches to personality research: a tentative comparison

In many cases we ‘measure’ traits that do not exist in subjects and do not even notice it. Someone who does not ‘have’ the particular trait measured gets the same score as the person who exhibits a medium (or neutral) ‘amount’ of this trait. E.g., a person who is as much introvert as extrovert in his responses to individual items of a scale receives the same, average, score as a person who responds consistently medium extrovert, or medium introvert.

In assessing the degree to which a hypothesized dynamic component organizes instrumental acts into a functional whole, Experimental Questionnaires avoid this pitfall. This method provides a measure for ascertaining the degree of categorical overlap between an individual for classifying this person. The person is thus not forced into a descriptive system which actually does not tally with his, or her, way of perceiving the world. The measure of ‘cognitive anchoring’ of a trait thus gives us also an opportunity of becoming aware of the descriptive adequacy or inadequacy of the theory on which the assessment technique is based. It tells us something about the ability of a psychological theory to understand particular persons and

persons in general. (I am not speaking of the psychologist's ability to understand since this may deviate greatly from the common knowledge as codified in a theory. Certainly, and hopefully, many psychologists also make successful use of sources of knowledge).

Notes

1. There are similar problems in natural sciences. As Piaget (1973) observes, also in physics, chemistry and biology "at the outset the problem is always to isolate the components out of the bewildering mass of phenomena" (36; translation mine). See also Popper (1979).
2. This problem is not unique for psychology. Also in the natural sciences the "observable is only the result of an interaction between experiment and reality" (Piaget 1973, 53). Hence the scientist's major task is to disentangle the nature of the observation (i.e. the observation mediating processes) and the nature of the observed. This is, even in physics, less trivial than might be assumed. For some time, physicists interpreted his indeterminacy principle for a long time as pertaining to the measurement process while it was only eventually realized that it pertained to the nature of a quantum (cf. Popper 1979, 305).
3. Unfortunately 'nomothetic science' is sometimes misunderstood as postulating a priori the lawfulness of any behavior. This is, of course, not the case. Nomothetic sciences are rather law-seeking. They critically test the question to which extent the hypothesis of lawfulness is empirically valid. This implies also that sciences have to be aware of idiosyncracies.
4. Note that EQs always require the existence of an explicit substantial theory (which is more than a statistical account of singular investigations or a mathematical model). Otherwise there is no reason for employing EQs. On the other hand a theoretically valid EQ can be regarded as the explication of a theoretical concept and, therefore, outset, such an 'operational' definition would be meaningless.

References

- Allport, G., 1961: Pattern and growth in personality. New York: Holt.
- Bem, D.J. & Allen, A., 1974: On predicting some of the people some of the time. *Psychological Review* 81, 506-520.
- Block, J., 1968: Personality measurement. Overview. In: *International Encyclopedia of the Social Sciences*, Vol. 12, 30-36.
- Brunswik, E., 1955: Representative design and probabilistic theory in a functional psychology. *Psychological Review*, 62, 193-217.
- Cattell, R.B., 1946: The description and measurement of personality. New York: World Book.
- Campbell, D.T., 1963: Social attitudes and other acquired behavioral dispositions. In: S. Koch (Ed.), *Psychology: A study of a science*, Vol. 3. New York: McGraw-Hill, 94-172.
- Cronbach, L. & Meehl, P., 1955: Construct validity in psychological tests. *Psychological Bulletin* 52, 281-301.
- Fiske, D. & Pearson, P.H., 1970: Theory and Techniques of Personality measurement. *Annual Review*, 49-86.
- Graumann, C.F., 1960: Eigenschaften als Problem der Persönlichkeitsforschung. In: Ph. Lersch & H. Thomae (Eds.), *Handbuch der Psychologie*, Band 4, Persönlichkeitsforschung und Persönlichkeitstheorie, Göttingen: Hogrefe, 87-154.
- Hartshorne, H. & May, M.A., 1928: Studies in the nature of character. Volume 1. Studies in deceit. New York: Macmillan.
- Helm, J., 1960: Über Gestalttheorie und Persönlichkeitspsychologie. In: Ph. Lersch & H. Thomae (Eds.), *Handbuch der Psychologie*, Band 4, Persönlichkeitsforschung und Persönlichkeitstheorie, Göttingen: Hogrefe.
- Hempel, C.G., 1958: The theoretician's dilemma: A study in the logic of theory construction. In: H. Feigl, M. Scriven & G. Maxwell (Eds.), *Minnesota Studies in the philosophy of science*. Vol. II. Minneapolis: University of Minnesota Press, 37-98.
- Hilke, R., 1980: Zu den impliziten "anthropologischen" Voraussetzungen der klassischen Testtheorie. *Diagnostica* XXVI, 99-111.
- Keasey, Ch. B., 1974: The influence of opinion-agreement and qualitative supportive reasoning in the evaluation of moral judgments. *Journal of Personality & Social Psychology* 30, 477-482.
- Kempf, W., 1978: Kritische Bemerkungen zu impliziten Voraussetzungen psychologischer Testtheorie und -praxis und ihre Angemessenheit zur Erfüllung der

- Aufgaben curricularer Evaluation. Vortrag zum Symposium "Aspekten von Leerplanevaluatie", Veldhoven (Holland).
- Kohlberg, L., 1969: Stage and sequence. The cognitive-developmental approach to socialization. In: D. Goslin (Ed.), *Handbook of socialization theory and research*. Chicago: Rand McNally.
- Lewin, K., 1963: *Feldtheorie in den Sozialwissenschaften*. Bern, Stuttgart: Huber (Original: *Field theory in social sciences*. New York: Harper 1951).
- Lind, G., 1977: Über die Grundlagen einer integrationistischen Meßtheorie zur Erforschung des moralischen Urteils. *Arbeitsunterlage 41*, Forschungsgruppe Hochschulsozialisation. Universität Konstanz.
- Lind, G., 1978: Wie mißt man moralisches Urteil? Probleme und alternative Möglichkeiten der Messung eines komplexen Konstrukts. In: G. Portele (Ed.) *Sozialisation und Moral*. Weinheim: Beltz 1978, S. 171-201.
- Lind, G., 1980 a: Zur Bestimmung des Entwicklungsstandes der moralischen Urteilskompetenz beim Übergang vom Gymnasium auf die Universität. In: J. Domnick (Ed.), *Aspekte grundlagenorientierter Bildungsforschung*. Forschungsberichte 18, Zentrum I Bildungsforschung, Universität Konstanz.
- Lind, G., 1980 b: *Toward an evaluation of the impact of university on students' moral development*. Unpublished manuscript, Universität Konstanz.
- Lind, G. & Wakenhut, R., 1980: Erfassung von moralischem Urteil mit standardisierten Fragebogen. *Diagnostica*, in press.
- Lind, G., Nielsen, A. & Schmidt, U., 1976: *Moralisches Urteil und Hochschulsozialisation Materialien, Beiträge*. *Arbeitsunterlage 40*, Forschungsgruppe Hochschulsozialisation. Universität Konstanz.
- Lind, G., Sandberger, J.-U., & Bargel, T., 1980: *Moral judgment, ego-strength and democratic orientations. Some theoretical contiguities and empirical findings*. *Arbeitsunterlage 64*, Forschungsgruppe Hochschulsozialisation. Universität Konstanz.
- Loevinger, J., 1957: Objective tests as instruments of psychological theory. *Psychological Reports* 9, 635-694.
- Loevinger, J., 1976: *Ego Development*. San Francisco, Jossey Bass.
- Lumsden, J., 1976: Test theory. *Annual Review of Psychology* 27, 251-280.
- Mead, G.H., 1968: *Geist, Identität und Gesellschaft*. Frankfurt: Suhrkamp (Originally published: 1934).
- Meehl, P.E., 1978: Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft Psychology. *Journal of Consulting and Clinical Psychology* 46, 806-834.
- Miller, G.A., Galanter, E. & Pribram, K.H., 1960: *Plans and the structure of be-*

- havior. New York: Holt.
- Mischel, W., 1968: Personality and assessment. New York: Wiley.
- Mischel, W., 1977: On the future of personality measurement. *American Psychologist*, 246-254.
- Nunner-Winkler, G., 1978: Probleme bei der Messung des moralischen Urteils mit standardisierten Verfahren. In: L.H. Eckensberger (Ed.), *Entwicklung des moralischen Urteilens*. Universität des Saarlandes, Saarbrücken: Universitätsdruck.
- Olweus, D., 1976: Der moderne Interaktionismus von Person und Situation und seine varianzanalytische Sackgasse. *Zeitschrift für Entwicklungspsychologie und pädagogische Psychologie* 8, 171-185.
- Perry, W.G., 1970: Forms of intellectual and ethical development in the college years. New York: Holt.
- Piaget, J., 1973: Erkenntnistheorie der Wissenschaften vom Menschen. Frankfurt: Ullstein (Original 1970).
- Popper, K., 1979: Objective Knowledge. An evolutionary approach. Oxford: At The Clarendon Press (Revised Edition).
- Rest, J.R., 1973: The hierarchical pattern of moral judgment: a study of patterns of comprehension and preference of moral stages. *Journal of Personality* 41, 86-109.
- Stern, W., 1911: *Die Differentielle Psychologie in ihren methodischen Grundlagen*. Leipzig: Barth.
- Travers, R.M., 1951: Rational hypotheses in the construction of tests. *Educational and psychological Measurement* 11, 128-137.
- Warren, H.C. & Carmichael, L., 1930: *Elements of human psychology*. Boston: Houghton Mufflin (Revised edition), pp. 11, 128-137.