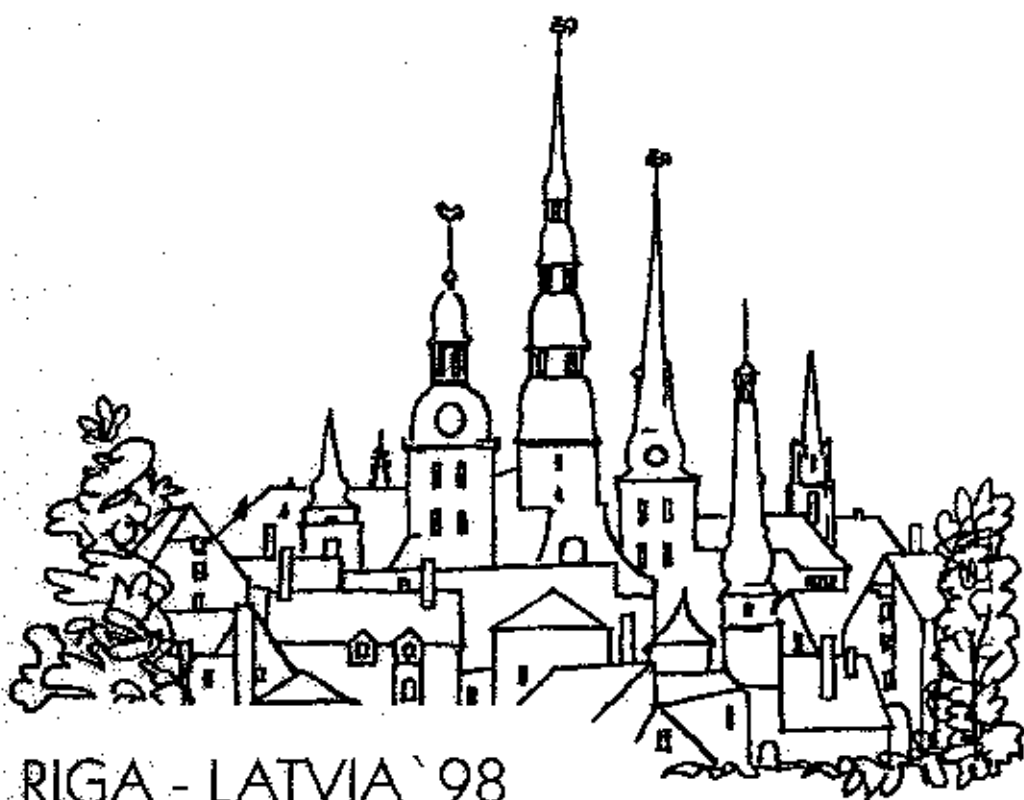




IDENTITY AND SELF-ESTEEM

Interactions
of Students, Teachers,
Family and Society



RIGA - LATVIA '98

Scientific Papers of XXI International
School Psychology Colloquium

PSYCHOLOGICAL EXPERT ASSESSEMENTS FOR FAMILY COURT DECISIONS IN GERMANY

Werner G. Leitner, Germany

Family psychological reports represent a basis for decisions with lasting consequences for the life of children after separation and divorce. Through the link between custody and educational rights they also have a decisive influence on the cooperation between the custodial person and school, and between custodial person and school psychological service.

Within the framework of interviews across Germany, family psychological assessments used by courts were evaluated in two steps, using among others the following criteria: quality criteria for the various test procedures, taxonomy of the behaviour study and the exploration and transparency of the underlying scientific theories and diagnosis. According to the results, the majority of assessments do not satisfy the imposed criteria because reliable, valid and at the same time objective tests are lacking. A systematic behaviour observation has hardly been carried out and the scientific concepts (missing references to literature) are mostly not exhibited. New relevant research regarding, for example, the Parental-Alienation Syndrome (see *Gardner*, 1992) is hardly included. The validity of such decision aids can not be regarded assured and should also be critically questioned by school psychologists.

According to the guidelines for making psychological assessments (*Foederation Deutshner Psychologenvereinigungen*, 1994) such an assessment represents "a scientific accomplishment which consists of making statements on specific questions based upon scientifically recognized methods and criteria, according to fixed rules for the exploration and interpretation of data".

Therefore, such a report must fulfill several necessary conditions of elementary scientific requirements. It has to be mentioned in connection with this that a psychological assessment has to be based not only on a single but always on several scientifically provable data sources (see *Foederation Deutshner Psychologenvereinigungen*, 1994, 11).

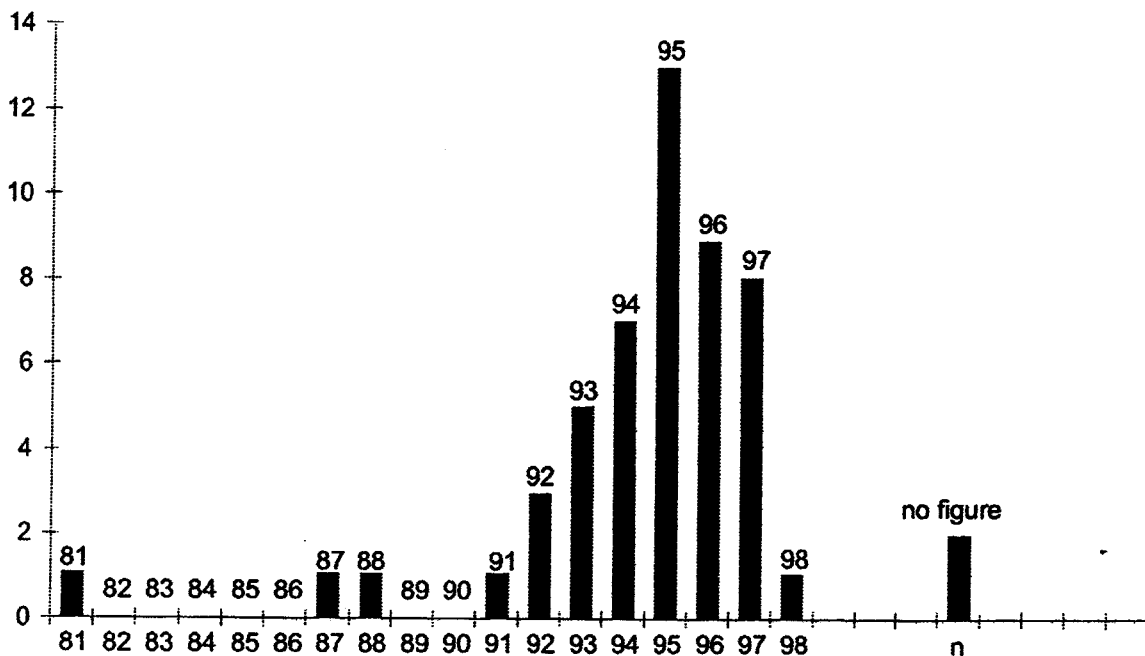
The study presented here shall, after several qualitative preliminary remarks with regard to further basic requirements and their realisation, give information with a quantitative analysis. It is to inform to what

extent the basis for test diagnosis, that means the tests used in a sample survey of N=52 reports, corresponds to the elementary standards of the main test criteria (reliability, validity, objectivity) and the side test criteria standardization, and with regard to the validity of their expression.

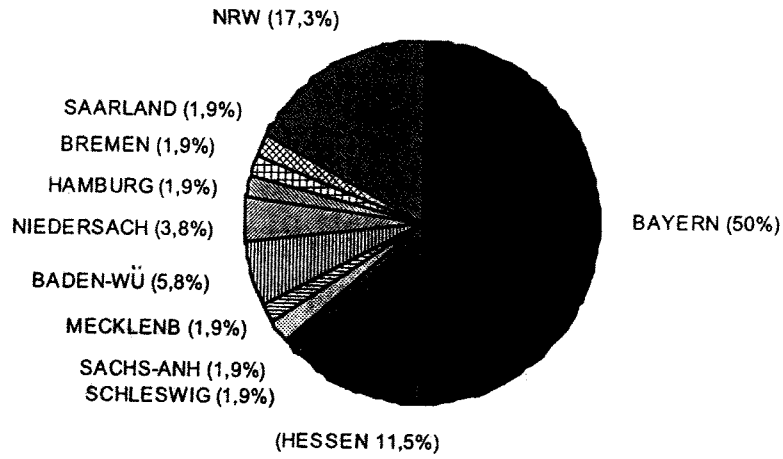
The sample that was made in cooperation with an association of family law is based on a federal survey (see *Leitner & Goethe, 1997/1998*) and it is dedicated to the memory of Siegbert Goethe who died in June 1998 shortly before the results were evaluated. The survey was carried out between 1994 and 1998 and consists of 52 reports that are from Bavaria (50%) and from the other "Bundeslaender" (German federal states) including the new German federal states (former German Democratic Republic) (50%).

The symmetric structure of the survey can also be seen in the organisational structure as the following graph shows. Copies of the original reports were submitted so that the results can be proved this way.

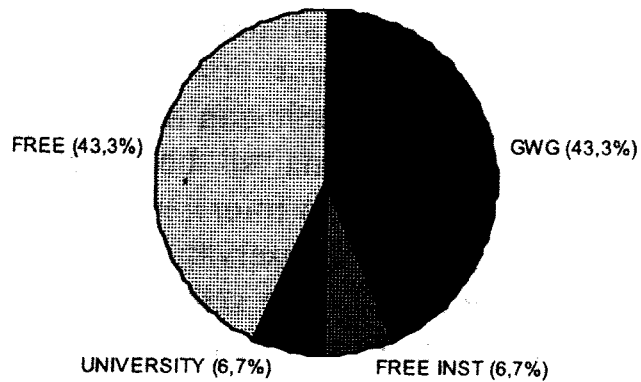
ANNUAL FIGURES N=52 family psychological assessments



SAMPLE
N=52 reports: German federal states



SAMPLE
N=52 expert assessments



Qualitative analysis

Document analysis, exploration, and behaviour observation

In order to speak of scientific research concerning the basic data, these data must have been assimilated with systematic methods. A casual behaviour observation would not be considered to be a systematic method as it is explained later more clearly.

Questions

Was a systematic document analysis carried out on the basis of a category system or was the viewing and representation of the documents on hand unsystematic and obviously at the expert's discretion?

Were the interviews carried out on the basis a systematic exploration manual and/or recorded and analysed with a category system or are no specific methods mentioned here?

Was a systematic (on the basis of a rating scale or a category system) or only casual (unsystematic, naive) behaviour observation carried out?

DORSCH et.al. (1994, 100) states the following:

“The taxonomy of behaviour observation is done with several viewpoints: (1) systematic (scientific) observation – unsystematic (naive, unscientific) observation”. Therefore, it is clear that a casual or unsystematic behaviour observation does not represent a scientific method because it does not systematically try to leave out “observation errors” (see *Dorsch et al 1994, 100*). This is similar to document analysis and interviews.

Not in a single report of the sample can the comprehensible use of an especially developed scientific analysis be observed regarding the analysis of documents. With regard to interviews, scientific approaches – as they are demonstrated in the conservation manual of *Westhoff & Kluck (1991/1992)* – are not used. The scientific and methodological paradigm or experts are seldom mentioned. It is almost unclear or not recognizable to see a theory with corresponding citations of sources and the mentioning of relevant findings (apart from the tests that we will talk of later). Therefore, the use of a rating scale is not shown in a single report (N=52). The reports about this sample do not show systematic (scientific) behaviour observation. Elements of a unscientific

(unsystematic) behaviour observation appear in a relatively unstructured form; those elements can not be evaluated as a scientific method and they do not try to exclude observation errors. In the sample there are no scientific methods which were applied with document analysis, interviews and behaviour observation and which would do justice to scientific basic requirements.

Literary assimilation principles of a scientific study

A family-psychological report is to fulfill basic scientific requirements also with regard to elementary literary principles. It is therefore required that there are bibliographical references also about the theories and concepts of interpretations. Literal citations from sources (documents, literature) have to be indicated in the text. A report that disregards those requirements can not expect to be “a scientific achievement” (*Foederation Deutscher Psychologenvereinigungen*, 1994, 8) – as it is required for such a report in the guidelines of the Federation of German Psychological Associations. Only a small part of the reports take those requirements into adequate consideration.

Quantitative Analysis

Test

The term “test” is often used colloquially for situations that can not scientifically be called by that term. A special example are the various tests in magazines that at best entertain but they have nothing in common with a test in a scientific sense.

Leinert (1969, cited in *Dorsch et al.*, 1994, 793f) decides about the criteria for test quality whether the tests are carried out in a scientific sense or not: “According to *Lienert* (1969) it can be differentiated between main and side test criteria. Objectively, reliability and validity are main test criteria. Economy, usefulness, standardization (scale) and comparability are side test criteria. Test methods that do not fulfill those criteria cannot be called a test in a scientific sense because they lack the foundations that are examined with scientific methods and they lack the control examinations that are necessary for test constructions”.

Test criteria

According to *Leinert* (1969) test criteria are important for the decision whether a test is a test in a scientific sense.

Objectivity

In Anglo-American literature, objectivity as a criterion is not always used consistently. *Catell* (see *Dorsch et al.*, 1994, 525), for example, called those tests objective, which are not transparent. Therefore, some authors emphasize objectivity as a uniformity for tests. *Lienert* (1969), cited in *Dorsch et al.*, 1994, 525) defines objectivity in German-speaking countries as “the degree to which the test results are independent from the examiner” (see above). This definition emphasizes “interpersonal correspondence” (see above). This correspondence plays a major role in the carrying-out (carrying-out objectivity), the evaluation (evaluation objectivity) and interpretation of tests.

Reliability

Reliability is about the reliability of a measuring method or a standardized test and it shows to which degree of accuracy the results are obtained with such a method. The accuracy of measuring is examined regardless of the validity of its results. The concept of reliability takes into account that each measuring has some errors and the results contain a part of true and error variance. The reliability coefficient is the quotient of true variance and the total variance. With regard to the method of reliability determination, various aspects of reliability can be differentiated with the retest method, the equivalence method and the split-half method (see *Dorsch et al.*, 1994, 656f).

Validity

In contrast to reliability, validity examines not only the degree of accuracy but also how the test method measures what it wants to measure (see *Lienert*, 1969, cited in *Dorsch et al.*, 1994, 838). Some kinds of validity can be distinguished: intrinsic validity, criterion-oriented validity and construct validity (see *Dorsch et al.*, 1994, 838).

Side test criteria

Apart from the three main test criteria objectivity, reliability and validity, standardization (scale) is the most important side test criterion.

The normative scales on hand are therefore called “standardized test scales for the economic comparability of test values” (*Dorsch et al.*, 1994, 519). These standards of comparison can be drawn up because of the distribution of raw scores and they are the basis for diagnosis whose usefulness contributes to the decision about quality and value of a test method.

*Family psychological Assessments for courts in Germany:
Do the applied tests fulfill the test criteria?*

Frequency of the tests in N=52 family psychological reports

Only the tests that were used in at least two different reports were evaluated. The following graphic shows the frequency of the methods with a differentiation with regard to its origins and organizational structure.

Test abbreviations according to their frequency

- | | |
|--|--|
| 1. 1. FiT
Familie in Tieren (family as animals) (<i>Brem-Graeser</i> , 1995)
Quality criteria according to <i>Brickenkamp</i> (1997, 964, 968ff)
Objectivity: no
Reliability: no
Validity: no
Standardization: partly | Not listed |
| 2. FRT
Family-Relations Test (<i>Bene & Anthony</i> , 1964)
Quality criteria according to <i>Brickenkamp</i> (1997 and 1983, 1997) | 3. DUESS
Fabelmethode (fable method) (<i>Duess</i> , 1964)
Quality criteria according to <i>Brickenkamp</i> (1975, 520ff)
Objectivity: no
Reliability: no
Validity: no
Standardization: yes |
| | 4. SERGT
Satzergänzungstest (incomplete sentences)
Quality criteria according to |

Brickenkamp (1975 and 1983, 1997)

Not listed

5. CAT

Children Apperception Test
(*Bellak & Bellak*, 1955)

Quality criteria according to
Brickenkamp (1997, 942, 949f)

Objectivity: no

Reliability: no

Validity: no

Standardization: no

6. MZT

Mann-Zeichen-Test (draw a
person) (*Ziler*, 1996)

Quality criteria according to
Brickenkamp (1997, 964, 973ff)

Objectivity: yes

Reliability: partly

Validity: yes

Standardization: yes

7. SCHLOSS

Schloss-Zeichen-Test (castle
drawing test)

Quality criteria according to
Brickenkamp (1975 and
1983, 1997)

Not listed

8. SCENO

Scenotest (*Von Staabs*, 1992)

Quality criteria according to
Brickenkamp (1997, 964, 981)

Objectivity: partly

Reliability: partly

Validity: partly

Standardization: no

9. FPI

Freiburger

Persoenlichkeitsinventar

(Freiburg Personality Inventory)

(*Fahrenberg, Hampel & Selg*,
1994)

Quality criteria according to
Brickenkamp (1997, 505, 557f)

Objectivity: yes

Reliability: yes

Validity: yes

Standardization: yes

10. FBT

Familien-Beziehungs-Test (fam-
ily-relations test) (*Howells &*

Lickorish, 1994)

Quality criteria according to
Brickenkamp (1997, 942, 945f)

Objectivity: no

Reliability: no

Validity: yes

Standardization: no

11. HAWIK

Hamburg-Wechesler-

Intelligenztest fuer Kinder

(WISC) (*Tewes*, 1985)

Quality criteria according to
Brickenkamp (1997, 49, 120f)

Objectivity: yes

Reliability: yes

Validity: yes

Standardization: yes

12. SF

Schwasrzfuss Test (*Corman*,
1992)

Quality criteria according to
Brickenkamp (1997, 942, 955f)

- Objectivity: no
 Reliability: no
 Validity: no
 Standardization: no
13. VF
 Verzauberte Familie (enchanted family) (*Kos & Bierman, 1995*)
 Quality criteria according to *Brickenkamp (1997, 964, 985f)*
 Objectivity: partly
 Reliability: no
 Validity: yes
 Standardization: partly
14. CFT 1
 Culture Fair Intelligence Test – Scale 1 (*Cattell, 1980*)
 Quality criteria according to *Brickenkamp (1997, 49, 103f)*
 Objectivity: yes
 Reliability: yes
 Validity: yes
 Standardization: yes
15. HANES
 Hamburger Neurotizismus und Extraversionsskala für Kinder und Jugendliche (Hamburg neuroticism and extraversion scale for children and teenagers) (*Buggle & Baumgartel, 1975*)
 Quality criteria according to *Brickenkamp (1997, 505, 561f)*
 Objectivity: yes
 Reliability: yes
 Validity: yes
 Standardization: yes
16. PFT
 Rosenzweig-P-F-Test (*Rosenzweig, 1957*)
 Quality criteria according to *Brickenkamp (1997, 942, 951f)*
 Objectivity: yes
 Reliability: yes
 Validity: yes
 Standardization: yes
17. CPM
 Coloured Progressive Matrices (*Becker, Schaller & Schmidtke, 1980*)
 Quality criteria according to *Brickenkamp (1997, 51, 192f)*
 Objectivity: yes
 Reliability: yes
 Validity: yes
 Standardization: yes
18. TAT
 Thematischer Apperceptionstest (Thematic Apperception Test) (*Murray, 1991*)
 Quality criteria according to *Brickenkamp (1997, 942, 957ff)*
 Objectivity: yes
 Reliability: yes
 Validity: no
 Standardization: no

Commentary on the Family-Relations-Test (Bene & Anthony)

Together with the test “family as animals” (*Brem-Graeser*) that does not meet the expectations concerning the test criteria, the Family-Relations-Test of *Bene* and *Anthony* (1957) is the most often used method. The FRT is not listed in the test manual by *Brickenkamp* (1997). The method holds this top position because it is often used in reports by the Gesellschaft für wissenschaftliche gerichts- und Rechtspsychologie (Society for scientific court and legal psychology) (GWG). Twelve out of the 16 applications of this method refer to those reports. This test shows especially that organisation specific peculiarities can be observed in certain assessments. Due to its frequent use and those organisation specific peculiarities a few comments should be made here:

The FRT is a test method that deals with four to five-year-old and six to eleven-year-old children according to a translation by *Flaemig* and *Woerner* (1997) (see *Beelmann*, 1995, 38). *Beelmann* talked and discussed about “new examinations with the Family Relations Test” at a conference of the expert group for developmental psychology of the Deutsche Gesellschaft fuer Psychologie e. V. (German Society for Psychology) in Leipzig in 1995. It became apparent that the validity of this method can not be assured at the moment. In his speech and the following discussion *Beelmann* called the work with this method in diagnostic reality shocking and he continued that due to economic reasons inadequate modifications are done in the practical application.

Conclusions

Test Diagnostic Basis

If one of the main quality criteria can not be assured, the validity of the results based on the test diagnosis are severely doubted. The impaired reliability is of a major consequence especially if the test was only carried out in the framework of a status diagnosis in a short period of time or only at a specific point of time.

The results obtained on this basis can not claim scientific requirements in a decision-oriented diagnosis because they are not provable valid, reliable, and objective.

Brickenkamp (1997, XXIV) says “that the hypothetical character of the test results increases with diminishing reliability and validity and

it needs to be secured by other results". Other scientifically based and comprehensible valid results can hardly be found in reports of the sample survey on hand. The speculative character of corresponding results becomes apparent when the main test criteria are not sufficiently guaranteed.

Unfortunately, this is often the case with frequently used test methods (e.g. Family as animals: position No.1)

Several Diagnostic Levels

Since the reports of the sample do not meet the scientific requirements, they prove to be inadequate as a basis for decisions in custody and visitation rights that have a significant influence on the future environment and education of children.

The results and suggestions made here are lacking elementary scientific requirements.

Besides the insufficient scientific basis of the psychological examinations, important pedagogical principles (pedagogical theories and concepts) can not explicitly be found in the reports of this sample. Not a single report deals with the new findings on Parental-Alienation-Syndrome (PAS) or with the concept itself (see *Gardner, 1992*).

The minister of justice of the Federal Republic of Germany told me in a personal letter of February 28, 1997 that he regards my "suggestion to emphasize educational reports worth considering". This concerns mostly courts (*Schmidt-Jortzig, 1997*, personal letter to the person signing). "The decision if the use of an expert with educational qualifications or a psychological expert is required is the work of the court" (*Schmidt-Jortzig, 1997*).

The statements made in the brochure of the GWG (Society for scientific court and legal psychology) "Ein Institut stellt sich vor" (an institut introduces itself) (GWG, without year, 16) can be confusing due to the results, because newly 50% of the reports of this sample were from the GWG. The standards mentioned in the brochure of the GWG, namely, "Comprehensibility of diagnostic procedures for the persons affected and the courts", "transparency of the diagnostic procedure for the persons affected" and the "strict concern with reliability of diagnostic procedures" are hardly confirmed by the results of this analysis.

Literature

Beelmann, W. (1995) Neure Untersuchungen mit dem Familz Relations Test – Normierungen und Anwendung in Scheidungsfamilien, in Ettrich, K.U., Friedrich, G., Fries, M. und Witruk, E.) ed., Abstractband. Tagungen der Fachgruppen Entwicklungspsychologie und Pädagogische Psychologie. Leipzig, 38,

Boerner, K. (1991) Das Psychologische Gutachten, Weinheim und Basel.

Brickenkamp, R. (1975, 1997) Handbuch psychologischer und pädagogischer Tests, Göttingen.

Brickenkamp, R. (1983) Erster Ergänzungsband zum Handbuch psychologischer und pädagogischer Tests, Göttingen.

Dorsch, F., Häcker, H. und Stapf, K.-H. (1987, 1994) Psychologisches Wörterbuch, Bern.

Ettrich, K.U., Friedrich, G., Fries, M. und Witruk, E. (Hrsg.) (1995). Abstractband. Tagungen der Fachgruppen Entwicklungspsychologie und Pädagogische Psychologie, Leipzig.

Föderation Deutscher Psychologenvereinigungen. Richtlinien für Erstellung Psychologischer Gutachten, Bonn 1994.

Gardner, R.A. (1992) The Parental Alienation Syndrome. A Guide for Mental Health and Legal Professionals, Cresskill; New Jersey.

Gesellschaft für Wissenschaftliche Gerichts- und Rechtspsychologie GWG-Ein Institut stellt sich vor, Without year and place.

Leitner, W.G. (1995) Pädagogisch-psychologische Diagnostik. Interventionsgeleitete Befunderstellung im exemplarischen Einzelfall, Bamberg.

Leitner, W.G. und Linsler, J. (1994) Das Recht des Kindes in Trennungs – und Scheidungsfamilien. Tätigkeit einer neuen Bundesarbeitsgemeinschaft (ISUV/VDU). Abstract. Europäische Factagung zur Familienforschung, Universität Bamberg.

Leitner, W.G. und Linsler, J. (1994) Elternverlust – unabdingbares oder vermeidbares Kinderunglück nach Trennung oder Scheidung?, in: Zeitschrift

für Familienforschung. Familie und Familienprobleme im Wandel. Ed. For Staatsinstituts für Familienforschung/ Universität Bamberg. Sonderheft 1/6, 315-322.

Schmidt-Jortzig, E. (1997). Personal correspondence.

Westhoff, K. und Kluck, M.-L. (1991) Psychologische Gutachten schreiben und beurteilen, Heidelberg.

Westhoff, K. und Kluck, M.-L. (1992) Ein Beispiel für entscheidungsorientierte Hypothesenbildung bei gerichtlichen Fragen zur elterlichen Sorge, in: Westhoff, K. (ed.), Entscheidungsorientierte Diagnostik, Bonn, 28-34.

Westhoff, K. und Kluck, M.-L. (1992) Entwicklung eines Leitfadens für ein entscheidungsorientiertes Gespräch bei gerichtlichen Fragen zur Umgangsregelung, in: Westhoff, K. (ed.), Entscheidungsorientierte Diagnostik, Bonn, 35-44.

CERTIFICATE OF DISTINGUISHED CONTRIBUTION

The International School
Psychology Association

wishes to recognize
the outstanding contribution
made by

Werner G. Leitner

at the XXI Colloquium
in Jurmala, Latvia

ISPA * 98

