The Munich High Ability Test Battery (MHBT): A multidimensional, multi-method approach for diagnosis of gifted students

Christoph Perleth, Kurt A. Heller and Virginia Danahe Sanchéz Resa

Prof. Dr. Christoph Perleth, Universidad de Rostock, Instituto de la Psicología Educativa “Rosa y David Katz”,

Traducción: Virginia Danahe Sanchéz Resa, Madrid

Dirección de contacto:

Prof. Dr. Christoph Perleth
University of Rostock
Institute of Educational Psychology “Rosa and David Katz”
August-Bebel-Str. 28, D-18051 Rostock
Germany
Correo electronico: christoph.perleth@uni-rostock.de

Abstract/Resumen

After a brief introduction the theoretical basis of the Munich High Ability Test-Battery (MHBT) will be outlined in the first part of the article. The MHBT has been developed in the framework of the Munich longitudinal study of giftedness and talent. The MHBT includes not only cognitive predictors measuring several dimensions and types of giftedness concerning intellectual, creative or social abilities etc., but also giftedness-relevant non-cognitive personality and social moderators measuring interests, motivations, learning emotions, self-concepts or family and school climate, educational style, quality of instruction, etc. The MHBT-instruments (different scales and dimensions) are described in greater detail.

Before describing the characteristics of the MHBT in more details we discuss problems and principles of diagnostics with respect to the diagnostic of gifted children and youth.

In the third part of the article, after dealing with the objectivity, the reliability, and the validity of the MHBT, we discuss the standardization procedure including the development of grade-based T-norms respectively as well as several talent-profiles, e.g. of gifted achievers vs. underachievers, intellectual, creative, social talents or linguistic, math, science talent profiles etc. Finally, examples of talent search for gifted programs and case studies on the basis of MHBT should illustrate multidimensional identification procedures.

The MHBT fulfills the most relevant assessment tasks belonging to the gifted educational and counseling practice. The usefulness of the MHBT in the framework of giftedness research as well as of gifted program evaluation studies has also been proven in the last decade. Hence the MHBT offers many opportunities to assessing giftedness and talent.

At the end a short overview is given on other traditional instruments for the diagnosis of gifted children and youth. We end with some conclusions summarizing possibilities and limits of the diagnosis of the gifted.

Objectivos

En este texto presentamos una batería de tests psicológicos, the Munich High Ability Test Battery (MHBT), as a paradigmatic example for a multidimensional, multi-method approach for diagnosis of gifted students. For better understanding we give also short overviews about the underlying theoretical concepts and the Munich Giftedness Study as well as the psychometric principles of psychological measurement in general. Al final del texto mostramos las perspectivas de otros métodos de diagnosis de altas capacidades.

Índice

1. Introduction
2. Conceptual and Theoretical Perspectives
3. Problems and prerequisits for the diagnostic of gifted children and youth
3.1 General methodological considerations concerning the identification of gifted children and youth
3.2 Prerequisites for psychological measurement and identification methods
3.2.1 Objectivity
3.2.2 Reliability
3.2.3 Validity
3.2.4 Norms
3.2.5 Auxiliary quality criteria
4. The Munich High Ability Test-Battery (MHBT)
4.1 Instruments of the MHBT
4.2 Scoring of the MHBT-results
4.3 Psychometric quality of the MHBT-scales
4.4 Standardization of the MHBT battery
4.5 Talent search with the use of MHBT
5. Further, traditional methods for the identification of gifted children and youth
5.1 Psychometric tests
5.2 Teacher-Checklists
5.3 Nomination by parents
5.4 Nomination by peers
6. Conclusions

1. Introduction

The instruments of the Munich High Ability Test Battery – in German: *Münchner Hochbegabungstestbatterie (MHBT)* – have been developed within the framework of the Munich longitudinal study of giftedness (Heller & Hany, 1986; Heller, 1990, 1991, 1992/2001; Heller & Perleth, 1989; Heller, Perleth & Sierwald, 1990; Perleth, Sierwald & Heller, 1993; Perleth & Heller, 1994). Meanwhile selected scales used in the mentioned study are published in the MHBT by Heller and Perleth (2007a/b). Two MHBT-forms are available (in German): the MHBT-P for primary school level (grades 1-4) and the MHBT-S for secondary school level (grades 5-12). MHBT-translations into Chinese, Korean, and Thai are in process since 2006. But several MHBT-scales including KFT-HB (German version of the Cognitive Abilities Test (CogAT) for highly gifted students) are used not only in the mentioned Munich longitudinal study of giftedness started in 1985/86 but also in other investigations, e.g. on the role of creativity in science and technology (Heller, 1995a/b, 2002a, 2007; Hany, 1994), in several gifted program evaluation studies (Heller, 2002a/b, 2004; Neber & Heller, 2002; Heller & Reimann, 2002) or with respect to cross-cultural studies (Heller & Perleth, 2004; Perleth & Heller, 2007), among others.

The Munich longitudinal study of giftedness pursued three main goals:

(1) the development and trial of assessment instruments for the reliable and valid identification of gifted students (grades 1 to 12+);

(2) the analysis of achievement behaviors of gifted students under various conditions (variations of situations and demands);

(3) the longitudinal analysis of individual developmental processes of gifted children and adolescents including positive and negative socialization influences, critical life events, etc.

Methodological problems of identification depend not only on the definition of giftedness and talent but also on the employment purpose. Hence at first the Munich Model of Giftedness (MMG) will be described as reference model of the MHBT; for greater detail see Heller and Hany (1986); Heller (1992/2001, 2004, 2005); Heller, Perleth and Lim (2005). Then the structure of the MHBT including the scales (tests and standardized questionnaires) and factors analyzed will be described in greater detail. Examples of talent searches for gifted programs and individual case studies illustrate the identification design using the MHBT for different diagnostic purposes. Finally, the function of the MHBT in the practice of gifted counseling and education as well as in gifted program evaluation and talent research will be discussed.

2. Conceptual and Theoretical Perspectives

If one considers “giftedness” or “talent” – both terms used here simultaneously – to be the product of interaction between genetic and environmental factors, then – assuming differential influences on both sides – different types of giftedness or talent are to be expected. Gardner (1983), with his multiple intelligence theory, postulates seven (recently even nine or ten) types of giftedness. Renzulli's three-ring model of giftedness (1978) has been expanded by Mönks and van Boxtel (1985) to six factors including the social settings family, school, and peers. As seen in Figure 1, a general causal model can be sketched which also includes environmental factors. Conceived as a diagnostic-prognostic model, the predictor (giftedness) is on the left side with the performance behavior as criterion on the right.

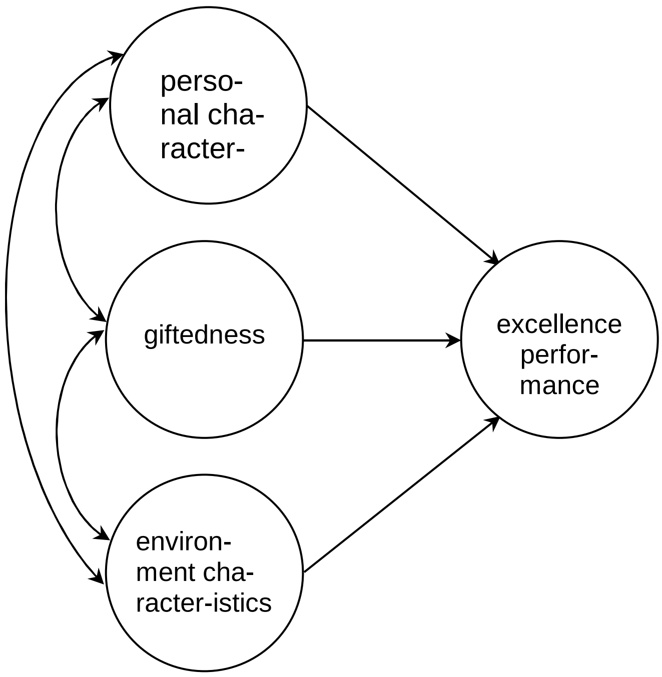

Figure 1: Causal model of performance behavior in the gifted and talented (according to Heller & Hany, 1986, p. 69)

Motivational and other non-cognitive personal traits which influence the relationship between ability or talent factors (predictors) and performance areas (criterion) in a relatively constant manner are important “mediators”, i.e. they serve as so-called moderators comparable to the “catalysts” in Gagné’s (2000) DMGT-conception. The moderators influence the transition of individual potentials (predictors) into performance (criterion) in various domains. For diagnostic purposes the moderators often play an indispensable role for explaining the relationship between predictors and criteria, e.g. the causal analysis of underachievement; see Figure 2. According to the MMG, giftedness or talent is conceptualized as a multifactorial ability construct within a network of non-cognitive (motivations, self-concepts, control expectations, coping strategies, etc.) and social moderators, as well as performance-related factors. For diagnostic purposes, the differentiation between predictor, criterion, and moderator variables is of particular interest.

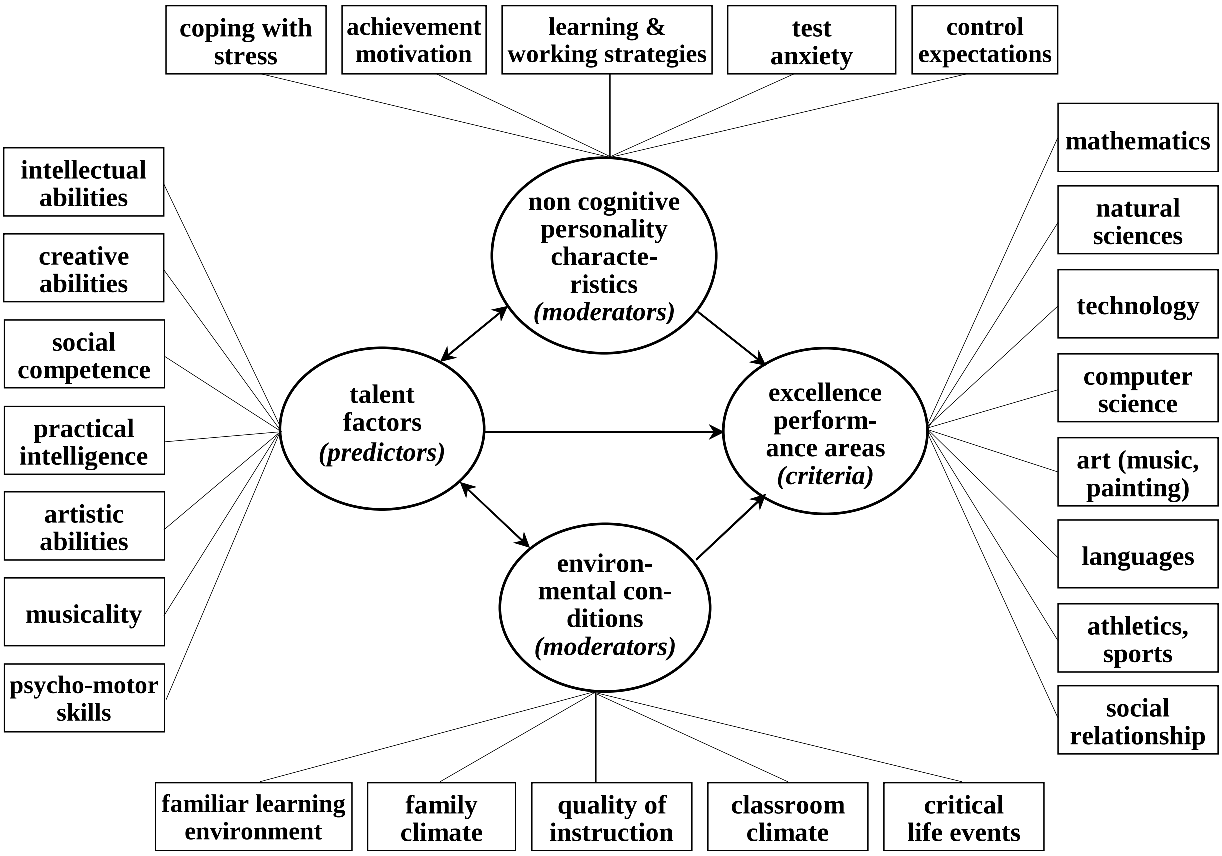

Figure 2: The Munich Model of Giftedness (MMG) as an example of multidimensional, typologicalconceptions of giftedness

Legend:

Talent factors (predictors)
– intelligence (language, mathematical, technical abilities, etc.)
– creativity (language, mathematical, technical, artistic, etc.)
– social competence
– musicality
– artistic abilities
– psycho-motor skills
– practical intelligence

(Noncognitive) Personality characteristics (moderators)
– achievement motivation
– hope for success vs. fear of failure
– control expectations
– thirst for knowledge
– ability to deal well with stress (coping with stress)
– self-concept (general, scholastic, of talent, etc.)

Environmental conditions (moderators)
– home environmental stimulation (“creative” environment)
– educational style
– parental educational level
– demands on performance made at home
– social reactions to success and failure
– number of siblings and sibling position
– family climate
– quality of instruction
– school climate
– critical life events
– differentiated learning and instruction

Performance areas (criteria variables)
– mathematics, computer science, etc.
– natural sciences
– technology, handicraft, trade, etc.
– languages
– music (musical-artistic area)
– social activities, leadership, etc.
– athletics/sports

3. Problems and prerequisits for the diagnostic of gifted children and youth

The objective of this excursus is to give an overview about methods and ways to identify gifted children and youth. There are different reasons why we want to know if a child has special talents and gifts. Generally we want to promote every student adequately according to his/her abilities and interests. With respect to gifted students the same principle holds through: We want to foster him/her according to his/her extraordinary abilities, performance and/or special interests. To choose the best strategy for promotion, however, it is necessary to identify specific talents or the giftedness profile of the individual. Another important and most frequent reason is the need to identify gifted students for special promotion programs. Ways to promote gifted can be enrichment courses for gifted children, competitions or acceleration programs such as special classes or schools (see Campbell, Wagner, & Walberg, 2000; Heller, 2001, 2002a; Neber & Heller, 2002; Renzulli & Reis, 2000).

In educational and counselling psychology it is also necessary to clarify adverse individual and social development conditions in order to offer special actions, for instance to improve attribution styles or motivation (Heller & Ziegler, 1996, 2001; Ziegler & Heller, 2000a, 2000b, 2000c). This means that the assessment of motivational and personal characteristics as well as of aspects of the learning environment is an important part of the diagnosis of giftedness. Further on, we want to analyse causes for underachievement in gifted children and youth (see Butler-Por, 1993; Peters, Grager-Loidl, & Supplee, 2000; Ziegler, Dresel, & Schober, 2000; Ziegler & Stoeger, 2003). It is also necessary to analyse social conflicts or behavioural and educational problems which could possibly be caused by giftedness (see Elbing, 2000; Freeman, 2000; Kaufmann & Castellanos, 2000; Webb, Meckstroth, & Tolan, 2002). Finally, special attention should be directed at risk groups, that is at children and youth whose giftedness is easily overlooked. Such groups are for instance gifted girls, handicapped gifted, children from ethnic minorities or children from underprivileged social groups (Borland & Wright, 2000; Kerr, 2000; Stapf, 2003; Yewchuck & Lupart, 1993).

3.1 General methodological considerations concerning the identification of gifted

The greatest concern when identifying gifted children and youth is to get as much “hits” as possible. “Hit” in this context means that somebody identified as gifted (blue subset in fig. 2) is gifted indeed. Reversed we want as little as possible to overlook gifted children (red subset in fig. 2) and to identify children falsely as gifted (Figure 2, see also Heller, 2000; Perleth & Sierwald, 2000).

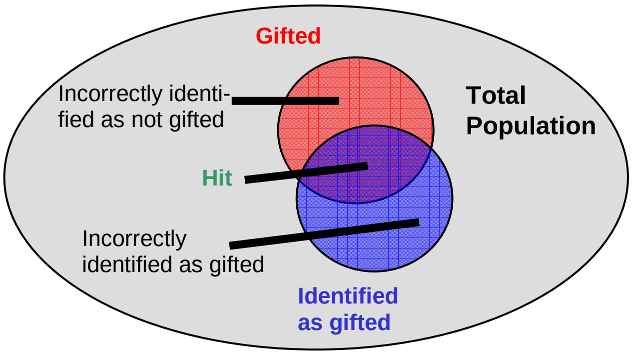

Figure 3: Problems with the identification of gifted children and youth

When identifying giftedness, we get in total four diagnostic groups (see table 1). Group A and D represent individuals who are correctly identified – either as gifted or not gifted. Diagnostic errors are represented in group B and C: students falsely identified as gifted and students whose giftedness is overlooked. It would be desirable if all individuals would be assigned to group A or D and diagnostic errors could be minimized. Unfortunately this is difficult to realise because diagnostic methods are never absolutely reliable and valid (see below). Depending on the aims of the diagnostic process we can only try to minimize errors in group C or B but have to accept that this leads to an increase of errors in group A or D.

Table 1: Diagnosis of giftedness

	Identified as		
	gifted	not gifted	total
gifted	A	B	A+B
not gifted	C	D	C+D
total	A+C	B+D	P=A+B+C+D

To exemplify this assume that a school wants to select a team of gifted students in mathematics to promote these students over a period of time. Finally they are going to take part in a math competition. In this case it would be useful to minimize group B in order to identify and promote every gifted student (unless the promotion would be very expensive). However, assume on the other hand a research project where we want to examine differences between highly gifted students and students with average abilities. Under these circumstances the group of gifted students should ideally contain only students who are gifted indeed. Therefore group C should be minimized, but inevitably there will be more cases in group B.

With this table we can assess the *effectivity* and *efficiency* of a measure to identify gifted children and youth. The effectivity can be described as "A/ (A+B) x 100" (see Table 1). This formula gives the percentage of gifted students who are identified correctly with the respective measure. On the contrary, the efficiency of a measure, calculated as A / (A+C) x 100 gives the percentage of gifted students in the as gifted identified group.

3.2 Prerequisites for psychological measurement and identification methods

When applying a psychological test we intend to get some information on the intellectual, emotional or motivational characteristics of the testees. These characteristics are conceptualized as constants in each person. Unfortunately we cannot directly measure the true values of these characteristics. Instead we have to estimate from test results the characteristics we are interested in. These tests are regarded as realisation of random variables; therefore they can differ more or less from the true values. In other words: The measured scores differ by a certain (normally distributed) error from the true values (see Perleth & Sierwald, 2000 for a more detailed overview on the mathematical and statistical considerations that underlie psychological tests). To minimize differences between test results and true values certain quality criteria for (psychological) research have to be met. These prerequisites for psychological measurement, including objectivity, reliability, validity, norms and test fairness, are introduced in the following section.

3.2.1 Objectivity

Objectivity is supposed to guarantee that the result of a test or diagnostic measurement does not at all or as little as possible depend on the test situation and the person conducting and/ or analyzing the test. Test authors usually try to ensure objectivity by prescribing in detail how to conduct, analyse and interpret the test. This aims at the highest possible comparability of test results between different students.

In order to obtain high objectivity during the conduction of a test instructions are often prescribed literally by the test author. If this would not be the case, different instructions could lead to different motivation and thus to different (i.e. wrong or misleading) results. In a drawing test on creativity for example an instruction can be focused rather on performance ("We want to know who can draw best.") or rather on pleasure ("You can now draw whatever you like."). Good tests also give detailed instructions about the best point in time for the conduction of the test (usually in the morning when students can reach their performance peak) and the arrangement of the test situation.

Another point concerns the analysis of the test results. Multiple choice tests usually ensure high objectivity. Tests with open answers need categorisation systems with detailed instructions how to classify answers as right or wrong (for instance free answers to the question "What have cat and dog in common?"). Answers like "both are pets" or "mammals" are surely correct, whereas answers like "Both lie on the sofa" or "Are called Garfield and Odie" are definitely wrong. But what about answers like "Have four legs and a fur" or "Both eat meat"?

The aim of an achievement test is to get to know if the result of a student is rather above or below average, that is if s/he is gifted or has intellectual deficiencies. For the objective interpretation of test results the handbooks of most tests contain norm (standard) tables with reference values. By means of these tables raw values can be transformed in standardized values which make the direct comparison possible between the individual result and a reference group or population (for instance the population of all twelve years old pupils). This ensures that test values are interpreted in the same way for all members of a reference group.

3.2.2 Reliability

The next quality criteria, reliability refers to the accuracy of a diagnostic measure. As stated above each test result is infected by a certain error. Therefore the measured value differs more or less from the true value of an individual. If the test user has information about the reliability of a test he can, for a given certain probability, estimate an interval around the measured value where the true value of the testee might be located. With other words: The reliability is necessary to compute the confidence interval which contains the true value of a person considering a specific probability value.

3.2.3 Validity

The validity of a test indicates how accurately a test measures or predicts the personality characteristics or behaviours that it is supposed to measure or predict. Only tests with high validity can be meaningful interpreted. Therefore the improvement of validity is one of the most important (and most difficult) aims when constructing a diagnostic measure. Apart from the fact that the test items must fit to the conception of the measured variable (a series of simple calculations would surely not measure intelligence) the following validity aspects can be differentiated.

Depending on the time between the collection of test and criterion scores *concurrent* and *predictive validity* can be distinguished. If test scores and criterion values are measured simultaneously or consecutively, concurrent validity can be obtained. With this validity coefficient conclusions can be drawn about simultaneous criteria. This is for instance important to analyse current achievement problems. If the criterion score is measured after the test score we refer to predictive validity. With this validity we can prognose from the test score (about academic achievement or job performance). Although both, predictive and concurrent validity, are computed in the same way, their different diagnostic meaning is to bear in mind.

3.2.4 Norms

The interpretation of test results is not possible without a frame of reference. There are three reference norms in psychological diagnostics:

- *Social reference norm* means that the test score is related to the scores of a reference group, e.g. all eight years old children. Measured here is the distance between test score and the mean of the reference group.

- *Individual reference norm* means that the result is related to previous performance of the student, measured is the individual advancement or deterioration of performance.
- *Task-oriented reference norm* finally means that the result is related to a criterion defined before (e.g. learning goal), measured here is the distance between performance and learning goal.

Most intelligence and performance tests use social reference norms, even if for the assessment of school performance the task-oriented norm would be more appropriate.

Standardisation of a test means the construction of a numerical reference system (usually tables) which allows to compare individual test scores (e.g. right answers) with the test scores of a reference population. This reference values are referred to as norms. Two points are crucial for the quality of standardisation. First, the sample for the standardisation of the test has to be chosen carefully. The sample must be representative of the test population. Second, during the standardisation test objectivity must be ensured. Therefore well trained examiners are required.

To construct norm tables test authors choose one of the common norm scales (see table 2). The mean of the reference group is then assigned to the mean of the scale; the variance of the reference group is assigned to the variance of the scale etc. Crucial for the interpretation of standard scores is not only the reference group but also how old the norms are. Norms conducted more then 10 years ago have to be considered as obsolete and must be verified by empirical studies.

Table 2: Important standard scales the interpretation of standard scores

Scale	M_X	s_X	<<Ø	<Ø	Ø	>Ø	>>Ø	example
IQ-Scale	100	15	x<70	70<=x<85	85<=x<=115	115<x<=130	130<x	WISC, CFT, SB V
WP-Scale	10	3	x<4	4<=x<7	7<=x<=13	13<x<=16	16<x	WISC
T-Scale	50	10	x<30	30<=x<40	40<=x<=60	60<x<=70	70<x	CAT
PR	50%	--	PR<2	2>PR<16	16<=PR<=84	84<PR<=98	98<PR	

Legend: <<Ø, <Ø, Ø, >Ø, >>Ø: far under average, under average, average, over average, far over average. Bold: important scales. *Note:* Of course there are more scales not mentioned in the table, e.g. SN-Skala (M_X=5, s_X=2) or the old scale of the Stanford-Binet IV (M_X=100, s_X =16).

3.2.5 Auxiliary quality criteria

Fairness of a diagnostic measure refers to the degree that test persons from different social, cultural, ethnic or psychological groups are treated fairly during the diagnostic process and that decisions based on this diagnostic process do not favour or discriminate a group. Since the 1970 years it is stipulated to include fairness as an additional quality criterion in psychological testing, especially for intelligence and achievement tests (see for instance Möbus, 1983). Some researchers consider test fairness to be a main quality criterion besides objectivity, reliability and validity (e.g. Stumpf, 1996).

Just like test fairness other auxiliary quality criteria in psychological diagnostics address as well the applicability of diagnostic measures in certain diagnostic contexts. First, a diagnostic measure or test should be *comparable* to other tests which measure the same construct. A good diagnostic measure should also be *economical*, that is it should be fast and inexpensive to apply (in terms of test material, time, number of examiners needed, individual- or group assessment) and analyse. It should also be *useful*: it should serve the target purpose and help to make decisions.

4. The Munich High Ability Test-Battery (MHBT)

4.1 Instruments of the MHBT

The MHBT has been developed on the basis of MMG which served as reference model. Therefore, the tests and questionnaires of the MHBT represent different scales measuring not only various aspects and types of giftedness (which serve as predictors) but also various non-cognitive personality and social-environmental learning conditions (which serve as moderators). The MHBT contains two dozen tests and standardized questionnaires for the differential assessment of the predictor and moderator variables illustrated in Figure 2. These variables are mostly relevant to the promotion and development of giftedness and talent. Multiple predictors and moderators are advocated because the excellence criterion is considered to be complex (see Figure 2).

The criterion excellence performance can be measured by means of school achievement tests and/or teacher ratings (e.g. school grades), etc. For a new performance-based assessment system see VanTassel-Baska, Feng and Evans (2007). Such scales are not included in the MHBT. In the diagnosis-prognosis paradigm, the criterion is to be predicted; see Heller (1989).

Furthermore, the checklists of the MHBT facilitate a rough estimation of individual talent levels for children and adolescents in the following six areas: intelligence, creativity, musicality, social competence, and psycho-motor ability, and can be used in the screening phase (see Table 2 below). For the complete MHBT including information about the test and questionnaire dimensions as well as the target age groups (grades) see Table 1.

Table 2: The complete MHBT (sub)scales and selected dimensions of MHBT-P/MHBT-S

MHBT-scales	MHBT-dimensions (selection)	Grades
Checklists (teacher ratings):		
Checklist re. **intellectual giftedness**	Thinking, learning capabilities, knowledge, etc.	1-12+
Checklist re. **creative giftedness**	creative thinking, originality, etc.	
Checklist re. **musicality**	acoustic sensibility, pitch differentiation, etc.	
Checklist re. **social competence**	cooperation skills, leadership, etc.	
Checklist re. **psycho-motor ability**	dexterity, hand skillfulness, etc.	
KFT-HB 3:		
V1 (vocabulary)	verbal abilities	3
V2 (word-classification)	verbal abilities	
Q1 (comparison of quantities)	quantitative (mathematical) abilities	
Q2 (equation forming)	quantitative (mathematical) abilities	
N1 (figure classification)	nonverbal (technical-constructive) abilities	
N2 (figure analogy)	nonverbal (technical-constructive) abilities	
GL (total score)	cognitive abilities level (intelligence)	
KFT-HB 4:		
V1 (vocabulary)	verbal abilities	4
V2 (word-classification)	verbal abilities	
Q1 (comparison of quantities)	quantitative (mathematical) abilities	
Q2 (equation forming)	quantitative (mathematical) abilities	
N1 (figure classification)	nonverbal (technical-constructive) abilities	
N2 (figure analogy)	nonverbal (technical-constructive) abilities	
GL (total score)	cognitive abilities level (intelligence)	
MHBT-inventory for primary school level (MHBT-P):		
KRT-P (questionnaire of creativity)	originality, flexibility, etc.	1-4
SK-P (questionnaire of social competence)	social cognitions	
LM-P (questionn. of achievement motivation)	hope for success vs. fear of failure	
AV-P (questionnaire of working behavior)	attentiveness, control of thinking processes, etc.	
KA (questionnaire of causal attribution)	success vs. failure attributions	
KFT-HB 4-12:		
V1 (vocabulary)	verbal abilities	4-12+
V2 (word-classification)	verbal abilities	
Q1 (comparison of quantities)	quantitative (mathematical) abilities	
Q2 (equation forming)	quantitative (mathematical) abilities	
N1 (figure classification)	nonverbal (technical-constructive) abilities	
N2 (figure analogy)	nonverbal (technical-constructive) abilities	
GL (total score)	cognitive abilities level (intelligence)	
MHBT-inventory for secondary school level (MHBT-S):		5-12+
AW (unfolding test)	spatial reasoning	
SP (mirror images)	spatial cognition	
APT (tasks of physics and technology)	problem solving in physics and technology	
KRT-S (questionnaire of creativity)	originality, flexibility, etc.	
SK-S (questionnaire of social competence)	social cognitions	
IFB (questionnaire of interests)	preferences of interests	
FES (questionnaire of thirst for knowledge)	curiosity as a preliminary form of striving for knowledge	
LM-S (question. of achievement motivation)	hope for success vs. fear of failure	
AV-S (questionnaire of working behavior)	attentiveness, control of thinking processes, etc.	
SCHUL (questionnaire of school climate)	aspects of school climate	
FAM (questionnaire of family climate)	aspects of family climate	

Legend:

KFT-HB = Cognitive Abilities Test for Highly Gifted Students
V = Verbal abilities
Q = Quantitative (mathematical) abilities
N = Nonverbal (technical-constructive) abilities

-P = Primary school level
-S = Secondary school level

4.2 Scoring of the MHBT-results

The scoring of the tests and questionnaires of the MHBT-battery is exclusively done with the help of a computer software. Therefore, the usual scoring with the help of stencils and norm tables is not possible. After entering each answer of the respective student in a formular (see Figure 3) one gets at once a lucid profile evaluation for founded analysis in the frame of the respective diagnostic problem – without complicated calculation and long winded work with norm tables (see Figure 4).

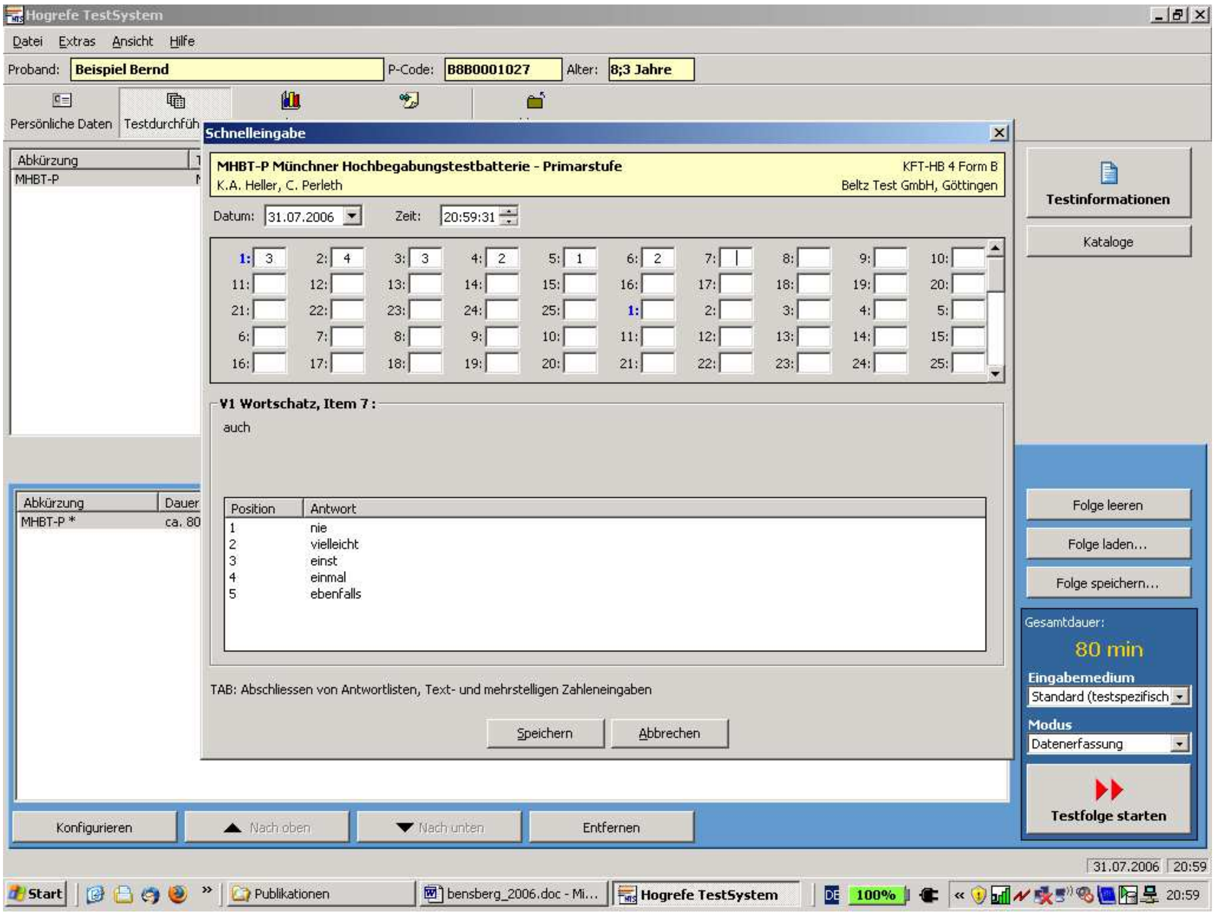

Figure 4: Data gathering with the scoring computer program of the MHBT (fictive example)

Legend: Software is available in German only. The screenshot shows the input mask where the chosen alternatives for the 25 items of subtest "V1 Wortschatz" (vocabulary) of the KFT-HB can be entered. You can either save ("Speichern") or cancel ("Abbrechen") the data of this mask.

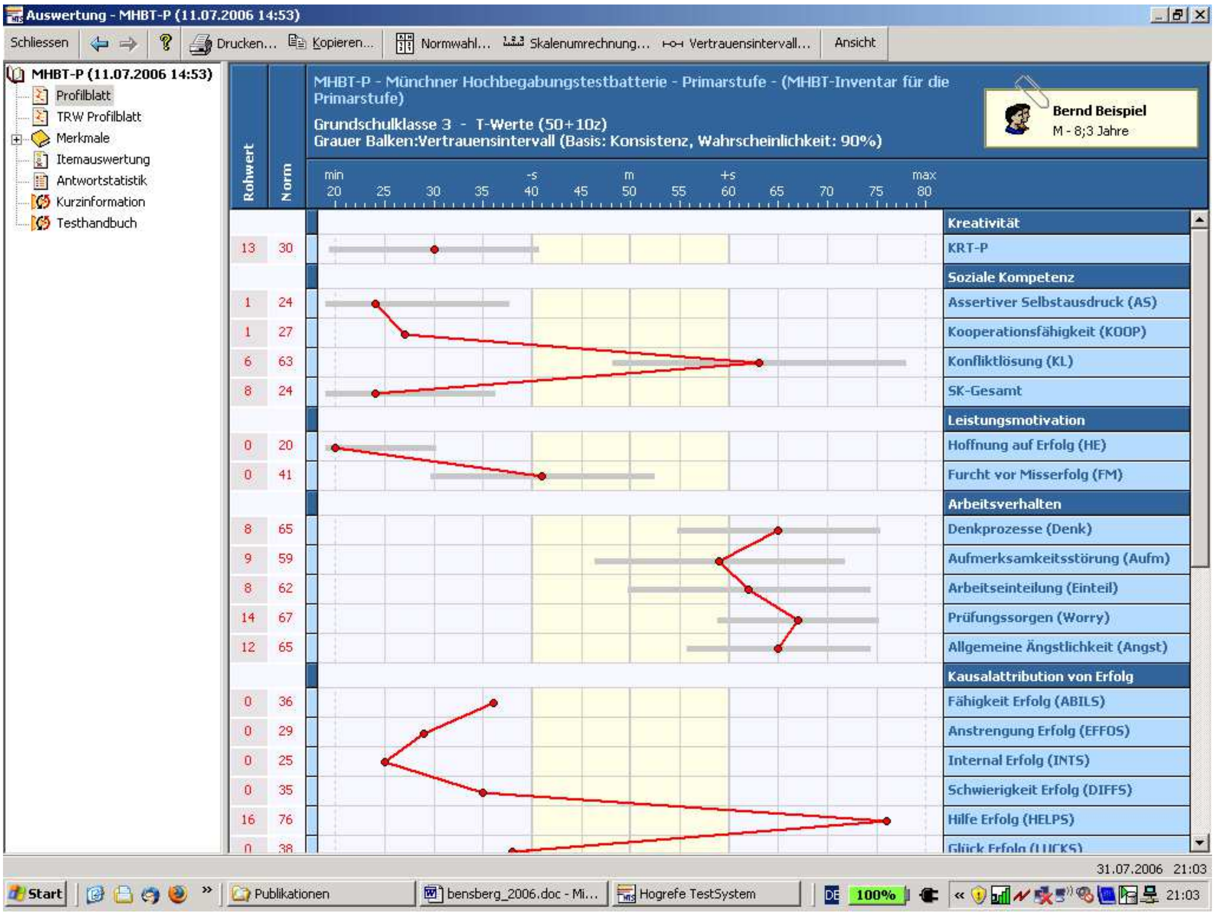

Figure 5: Profile analysis with the scoring computer program of the MHBT (fictive example)

Legend: Software is available in German only. For an explanation of the abbreviations of the different scales (blue fields on the right) see table 1 above. The red points and lines show the profile (T-scores), the grey whiskers represent the90 % - confidence intervals on the basis of consistence reliability.

This scoring computer program is integrated in the Hogrefe-Testsystem and offers a number of advantages for the practitioner:

- Scoring can be more easily done and mistakes can be avoided – as long as one correctly enters the data.
- In view of the complex structure of the MHBT with many dimensions and subscales this facilitation of the scoring gains even more importance.
- The computer program provides diverse possibilities of data management and
- the results or profiles of several students can be shown at the same time and this way can be very easily compared.

The scoring program can be installed only once on exactly one personal computer and for each case one has to pay a certain amount of licence fee. For this purpose one has to buy a certain amount of scoring cases. This procedure may be unfamiliar to those who have collected experiences only with traditional scoring programs which were available extra to the conventional scoring „by hand". However, even if one has to pay a licence fee for each use or scoring process, the scoring of the MHBT is not per se more expensive in comparison with conventional tests and questionnaires. For scoring the data of each student one has to pay just 3 Euro. Together with the costs of the maximum of 3 answer sheets (0.25 Euro each) the total costs for each case is rather moderate, especially if one considers the time one usually needs for scoring.

To prevent abuse and also because of the layout of the computer program in the framework of the Hogrefe-Testsystem it is not possible to edit the data of a single answer after the answer record of a certain case/student has been stored. It is, however, possible to inspect the answer pattern of a single case in detail. With other words: One can exactly reproduce which alternative a certain student crossed out but one cannot change or correct the answer of a single item after storing the respective case.

In the manual of the MHBT one can find a number of examples for individual diagnostics with varying contexts and for different counseling problems. This should support the practical use of the MHBT. The manual also contains examples for talent search, an example of which is given at the end of this article.

4.3 Psychometric quality of the MHBT-scales

With respect to the objectivity no bigger problems should arise with trained test instructors as detailed instruction are available. All tests and questionnaires are suited for application in groups so that the instructor-testee-interaction is reduced to a minimum. Nearly all scales use multiple-choice-format, the evaluation of the answers is done with the help of a special computer software. Of course, the teacher checklists available with the MHBT own a lower degree of objectivity, above all because the ratings given depend on the experience of the specific teacher.

Depending on the relative test or questionnaire in the framework of the Munich Study of Giftedness as well as for the standardization sample reliability coefficients between r = .40 (for example for some scales of the questionnaire for family climate, FAM) and r = .95 (e.g. for the scales of cognitive abilities, KFT-HB) were found. The KFT-HB-scales also showed astonishingly high stability coefficients over periods of one or two years. For the rather rough teachers' checklists (screening procedure) no systematic results concerning their reliability could be collected. However, there are some findings for very differentiated teachers' ratings (some had more than 100 items for 5 domains of giftedness) showing that their predictive validity are only little (not meaningfully) higher than the rather rough ones (Perleth & Sierwald, 2001).

With respect to factorial validity we found patterns which we expected, for example the three material factors typical for the KFT (Cognitive Abilities Test): A verbal, a quantitative, and a nonverbal-figural factor. The analysis of concurrent and predictive validity of the MHBT tests and questionnaires showed middle to high coefficients with teachers' ratings, school grades, Abitur grades (final school exam), first achievement during university study as well as diverse activities and achievement in leisure time activities (i.e. extracurricular activities). For some methods and groups of predictors quite high coefficients could be found for longer periods of time: For example in the framework of the follow-up-studies validity coefficients up to $r = .79$ could be found between the KFT-HB and first achievement at university; with respect to Abitur grades we found coefficients for predictive validity up to $r = .80$ for some of the subjects. The teachers' checklists showed sufficient concordance with test results. See Figure 5 for an overview on the Munich giftedness study including the follow up, and the standardization studies.

4.4 Standardization of the MHBT battery

For the MHBT grade specific norms have been computed on the basis of an unselected standardization sample of more than 4,000 students in total. Tables 2 and 3 should give an impression how the total standardization sample was divided for the standardization of the different scales and how the students were distributed with respect to school level, grade and sex. The German secondary school system is built up of three school types of different level: The "Gymnasium" (grade 5 to 12 or 13) is attended by about 30-40 percent of the students (11-18/19 years of age) and represents the highest level leading to university. This school form is chosen by a relatively high number of students with above average cognitive abilities. The "Realschule" (grade 5 to 10, i.e. age 11-16) represents a middle level, while the students of the Hauptschule, all in all, show lower school achievement; here you find also a high percentage of students from migrant families.

In order to get grade specific norms which differentiate good in the upper range of the respective scales a similar technique (stratification and rectification) was used as was done in the PISA-study. That means that the sample was recruited in a way that

- an over proportional percentage of students from the Gymnasium was included,
- and a relatively small percentage of students from the Hauptschule.

For the calculation of the norms, the sample was weighted following the correct percentage of the students of the different school types of the respective federal state. All norm tables contain T-norms which have been normalized by McCall's procedure. All tables give T-norms up to a maximum of *T* = 80, no extrapolation was used to get even higher T-norms because we are convinced that all norm tables should have an empirical foundation and should not belong to the genre of science fiction.

Further on, one finds in the manual of the test battery MHBT-profiles or standards on the basis of 332 gifted, highly achieving students as well as 134 underachieving students. These profiles or standards are given for both primary and secondary school age students. These gifted, highly achieving students and underachievers stem from different studies in which the MHBT was used.

The procedure used for standardization and computation of the norm tables was chosen in order to get a good differentiation especially in the upper range of the different scales, above all the abilities and achievement tests. For detailed profile analysis the standards/profile of the gifted and underachievers (see above) can be used. These standards/profile can not only be useful for identification and counseling of individuals but also for the identification of giftedness types as well as for talent searches (see below). The profile can also be useful for a detailed analysis of moderators or factors (or catalysts) which are useful for transformation of abilities in achievement. As shown above the MHBT provides a good number of scales for different motivational and other personality factors as well as scales for relevant variables of the family or school learning environment.

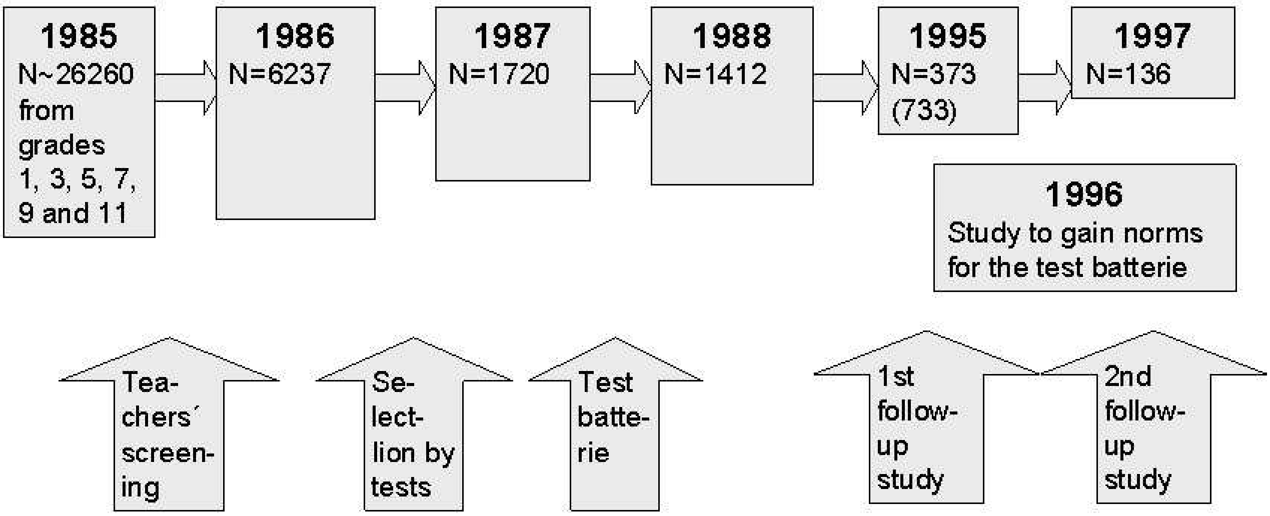

Figure 6: Sample and design of the Munich Giftedness Study

Table 3: Sample 1 (Standardization of the KFT-HB)

	Primary school/ Hauptschule	Realschule	Gymnasium	Total
Grade	Sex	Sex	Sex	Sex

	♂	♀	♂	♀	♂	♀	♂	♀
3	80/76 + 318	71/85 + 319					156 + 318	156 + 319
4	161	161					161	161
5	30/33	35/27	11/17	22/14	54/32	45/35	179	178
7	38/35	36/27	20/8	9/9	55/46	61/47	204	190
9	8/17	7/12	45/35	36/34	62/56	58/55	224	205
11					80/89	91/77	170	168

Legend: See the main text for the different German school types; ♂ = male, ♀ = female.

Table 4: Sample 4 (Standardization of the questionnaires SK-S, SP, AW, Fam, LM-S)

	Primary school/ Hauptschule		Realschule		Gymnasium		
Grade	Sex		Sex		Sex		Total
	♂	♀	♂	♀	♂	♀	(♂/ ♀)
5	58	67	28	30	16	13	235 (102/110)
7	43	30	30	26	55	43	232 (128/99)
9-11	10	9	22	16	69	74	206 (101/99)

Legend: See the main text for the different German school types; ♂ = male, ♀ = female.

4.5 Talent search with the use of MHBT

When regarding the diagnostic function of talent searches, it is necessary to be aware that the individual prerequisites and the demands of the new learning content in the advancement gifted program fit together (Heller, 1999, 2005; Pfeiffer & Jarosewich, 2007). Talent search in this sense means individual developmental help. A comprehensive assessment approach should, therefore, be an indispensable component of every talent search (Hany, 1993; Feldhusen & Jarwan, 2000). The MHBT-instruments represent the most relevant cognitive abilities (verbal, quantitative, nonverbal, technical, space and other factors) and noncognitive personality moderators (self-concept, action control, task commitment, achievement motivation, etc.) as well as social conditions of the learning environment (family and school climate, “creative” stimulation in the classroom, quality of instruction, etc.).

The first step in the identification process is usually a *screening* on the basis of teacher checklists (with rating scales) based on the operationalism of behavioral characteristics of domain-specific talents. In this way, a range as broad as possible of cognitive and motivational traits is determined which provides information about the presumed talent and assessed performances. Since ratings and other “soft” data can be assumed to be less accurate than *test* data, the screening should attempt to “lose” as few gifted candidates as possible for the concerned gifted program. This occurs through the conscious inclusion of non-too-small “false hits”.

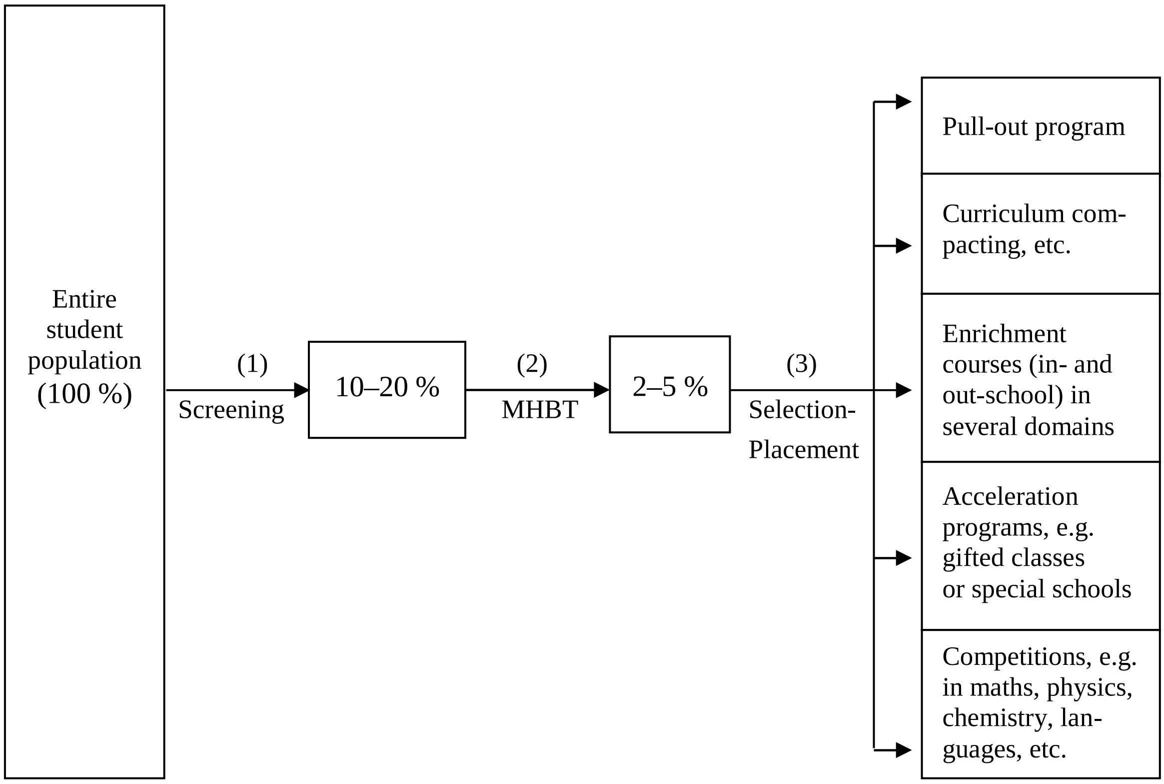

Figure 7: A sequential strategy model of the identification of gifted and talented students for educational programs according to Heller (2000, p. 252)

Legend:

(1) = Screening phase (e.g. by teacher checklists): Nomination of the 10-20 % class leaders with respect to different dimensions of giftedness and talent.

(2) = Tests and standardized questionnaires (MHBT) measuring different factors of giftedness and talent in the preselected group of the 10-20 %.

(3) = Final decision and assignment to various nurturing programs.

It will not be until the second or, if necessary, the third selection step – with the aid of more accurate measurement instruments that are, however, more limited in breadth – that a final selection can be made; see Figure 7 above. For greater detail see Heller (2004, 2005).

The following example of talent search illustrates the identification steps mentioned above. This paradigm has been applied and validated among very able students identified through MHBT for the “Hector-Seminar”, a gifted program in Mathematics, Informatics, Natural sciences, and Technology (MINT) carried out in the state of Baden-Württemberg (Germany). Depending on the main goal of the *Hector-Seminar* (furtherance of MINT-talents), at the first step, checklists have been applied for pre-selection of the top 10 % of the students in the German Gymnasium. The checklists focused on several aspects of intellectual, creative and social giftedness which are mostly relevant to MINT (see Table 4).

At the second step, the pre-selected top 10 % of the candidates of the gifted program "Hector-Seminar" have been tested by following MHBT-scales: KFT-HB V1, V2, Q1, Q2, N1, N2, AW, SP, APT, KRT-S, and SK-S (see Table 1 above). At the final step, the "Hectorians" could be recruited on the basis of a combinatory decision strategy (Heller, Senfter & Linke, 2006, pp. 13-15; Heller & Perleth, 2007b, pp. 133-140).

Table 5: The first step of talent search (here in the recruiting of the "Hectorians") based on teachers' checklists

	Possible criteria for	
intellectual giftedness	**creative giftedness**	**social competence**
o Logical/analytical thinking o Abstract thinking o Mathematical thinking o Scientific/technical thinking o Language skills (rich vocabulary, fluency of expression, talent for foreign languages) o Learning ability (quick understanding, retentive memory, accurate reproduction, active learning) o Powers of deduction, combination etc. o Broad knowledge o Consolidated special knowledge in one or more domains	o Curiosity, quest for knowledge o Imagination, ability to think in alternatives o Creative and inventive thinking o Originality, search for extraordinary problem/task solutions o Flexible thinking, spiritual agility, ability to consider a problem from various points of view o Self-sufficiency, independence of thinking and opinion o Interest-oriented, independent solving of problems o Multiplicity of interests o Stability of interests	o Social adaptability o Social cognitions o Self assertion, self confidence o Cooperation, conflict solving, etc. o Capacity for understanding, empathy, etc. o Initiatives in social contexts o Social accomplishments o Leadership o Social responsibility, integrability, etc. o Popularity with classmates, etc.

Note: To assess MINT-related types of giftedness, please nominate the top 10 % of the students in your class referring to the dimensions listed above. The criteria need not be all present; it is sufficient if the student excels in some of them.

Using the identification selection strategy described above, one runs into the bandwidth-fidelity dilemma according to Cronbach and Gleser (1965). All selection decisions are fallible; one can only attempt to choose the lesser evil in the personnel decision. The risk of *type I* errors exists here in identifying someone as gifted when he or she is not gifted. The risk of *type II* errors exists here in failing to identify someone as gifted when indeed they are. The *type I* error can be reduced by making the criteria more rigid, the *type II* error by making them less strict. Unfortunately, simultaneous reduction of both types is not possible. In order to maximize *individual* usefulness, it is better to minimize the *type II* error. For maximizing the gifted program usefulness, the *type I* error should be minimized; for greater detail see Heller (2004, 2005), Heymans and Mönks (2004).

5. Further, traditional methods for the identification of gifted children and youth

5.1 Psychometric tests

Psychometric measures (tests and questionnaires) for the assessment of high ability or rather intelligence are widely-used in research and practice to identify and nurture gifted children and youth. A test is a psychological instrument to measure a clearly defined personality characteristic (e.g. intelligence, retentiveness, concentration or anxiety) and is designed by scientists for practical application. Tests usually give numerical results, so that test results or parameter values of different persons can be compared directly (for an overview see Perleth, Schatz, & Mönks, 2001). It is particularly interesting to examine to what extent the test score of an individual differs from the mean score of a reference group. To get meaningful values it is necessary that all examinees are evaluated on equal terms. To ensure objectivity test items and test situation have to be standardized. Most test authors assume that psychological characteristics underlie a Gaussian or normal distribution. Most statistical measures to analyse empirical data (especially in the context of test construction) require a normal distribution as well. Taken intelligence gifted children and children with intellectual deviancies represent the extreme points in this distribution. So the gifted deviate just as far as the mentally retarded from normality.

Tests and questionnaires are still the best methods to predict performance. Intelligence tests for instance are the strongest predictors for school achievement. It must be pointed out that tests have more prognostic strength in the beginning of the academic career (during the education increases the importance of previous knowledge) and the intelligence structure becomes more differentiated over the years. Therefore tests of general intelligence are useful at the beginning of primary school, whereas during secondary school more differentiated tests should be applied which make profile analyses possible (see Hany, 2001; Heller, 2000; Perleth & Sierwald, 2000; Rost, 1993; Wild, 1991).

Psychometric tests should be applied retentively in very young children (pre-school age) because the prognostic validity is quite low and measured traits are rather unstable over time (Lubinski, Benbow, Webb, & Bleske-Rechek, 2006; Lubinski, Webb, Morelock, & Benbow, 2001). In particular there is a lack of longitudinal studies which could give useful information about the development of giftedness from an early age on. First major longitudinal studies like the Study of Mathematically Precocious Youth show however that there is a positive relationship between exceptional test scores (here the SAT was applied) before age 13 and occupational and life accomplishments in adulthood (Wai, Lubinski, & Benbow, 2005). Similar results have been reported by Perleth and Sierwald (2001) and particularly Perleth (2001; see also Heller & Perleth, 2004) for the Munich Giftedness Study. In the following paragraph we give examples of the most important tests used for identifying gifted children and youth.

Stanford-Binet Intelligence Scale

The Stanford-Binet Intelligence Scales: Fifth Edition (SB5) by Roid (2003) is the current version of the Stanford-Binet intelligence test originally developed by Terman on the basis of Binet and Simon's scale (Binet, 1905). The norm sample of the SB5 consists of 4,800 individuals between the ages of 2.0 and 96 years and matches the U.S. census from 2000 (Becker, 2003). The SB5 provides the examiner with an overall score for general intelligence which is composed of five factors: knowledge, fluid reasoning, quantitative reasoning, visual-spatial processing and working memory (Roid, 2003). Each factor contains both verbal and nonverbal content. The test can be used with children from the age of 2 years.

Wechsler Intelligence Scale

The Wechsler Preschool and Primary Scale of Intelligence – III (WPPSI-III) (Wechsler, 2002) and the Wechsler Intelligence Scale for Children, now in its Fourth Edition (WISC-IV) (Wechsler, 2003) are the second most widely used intelligence tests in gifted children. They measure "the aggregate or global capacity of the individual to act purposefully, to think rationally and to deal effectively with his environment" (Wechsler, 1958, p. 7). In consequence, the most frequently used indicators derived from the tests are the total score and until the third edition the global scores for the verbal and performance scale. Although there have been some studies of the factor structure of the WISC-III (see Brown, Hwang, Baron, & Yakimowski, 1991; Masten, Morse, & Wenglar, 1995; Wilkinson, 1993) alternative scores or the profile are seldom used for scientific or practical purposes. As in the Stanford-Binet, the Wechsler scales have been shown to be sufficiently reliable and valid for the assessment of general intelligence in gifted children (Bracken & McCallum, 1993; Kaplan, 1992; Kaufman, 1992; Spangler & Sabatino, 1995; Sparrow & Gurland, 1998). The main change of the current edition of the WISC to the former edition was the omission of the familiar verbal and performance index scores. Instead there are now four composite scores: Verbal Comprehension, Perceptual Reasoning, Working Memory and Processing Speed. As before one total score can be computed. Since the comparison between the verbal and performance index scores were often used the new structure will require time for users to adjust to and will have to prove its usefulness and practicability (see Burns & O'Leary, 2004). The test has been normed on normal children (2,200 children) as well as on special populations (550 children, e.g. children with Attention-Deficit/ Hyperactivity Disorder, Learning Disabilities or intellectual gifted children). However, the sample sizes of the special groups are quite small, they range from 16 to 89 children per group (see Burns & O'Leary, 2004).

Kaufman Assessment Battery for Children

The Kaufman Assessment Battery for Children, Second Edition (K-ABC II) (Kaufman & Kaufman, 2003) can be used with children from 3 to 18 years. It consists of all in all 20 subtests from which a certain subset is chosen for different age groups and which can be aggregated to five scales: Learning, Memory, Simultaneous Processing, Planning and Knowledge. Before administration the examiner must select the underlying theoretical model for interpretation, either the Luria-Model or the Cattel-Horn-Carroll-Model (CHC) (see Kaufman & Kaufman, 2003). The main difference between the two models is that the Luria-Model focuses on mental processing and excludes acquired knowledge as far as possible, whereas the CHC-Model distinguishes between fluid and crystallized intelligence. Therefore the Luria-Model contains only four scales, the subtest Knowledge (in the sense of crystallized intelligence) is omitted. The test authors suggest the CHC-Model as the model of choice, because they consider knowledge to be an important aspect of cognitive functioning. The K-ABC II was normed with 125 to 250 children for each age group. From a theoretical point of view, however, it seems to be a little queer that in the framework of the K-ABC different theoretical concepts as e.g. Luria´s Sequential Processing which stems from a neuropsychological perspective and Memory (a factor analytic perspective) should be identical. In any case they are operationalized identically.

Tests of nonverbal reasoning abilities

A cultural fair instrument to measure cognitive abilities is the Progressive Matrices Test (Raven, Court, & Raven, 1983, 1986). It measures nonverbal reasoning abilities (inductive reasoning) without relying on verbal, quantitative, or memory aspects. Because of the omission of verbal items the test is attractive for the assessment of children from minority groups or risk groups (e.g., foreign children, children from underprivileged families and so on). The test is sometimes classified as tests for general intelligence in the sense of Spearman (Heller & Perleth, 2000) since it measures solely nonverbal reasoning abilities.

The covered age ranges from 5 to 11 (Colored Progressive Matrices) and 6 to 17 years (Standard Progressive Matrices). For older gifted students and adults the Advanced Progressive Matrices are available.

Peabody Picture Vocabulary Test

The Peabody Picture Vocabulary Test – Third Edition (PPVT-III, Dunn & Dunn, 1997), successor of the PPVT-Revised (PPVT-R, Dunn & Dunn, 1981) or the British Picture Vocabulary Test (Dunn, Dunn, Wetton, & Burley, 1997) is sometimes used as a screening measure to select gifted children - despite it is a pure verbal test (see Hayes & Martin, 1986). Norms for different age groups are available from 2 1/2 years on to adulthood. The reliability and validity is judged as sufficient, but Sattler (1988) advices against the use of the PPVT-R as a "screening device for measuring intellectual level of functioning" (p. 350). The use of the test as a (screening) instrument for the identification of gifted children can be criticized from a methodological point of view (e.g. Hayes & Martin, 1986; Tarnowski & Kelly, 1987) as well as because of its exclusively verbal character.

5.2. Teacher-Checklists

For economic reasons in both diagnostic practice and scientific research a multi-level proceeding is preferred for the identification of gifted children and youth. On the first level a screening takes place with the help of so called check lists. Concerning the application of check lists in school it matters in the first place how appropriate they are, that is how good teachers can identify gifted children in their classroom on the basis of a check list. The recently finished Bavarian Primary School Study addressed among other questions the quality of identification by teacher check lists (Heller, Reimann, & Senfter, 2005).

As expected, primary school teachers were better at assessing crystallized intelligence (sensu Cattell, 1963) than fluid intelligence. Whereas fluid intelligence is supposed to apply to genetic dispositions, crystallized intelligence depends rather on socialisation. Teachers were also better at assessing crystallized intelligence in older students. Contrary to their expectations teachers did not very well at the assessment of creative and social skills; the hit rate was about 20 % (creativity) or below (social skills).

We can conclude that teacher check lists provide valid assessments with a certainty of about 50 % primarily for verbal and mathematical abilities. Difficulties arise concerning the assessment of gifted underachievers, that is students with notable worse performance than their intelligence would indicate. Teachers often fail to recognize these gifted underachievers (see also Rost & Hanses, 1997; Wild, 1991; Ziegler & Stoeger, 2003). One possible explanation might be that teachers (like anybody else in such a situation) orient in the first place by clearly visible performance behaviour and much less by performance potential, which is only indirectly observable. It is also difficult for teachers to differentiate between specific areas of giftedness, instead teacher judgements are rather global (Wild, 1993).

In favour of teacher nomination is to say that the diagnostics is based on extensive observations. However, normally performance rather than potential is assessed and high intelligence is detected poorly. Furthermore teachers underestimate giftedness in students with educational difficulties and overestimate the abilities of motivated students. Based on these problems standardized instruments should be deployed as well (see above; for a comprehensive review see also Perleth, 2010).

5.3. Nomination by parents and peers

Research shows that nomination by parents is as difficult as nomination by teachers and assessment errors are even worse. Another point is that information about the development of the child (language, motor or social development) is usually collected retrospectively, but retrospective data are not very reliable. A study comparing parents ratings with results of psychological (test) diagnostics using data from 3 counseling centers in Germany and Austria (N>300) showed that the information from of parents' checklists have a very limited validity and can by no means substitute professional psychological diagnostics (Perleth, 2010). Most interestingly Perleth and his colleagues found higher correlations between parents' ratings and crystallized intelligence, while the coefficients between parents' judgements and fluid intelligence were considerably lower.

Peer nomination (diagnostics based on the assessment of classmates) also could not qualify for the identification of gifted. This procedure was suggested in the first place to identify creativity and social skills. With regard to intelligence could be shown that assessment is based on previous grades and that it is strongly consistent for different areas of intelligence. In addition younger children's assessments proved to be highly unreliable.

6. Conclusions

In the context of a re-analysis of the data of the Munich High Ability Study (Perleth, 2001; Perleth & Sierwald, 2000) ability ratings from different sources (test, questionnaires, and teacher nomination) were compared. Based on these comparisons we can conclude:

- Giftedness is not unidimensional. There is no such thing as „The gifted“ but rather different areas of giftedness must be differentiated.
- Depending on the source of information and the diagnostic measure different groups of people will be identified as gifted within a certain talent area.
- It could be supported that teacher assessments are rather global and that teachers are more likely to identify students with high academic achievement as gifted.

Therefore we can conclude that there is no perfect identification method. Depending on the aim of the identification process different diagnostic measures can be the best.

Recommended literature for further reading

Perleth, Ch. & Hoese, D. (in press). Proyecto Odysseus en el Departamento de Psicología de la Educación "Rosa y David Katz" de la Universidad de Rostock y su consulta psicológica para superdotados. In Sastre y Riba, S. (ed.). *Naturaleza e intervención en altas capacidades*. Barcelona: Editorial Viguera.

Most important literature used for the present text

Heller, K.A. & Perleth, Ch. (2007a). *Münchner Hochbegabungstestbatterie für die Primarstufe (MHBT-P)* [Munich High Ability Testbattery, Primary Level]. Göttingen: Hogrefe.

Heller, K.A. & Perleth, Ch. (2007b). *Münchner Hochbegabungstestbatterie für die Sekundarstufe (MHBT-S)* [Munich High Ability Testbattery, Secondary Level]. Göttingen: Hogrefe.

Perleth, C. (2010). Checklisten in der Hochbegabungsdiagnostik [Checklists in the diagnosis of giftedness]]. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung* (S. 65-87). Göttingen: Hogrefe.

Perleth, C., & Heller, K. A. (2010). Adapting conceptual models for Cross-Cultural applications. *Faisca. Revista de altas capacidades, 14*(16), 76-95.

Perleth, C., & Sierwald, W. (2000). Testtheoretische Konzepte der Begabungsdiagnostik [Test theoretical concepts of the diagnostic of intellectual capacity]. In K. A. Heller

(Hrsg.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (S. 41-95). Bern: Huber.

Complete list of references

Becker, K. A. (2003). *History of the Stanford-Binet intelligence scales: Content and psychometrics.* Itasca, IL: Riverside Publishing.

Binet, A. (1905). A propos de la mésure de l'intelligence. *Année Psychologique, 11*, 69-82.

Borland, J. H., & Wright, L. (2000). Identifying and Education Poor and Under-Represented Gifted Students. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 587-594). Oxford: Pergamon.

Bracken, B. A., & McCallum, R. S. (Eds.). (1993). *Wechsler Intelligence Scale for Children: Third edition*. Brandon, VT: Clinical Psychology Publishing.

Brown, S. W., Hwang, M. T., Baron, M., & Yakimowski, M. E. (1991). Factor analysis of responses to the WISC-R for gifted children. *Psychological Reports, 69*, 99-107.

Burns, T. G., & O'Leary, S. D. (2004). Wechsler Intelligence Scale for Children-IV: Test Review. *Applied Neuropsychology, 11*(4), 233-236.

Butler-Por, N. (1993). Underachieving gifted students. In K. A. Heller, F. J. Mönks & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 649-668). Oxford, UK: Pergamon.

Campbell, J. R., Wagner, H., & Walberg, H. J. (2000). Academic Competitions and Programs Designed to Challenge the Exceptionally Talented. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 523-535). Oxford: Pergamon.

Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Educational Psychology, 54*, 1-22.

Dunn, L. M., & Dunn, L. M. (1981). *Peabody Picture Vocabulary Test - Revised*. Circle Pines, MN: American Guidance Service.

Dunn, L. M., & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test - Third Edition (PPVT-III)*. Bloomington, MN: Pearson Assessments.

Dunn, L. M., Dunn, L. M., Wetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale (BPVS II)*. Circle Pines, MN: American Guidance Service.

Feldhusen, J.F. & Jarwan, F.A. (2000). Identification of Gifted and Talented Youth for Educational Programs. In K.A. Heller, F.J. Mönks, R.J. Sternberg & R.F. Subotnik (Eds.), *International Handbook of Giftedness and Talent* (2nd ed., pp. 271-282). Oxford: Pergamon.

Freeman, J. (2000). Families: The Essential Context for Gifts and Talents. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 573-585). Oxford: Pergamon.

Gagné, F. (2000). Understanding the Complex Choreography of Talent Development Through DMGT-Based Analysis. In K.A. Heller, F.J. Mönks, R.J. Sternberg & R.F. Subotnik (Eds.), *International Handbook of Giftedness and Talent* (2nd ed., pp. 67-79). Oxford: Pergamon.

Gardner, H. (1983). *Frames of Mind. The Theory of Multiple Intelligences*. New York: Basic Books.

Hany, E.A. (1993). Methodological Problems and Issues Concerning Identification. In K.A. Heller, F.J. Mönks & A.H. Passow (Eds.), *International Handbook of Research and Development of Giftedness and Talent* (pp. 209-232). Oxford: Pergamon.

Hany, E.A. (1994). The Development of Basic Cognitive Components of Technical Creativity: A Longitudinal Comparison of Children and Youth with High and Average Intelligence. In R.F. Subotnik & K.D. Arnold (Eds.), *Beyond Terman: Contemporary Longitudinal Studies of Giftedness andTalent* (pp. 115-154). Norwood, NJ: Ablex.

Hany, E.A. (2001). Identifikation von Hochbegabten im Schulalter. In K. A. Heller (Ed.), *Hochbegabung im Kindes- und Jugendalter* (2 ed., pp. 41-169). Göttingen: Hogrefe.

Hayes, F. B., & Martin, R. P. (1986). Effectiveness of the PPVT-R in the screening of young gifted children. *Journal of Psychoeducational Assessment, 4*, 27-33.

Heller, K.A. (1989). Perspectives on the Diagnosis of Giftedness. *The German Journal of Psychology, 13,* 140-159.

Heller, K.A. (1990). Goals, Methods and First Results from the Munich Longitudinal Study of Giftedness in West Germany. In C.W. Taylor (Ed.), *Expanding Awareness of Creative Potentials Worldwide* (pp. 538-543). New York: Trillium Press.

Heller, K.A. (1991). The Nature and Development of Giftedness: A Longitudinal Study. *European Journal for High Ability, 2,* 174-188. – Reprint in A.J. Cropley & D. Dehn (Eds.), *Fostering the Growth of High Ability: European Perspectives* (pp. 31-56). Norwood, NJ: Ablex (1996).

Heller, K.A. (1995a). Ability and Creativity: Their Role in Science and Technology. In S. Cho, J.H. Moon & J.O. Park (Eds.), *Creativity for the 21st Century* (pp. 39-73). Seoul: The Korean Society for the Gifted.

Heller, K.A. (1995b). Evaluation of Programs for the Gifted. In M. Katzko & F.J. Mönks (Eds.), *Nurturing Talent: Individual Needs and Social Ability* (pp. 264-268). Assen/Maastricht: Van Gorcum.

Heller, K.A. (1999). Individual (Learning and Motivational) Needs versus Instructional Conditions of Gifted Education. *High Ability Studies, 9*, 9-21.

Heller, K. A. (2000). Hochbegabungsdiagnose. In K. A. Heller (Ed.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (pp. 241-258). Bern: Huber.

Heller, K. A. (Ed.). (2001). *Hochbegabung im Kindes- und Jugendalter* [High Ability in Childhood and Adolescence] (2 ed.). Göttingen: Hogrefe.

Heller, K. A. (Ed.). (2002a). *Begabtenförderung im Gymnasium* [Gifted Education at the German gymnasium]*. Ergebnisse einer zehnjährigen Längsschnittstudie*. Opladen: Leske + Budrich.

Heller, K.A. (Guest-Ed.). (2002b). Program Evaluation. *European Journal of Psychological Assessment, 18,* Issue 3 (Special Section), 187-241.

Heller, K.A. (2004). Identification of Gifted and Talented Students. *Psychology Science, 46*, 302-323.

Heller, K.A. (2005). The Munich Model of Giftedness and Its Impact on Identification and Programming. *Gifted and Talented International, 20*, 30-36.

Heller, K.A. & Hany, E.A. (1986). Identification, Development and Achievement Analysis of Talented and Gifted Children in West Germany. In K.A. Heller & J.F. Feldhusen (Eds.), *Identifying and Nurturing the Gifted* (pp. 67-82). Toronto: Huber Publ.

Heller, K.A. & Perleth, Ch. (1989). Forms of Giftedness in Children and Adolescents: Current Results of a Longitudinal Study. In M.A. Luszcz & T. Nettelbeck (Eds.), *Psychological Development: Perspectives Across the Life-Span* (pp. 241-249). Amsterdam: Elsevier Science.

Heller, K. A., & Perleth, C. (2000). Informationsquellen und Meßinstrumente [Information sources and measurement methods]. In K. A. Heller (Ed.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (pp. 96-216). Bern: Huber.

Heller, K.A. & Perleth, Ch. (2004). Adapting Conceptual Models for Cross-Cultural Applications. In J.R. Campbell, K. Tirri, P. Ruohotie & H. Walberg (Eds.), *Cross-Cultural Research: Basic Issues, Dilemmas, and Strategies* (pp. 81-101). Hämeenlinna, Finland: Research Centre for Vocational Education / University of Tampere, Finland.

Heller, K.A. & Perleth, Ch. (2007a). *Münchner Hochbegabungstestbatterie für die Primarstufe (MHBT-P)* [Munich High Ability Testbattery, Primary Level]. Göttingen: Hogrefe.

Heller, K.A. & Perleth, Ch. (2007b). *Münchner Hochbegabungstestbatterie für die Sekundarstufe (MHBT-S)* [Munich High Ability Testbattery, Secondary Level]. Göttingen: Hogrefe.

Heller, K. A., & Ziegler, A. (1996). Gender Differences in Mathematics and the Natural Sciences: Can Attributional Retraining Improve the Performance of Gifted Females? *Gifted Child Quarterly, 40*, 200-210.

Heller, K. A., & Ziegler, A. (2001). Attributional Retraining: A Classroom-Integrated Model for Nurturing Talents in Mathematics and the Sciences. In N. Colangelo & S. G. Assouline (Eds.), *Talent Development* (Vol. IV, pp. 205-217). Scottsdale, AZ: Great Potential (Gifted Psychology) Press.

Heller, K. A., Reimann, R., & Senfter, A. (2005). *Hochbegabung im Grundschulalter: Erkennen und Fördern*. Münster: LIT.

Heller, K.A., Perleth, Ch. & Lim, T.K. (2005). The Munich Model of Giftedness Designed to Identify and Promote Gifted Students. In R.J. Sternberg & J.E. Davidson (Eds.), *Conceptions of Giftedness* (2nd ed., pp. 147-170). New York: Cambridge University Press.

Heller, K.A., Perleth, Ch. & Sierwald, W. (1990). The Munich Longitudinal Study of Giftedness. In S. Bailey, E. Braggett & M. Robinson (Eds.), *The Challenge of Excellence* (pp. 37-57). Melbourne: Barker.

Heller, K.A., Senfter, A. & Linke, S. (2006). *Zweiter Zwischenbericht zur Evaluation des MINT-Projekts Nordbaden* [Second Interim Report on the Evaluation of the Gifted Program „Hector-Seminar“ (MINT-Project) in Nordbaden (Germany)]. München: Center for the Study of Giftedness at LMU.

Heymans, P. & Mönks, F.J. (2004). Identifying Promising Individuals: A Commentary. *Psychology Science, 46*, 398-403.

Kaplan, C. (1992). Ceiling effects in assessing high-IQ children with the WPPSI-R. *Journal of Clinical Child Psychology, 21*, 403-406.

Kaufman, A. S. (1992). Evaluation of the WISC-III and WPPSI-R for gifted children. *Roeper Review, 14*, 154-158.

Kaufman, A. S., & Kaufman, N. L. (2003). *K-ABC II: Kaufman Assessment Battery for Children, Second Edition*. Circle Pines, MN: American Guidance Service.

Kaufmann, F. A., & Castellanos, F. X. (2000). Attention Deficit / Hyperactivity Disorder in Gifted Students. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 621-632). Oxford: Pergamon.

Kerr, B. (2000). Guiding Gifted Girls and Young Women. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 649-657). Oxford: Pergamon.

Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Rechek, A. (2006). Tracking exceptional human capital over two decades. *Psychological Science, 17*, 194-199.

Lubinski, D., Webb, R. M., Morelock, M. J., & Benbow, C. P. (2001). Top 1 in 10,000: A 10-year follow-up of the profoundly gifted. *Journal of Applied Psychology, 86*, 718-729.

Masten, W. G., Morse, D. T., & Wenglar, K. E. (1995). Factor structure of the WISC-R for Mexican-American students referred for intellectually gifted assessment. *Roeper Review, 18*, 130-131.

Möbus, C. l. (1983). Die praktische Bedeutung der Testfairness als zusätzliches Kriterium zu Reliabilität und Validität. *Tests und Trends 3. Jahrbuch der Paedagogischen Diagnostik*, 155-203.

Mönks, F.J. (1985). Hoogbegaafden: een situatieschets. In F.J. Mönks & P. Span (Eds.), *Hoogbegaafden in de samenleving* (pp. 17-31). Nijmegen: Dekker & van de Vegt.

Perleth, C. (2001). Follow-up-Untersuchungen zur Münchner Hochbegabungsstudie. In K. A. Heller (Ed.), *Hochbegabung im Kindes- und Jugendalter* (pp. 357-446). Göttingen: Hogrefe.

Perleth, C. (2010). Checklisten in der Hochbegabungsdiagnostik [Checklists in the diagnosis of giftedness]]. In F. Preckel, W. Schneider & H. Holling (Hrsg.), *Diagnostik von Hochbegabung* (S. 65-87). Göttingen: Hogrefe.

Perleth, Ch. & Heller, K.A. (1994). The Munich Longitudinal Study of Giftedness. In R.F. Subotnik & K.D. Arnold (Eds.), *Beyond Terman: Contemporary Longitudinal Studies of Giftedness and Talent* (pp. 77-114). Norwood, NJ: Ablex, 2nd ed. 1998.

Perleth, C., & Sierwald, W. (2000). Testtheoretische Konzepte der Begabungsdiagnostik [Test theoretical concepts of the diagnostic of intellectual capacity]. In K. A. Heller (Hrsg.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (S. 41-95). Bern: Huber.

Perleth, Ch. & Sierwald, W. (2001). Entwicklungs- und Leistungsanalysen zur Hochbegabung [Analyses of development and achievement of the gifted].. In K. A. Heller (Hrsg.), *Hochbegabung im Kindes- und Jugendalter* (2nd ed., S. 171-355). Göttingen: Hogrefe.

Perleth, C., Schatz, T., & Mönks, F. J. (2001). Early indicators of high ability. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook for giftedness and talent* (pp. 297-316). Oxford: Pergamon.

Perleth, Ch., Sierwald, W. & Heller, K.A. (1993). Selected Results of the Munich Longitudinal Study of Giftedness: The Multidimensional/Typological Giftedness Model. *Roeper Review, 15,* 149-155.

Peters, W. A. M., Grager-Loidl, H., & Supplee, P. (2000). Underachievement in Gifted Children and Adolescents: Theory and Practice. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 609-620). Oxford: Pergamon.

Pfeiffer, St.I. & Jarosewich, T. (2007). The Gifted Rating Scales-School Form: An Analysis of the Standardization Sample Based on Age, Gender, Race, and Diagnostic Efficiency. *Gifted Child Quarterly, 51*, 39-50.

Raven, J. C., Court, J. H., & Raven, J. (1983). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 3) - Standard Progressive Matrices (1983 edition)*. London: Lewis.

Raven, J. C., Court, J. H., & Raven, J. (1986). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 2) - Coloured Progressive Matrices (1986 edition)*. London: Lewis.

Renzulli, J.S. (1978). What makes Giftedness? Reexamining a Definition. *Phi Delta Kappan, 60*, 180-184, 261.

Renzulli, J. S., & Reis, S. M. (2000). The Schoolwide Enrichment Model. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.), *International handbook of giftedness and talent* (2 ed., pp. 367-382). Oxford: Pergamon.

Roid, G. H. (2003). *Stanford-Binet Intelligence Scales: Fifth Edition. Technical Manual*. Itasca, IL: Riverside Publishing.

Rost, D. H. (1993). *Lebensumweltanalyse hochbegabter Kinder [Analysis of living circumstances of gifted children]*. Göttingen: Hogrefe.

Rost, D. H., & Hanses, P. (1997). Wer nichts leistet, ist nicht begabt? Zur Identifikation hochbegabter Underachiever durch Lehrkraefte. / Not achieving--not gifted? About the identification of gifted underachievers by teacher ratings. *Zeitschrift fuer Entwicklungspsychologie und Paedagogische Psychologie, 29*, (2) 167-177.

Sattler, J. M. (1988). Assessment of children.

Spangler, R. S., & Sabatino, D. A. (1995). Temporal stability of gifted children's intelligence. *Roeper Review, 17*, 207-210.

Sparrow, S. S., & Gurland, S. T. (1998). Assessment of gifted children with the WISC-III. In A. Prifitera & D. H. e. a. Saklofske (Eds.), *WISC-III clinical use and interpretation: Scientist-practitioner perspectives* (pp. 59-72). San Diego, CA: Academic Press.

Stapf, A. (2003). *Hochbegabte Kinder*. München: C.H. Beck.

Stumpf, H. e. (1996). Klassische Testtheorie. *Handbuch Quantitative Methoden*, 411-430.

Tarnowski, K. J., & Kelly, P. A. (1987). Utility of the PPVT-R for pediatric intellectual screening. *Journal of Pediatric Psychology, 12*, 611-614.

VanTassel-Baska, J., Feng, A.X. & Evans, B.L. (2007). Patterns of Identification and Performance Among Gifted Students Identified Through Performance Tasks: A Three-Year Analysis. *Gifted Child Quarterly, 51*, 218-231.

Wai, J., Lubinski, D., & Benbow, C. P. (2005). Creativity and occupational accomplishments among intellectually precocious youth: An age 13 to age 33 longitudinal study. *Journal of Educational Psychology, 97*, 484-492.

Webb, J. T., Meckstroth, E. A., & Tolan, S. S. (2002). *Hochbegabte Kinder - ihre Eltern, ihre Lehrer* (3 ed.). Bern: Huber.

Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Baltimore, MD: Williams & Wilkins.

Wechsler, D. (2002). *Wechsler Preschool and Primary Scale of Intelligence-III*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2003). *Wechsler Intelligence scale for Children - fourth Edition (WISC-IV)*. San Antonio, TX: Psychological Corporation.

Wild, K. P. (1991). *Identifikation hochbegabter Schüler. Lehrer und Schüler als Datenquellen*. Heidelberg: Asanger.

Wild, K. P. (1993). Hochbegabtendiagnostik durch Lehrer. In D. H. Rost (Ed.), *Lebensumweltanalyse hochbegabter Kinder* (pp. 236-261). Goettingen: Hogrefe.

Wilkinson, S. C. (1993). WISC-R profiles of children with superior intellectual ability. *Gifted Child Quarterly, 37*, 84-91.

Yewchuck, C., & Lupart, J. L. (1993). Gifted handicapped: A desultory duality. In K. A. Heller, F. J. Mönks & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 709-725). Oxford, UK: Pergamon.

Ziegler, A., & Stoeger, H. (2003). Identification of Underachievement: An Empirical Study on the Agreement Among Various Diagnostic Sources. *Gifted and Talented International, 18*(87-94).

Ziegler, A., Dresel, M., & Schober, M. (2000). Underachievementdiagnose - Ein Modell zur Diagnose partieller Lernbeeinträchtigungen. In K. A. Heller (Ed.), *Begabungsdiagnostik in der Schul- und Erziehungsberatung* (2 ed., pp. 259-278). Bern: Huber.

Aplicaciones y problemas, soluciones y problemas

Vease las explicaciones sobre buscar talentos en el texto.

Vease tambien los ejemplos de la consulta psicológica para superdotatos en Perleth & Hoese (in press):

Perleth, Ch. & Hoese, D. (in press). Proyecto Odysseus en el Departamento de Psicología de la Educación "Rosa y David Katz" de la Universidad de Rostock y su consulta psicológica para superdotados. In Sastre y Riba, S. (ed.). *Naturaleza e intervención en altas capacidades*. Barcelona: Editorial Viguera.

Glosario de términos más relevantes (all terms are explained in the text):

Diagnosis of Giftedness,
Identification of Giftedness,
Munich High Ability Test Battery (MHBT),
Munich Model of Giftedness (MMG),
Psychological Testing,
Psychometric Quality,
Objectivity,
Reliability,
Talent Search,
Test Norms,
Validity

Preguntas para la Autoevaluación

Cuáles de los siguientes comentarios de profesores están dentro del marco de comprensión para la superdotación intelectual, como el tal se usa dentro del MHBT?

☐ „Martin sacó en el último exámen de matemáticas la nota 3, por eso no creo que sea superdotado intelectual."

☐X „Parece que Ana sea superdotada intelectual en el campo verbal, aunque no entienda ni leches de las matemáticas."

☐ „Favor de leerles cuentos a los niños, porque altas capacidades verbales son la base para cada tipo de superdotación intelectual.

En cuanto a la interpretación del exámen de inteligencia: ¿Cuáles son las frases correctas? (Atención: ¡Podrían ser correctas varias!)

☐ Se describe la T-cuantía de 56 (aprox. un fáctor de inteligencia de 110) como rendimiento inferior al remedio?

☐X La T-cuantia de 39 refleja un rendimiento inferior al remedio.

☐ La T-cuantia de 35 refleja un rendimiento regular.

☐X La T-cuantia de 50 es igual a un fáctor de inteligencia de 100.

☐X Los que tienen una T-cuantía arriba de 60 (o un factor de inteligencia arriba de 115), pertenecen al 16 por ciento de los mejores.

☐X Un rendimiento del promedio tendría la T-cuantía de 56 (aprox. un factor de inteligencia 110).

¿Que son las características de pruebas psicológicas que encajan con los criterios de calidad del diagnóstico psicológico? (Atención: ¡Podrían ser correctas varias!)

☐X Para la percepción de la objetividad de la realización, el encargado del exámen tiene que contarle a la persona el contenido de las instucciones de la prueba.

☐ Alta reliabilidad tiene un exámen cuando mide exáctamente la característica psicológica que debería de medir.

☐X Pruebas psicológicas se diseñan para el uso de psicólogos ejeciendo su labor en la práctica.

☐X Es muy importante fijarse el mantenimiento de la objetividad durante el tiempo del exámen.

☐X Nada más una prueba reliable también puede tener alta validez.

☐ A traves de exámenes psicológicos se puede medir el „verdadero valor" del individuo o de una característica psicológica.

Conécte las siguientes frases con su norma relacionada. ¿Cuál es lo visible de cada ocasión?

	Norma relacionada		
	Social (*Social reference norm)*	Individual (*Individual reference norm)*	Referente a la característica (*Task-oriented reference norm)*

„Por diós, otra vez subí de peso.“	☐	**☐X**	☐
„¡Mi indíce de masa corporal ya alcanzó el 28!“	☐	☐	**☐X**
„Pues, para mi edad aún pertenece al promedio.“	**☐X**	☐	☐

¿Según Cattell, qué tipo de „inteligencias“ se puede medir através el KFT-HB? (Atención: ¡Podrían ser correctas varias!)

☐X Aspéctos de la inteligencia „líquida“.
☐ Conocimiento físico como parte de la inteligencia „cristalina“.
☐X Habilidades verbales como parte de la inteligencia „cristalina“.
☐X Capacidades mentales matemáticas/cuantitativas como parte de la inteligencia „cristalina“
☐ Velocidad del tratamiento de la información como aspécto de la inteligencia „líquida“.
☐ Rendimiento de la memoria como aspécto de la inteligencia „líquida“.