

On a Practitioner's Need of Further Development of Wechsler Scales. Adaptive Intelligence Diagnosticum (AID 2)

Klaus D. Kubinger
University of Vienna, Austria

Wechsler's intelligence test-batteries are still popular yet suffer from psychometric shortcomings and lack a certain content improvement and enlargement. In this paper a new approach will be presented that suits traditional Wechsler-testing. The approach in question is the Adaptive Intelligence Diagnosticum (AID 2; Kubinger & Wurst, 2000). Due to its "branched testing" design, AID 2 works out to be more economical and in this regard also offers other advantages, such as parallel tests and short forms. AID 2 offers a method of survey for identifying specific developmental disorders or learning disabilities. It includes an optional non-verbal instruction and a schedule for retrograde observation support of behavioral misfits. It also offers discriminative indicators for intellectual neglect vs. intellectual advancement.

Keywords: Adaptive Intelligence Diagnosticum, intelligence test batteries

Las baterías de tests de inteligencia de Wechsler todavía son populares a pesar de presentar limitaciones psicométricas y una cierta carencia de mejoras y ampliación de contenidos. En este trabajo se presenta un nuevo enfoque que cubre las necesidades de las tradicionales pruebas de Wechsler. El enfoque en cuestión es el Diagnóstico de Inteligencia Adaptativa (AID 2; Kubinger & Wurst, 2000). Debido a su diseño ramificado, AID 2 es más económico y ofrece otras ventajas, como pruebas paralelas y formas abreviadas. AID 2 ofrece un método de investigación para identificar trastornos específicos del desarrollo y dificultades de aprendizaje. Incluye una introducción no verbal opcional y un programa de apoyo para la observación retrospectiva de desajustes conductuales. También ofrece indicadores discriminativos de retraso intelectual vs. progreso intelectual.

Palabras clave: Diagnóstico de Inteligencia Adaptativa, baterías de tests de inteligencia

There is no doubt that Wechsler's intelligence test-batteries are of significant importance for psychological consulting. They have established themselves as state of the art test-batteries, particularly for children and juveniles up to the age of 16, and even for the purposes of lawsuits and any administration of justice. Within clinical research work their application is obligatory due to the scientific demands of comparability of research results.

Nevertheless, there are certain psychometric shortcomings and economic deficiencies that call for some kind of reformation. Apart from this, there are also various matters of content that appeal for a certain improvement and enlargement. That is to say, alternative approaches should also get a fair chance. This paper deals with such an approach.

The psychometric shortcomings of Wechsler test-batteries have already been proven elsewhere (cf. Kubinger, 1998). A highly economical approach, particularly with respect to the administration duration and the error of estimation of aimed-for ability parameters using adaptive testing, has also been presented in another paper (cf. Kubinger, 2004, submitted). In the following, some essential content developments will be discussed.

The given approach deals with a test-battery that has already been published: The test battery in question is, however, German-made since no other test battery exists as yet with regard to a such highly economical approach. The scientific advances made in this test-battery may serve either for a further development of the Wechsler scales, or this German-made test-battery may even be adapted for populations other than German speaking ones and be license distributed - one should note that a US-English item translation and adaptation, a British-English version and furthermore an Italian, a Hungarian, and a Turkish version are also available at request from the authors.

Adaptive Intelligence Diagnosticum

Above all, the Adaptive Intelligence Diagnosticum (AID 2; Kubinger & Wurst, 2000), published for the first time in 1985, is based on the Rasch model and therefore fulfills pertinent psychometric presuppositions in order to guarantee a fair scoring. For this, the items of every subtest have been conceptualized according to appropriate copy-right operational definitions and rules of item generation.

Besides this, item presentation was organized according to a copy-right scheduled branched testing design, according to adaptive testing theory. For instance, for five of the 14 subtests in AID 2, test administration occurs as follows: Test administration starts firstly with an age-appropriate subset of five items, then a second and a third subset of five items is administered, both of the latter depend on the preceding score of the testee. If the testee solves one item at the most, a subset one level below is recommended. If at least four

items are solved then a subset one level higher is advised, and only if two or three of the five items are solved is the subset of an equivalent level to be administered. Such a design guarantees that every subject is tested with almost optimal items. In the long run, it is not likely that the items administered will be too easy or too difficult, it is rather the case that items of an appropriate level are applied. In other words, the standard error of estimation comes close to the ideal minimum. One can be sure that this kind of administration does not require any particular effort from the psychologist as he/she is directed by a programmed instruction.

Furthermore, there are essential content developments which deal with:

- 1) (as already indicated) operational definitions for every subtest in order to unequivocally establish the measured aspect of intelligence.
- 2) a certain conceptualization of items so as to avoid interfering abilities or improper moderator variables.
- 3) a non-verbal instruction for the purpose of testing subjects with auditive deficiencies, those whose native tongue is a foreign language and those with social handicaps.
- 4) a pool of optional subtests which serve the purpose of diagnosing certain cognitive disorders.
- 5) a survey for identifying specific developmental disorders or learning disorders.
- 6) discriminative indicators for intellectual neglect vs. intellectual advancement.
- 7) the abandonment of the concept of the IQ and the general factor model.
- 8) a schedule which serves as a retrograde observation support for behavioral misfits.

Operational Definitions According to a Pragmatic Theory of Intelligence

Due to a more pragmatic approach towards testing intelligence, AID 2 was never based on any pertinent theory of intelligence: "What will intelligence tests be like in the year 2000? Instead of an IQ they will yield scores that reflect separately the various aspects of ability that are of interest, will express those abilities in a more manageable metric, and will report them in terms of greater educational utility." (Turnbull, 1979, p. 277). And even if any theoretical approach were then of relevance it would be the almost unknown, but Sternberg-like "information processing model of intelligence" by Roth, Oswald, and Daumenlang (1980) that claims four factors, those being speed of information processing, capacity of information processing, ability of reducing information, and ability to generate supra signs.

Hence, the pragmatic approach which leads to a kind of expert rating. Numerous psychologists engaged in the field of consulting were asked to comment on what kinds of cognitive performances they considered as necessary in order to deal with the demands of today's world. "Intelligence"

is thought of as the bundle of all cognitive requirements that are necessary in order to acquire knowledge and to develop the ability to perform – whereby “cognition” refers to each process, through which a human acquires knowledge of an object or through which it becomes aware of its environment: perception, recognition, imagination, judging, memory, learning, thinking and language. The result was a battery of eleven subtests whose aimed-for measured abilities are as follows. – Bear in mind that there was no doubt about the significant diagnostic relevance and practical importance of stimulating a testee’s interaction with the material as well as with the psychologist himself/herself:

- *Everyday Knowledge* refers to the ability to inform oneself about things that are commonly present nowadays – for example “What is the purpose of the blood circulation in the human body?”; the correct answers would be: distribution of oxygen, distribution of nutrients, removal of carbon dioxide, removal of waste products.

- *Competence in Realism* means the ability to understand the functional properties of everyday objects – given pictures with missing details which are essential for the object to work in an appropriate manner, the testee is asked to point out the psychologist of what is missing in the picture (cf. Figure 1a).

- *Applied Computing* deals with the ability to solve everyday numerical problems, primarily by reasoning, however not by using any high-level school techniques – for example “Phoebe runs across a meadow which is 10 yards long, once to its end and then back again. Charly runs twice as far. How many yards did Charly run?”

- *Social and Material Sequencing* means the ability to understand the sequence of everyday actions – given several picture cards of a story the testee has to arrange them in a sequence in which he/she thinks the events shown in the pictures would happen (cf. Figure 1b).

- *Immediately Reproducing* (numerical stimuli) demonstrates the capacity of the short term memory, this being with respect to consecutively presented information – for example “6-9-5-7-4”; missing or interchanged digits destroy any score profit.

- *Producing Synonyms* refers to the ability to understand the concept of a word – for example “agreement”; correct answers would be “arrangement”, “understanding”.

- *Coding and Associating* demonstrate speed of information processing and incidental learning ability – the testee has the task of coding simple figures into simple signs by hand according to a pattern sheet. He/she is then asked to code the same figures however this time without the pattern sheet at his/her disposal (cf. Figure 1c).

- *Anticipating and Combining* (the parts of a cut up “Gestalt”) establish figural reasoning abilities with regard to everyday objects – given pieces of a “gestalt”, including an anchor piece, the testee then has to put them together to form some kind of figure – he/she does not necessarily know from the very beginning what the pieces will become (cf. Figure 1d).

- *Abstracting* (the common functionalisms of things) refers to the ability of conceptualization – for example “What is common in ‘fish’ and ‘ship’?”; the correct answer is “they both move in water”.

- *Analyzing and Synthesizing* (abstract figures) refers to the ability to structurally decompose and rebuild a “Gestalt” – the testee has to form different patterns using a number of cubes, similar to the well-known test from Kohs, only the sides of which are colored in a more complicated way (cf. Figure 1e).

- *Social Understanding and Material Reflection* incorporates the comprehension of pertinent relationships and coherences – for example “Why shouldn’t you enter a construction site?”; the correct answer is “danger of injury”.

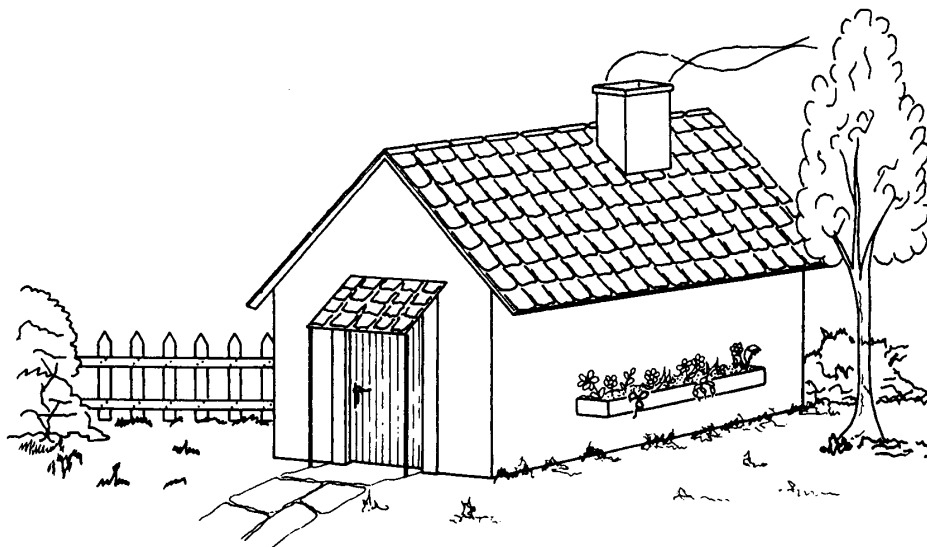


Figure 1a. Example of AID 2 item: Competence in Realism.

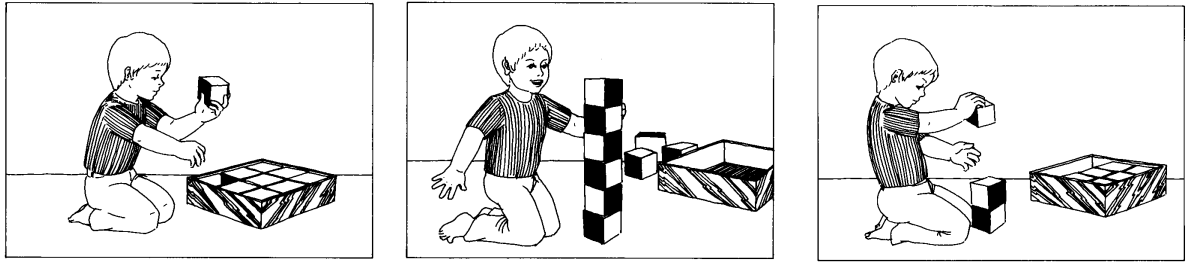


Figure 1b. Example of AID 2 item: Social and Material Sequencing.

Pattern sheet

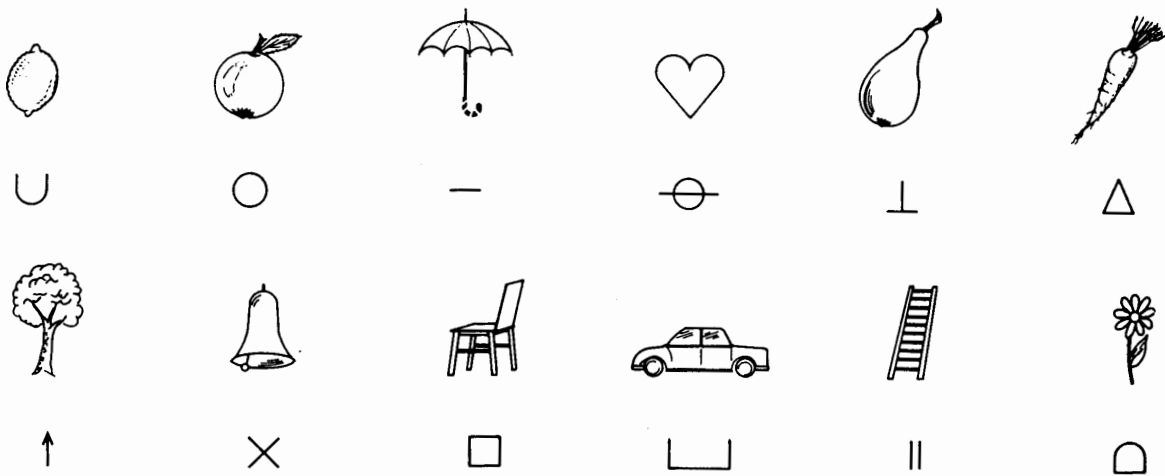


Figure 1c. Example of AID 2 item: Coding and Associating.

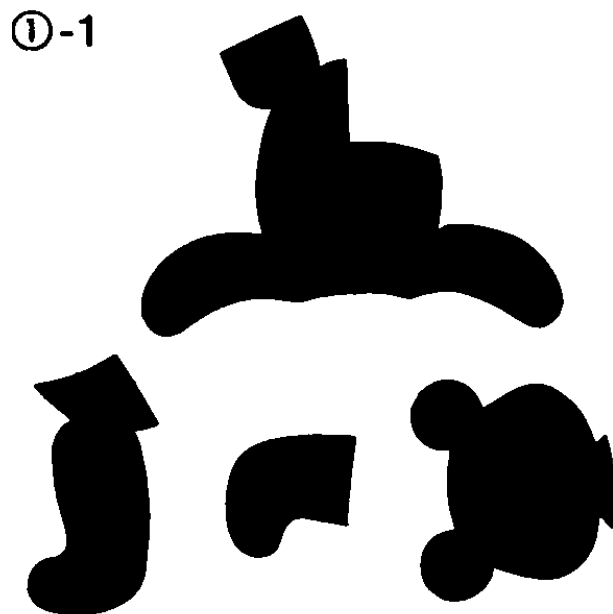


Figure 1d. Example of AID 2 item: Anticipating and Combining.

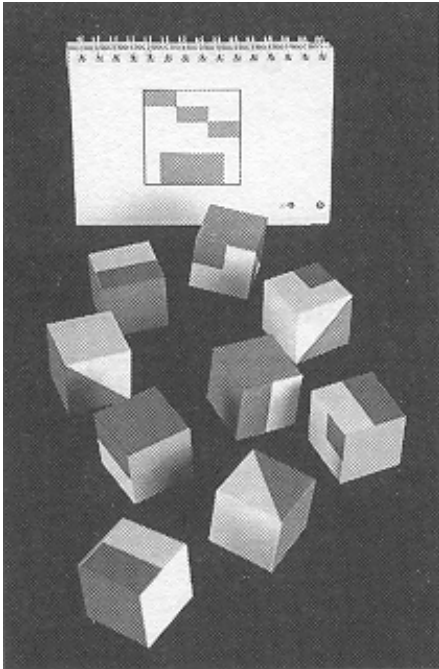


Figure 1e. Example of AID 2 item: Analyzing and Synthesizing.

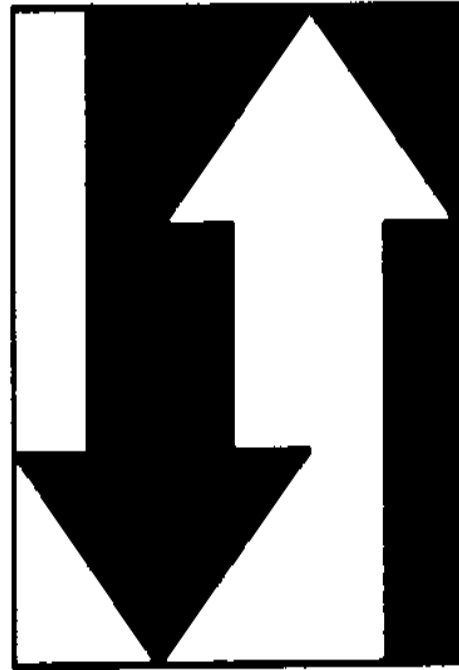


Figure 1f. Example of AID 2 item: Recognition of (figural) Structures.

Later on three further subtests were conceptualized, these being merely optional subtests for further diagnoses of cognitive disorders. The aimed-for measured abilities of these subtests are as follows:

- *Immediately Reproducing* (figural stimuli) demonstrates the capacity of short term memory, this being with respect to consecutively presented information – for example, four certain pictures from a 7 times 7 table that are pointed at; missing pictures or pictures that are interchanged destroy any score profit!

- *Learning and Remembering* demonstrates the capacity of memory after stimuli are presented twice in a row – nine senseless syllables are presented twice acoustically however always in a different sequence, so that no association of consecutive presentation may really help although the repetition in itself should help.

- *Recognition of (figural) Structures* refers to the ability to decompose abstract figures into given units – using patterns according to the cubes described above the testee now has to find the respective intersecting lines of the patterns built from the cubes, the testee however does not need to handle the cubes (cf. Figure 1f).

The Conceptualization of Items so as to Avoid Interfering Abilities

As a matter of fact, some of the subtests bear a likeness to various subtests from other intelligence test-batteries. However, there is evidence that similar looking

subtests do not measure unidimensionally or that they suffer from improper moderator variables. For instance, Gittler (1986) found in his re-analyses of a large sample data set of the German WISC a tempting explanation for the lack of Rasch model fit (cf. Kubinger, 1998) observed in the subtests Information and Picture Completion – the subtest “Information” from WISC corresponds to some extent to *Everyday Knowledge* and “Picture Completion” corresponds to *Competence in Realism*. For the first subtest, he intuitively defined two subgroups of items, i.e. „factual information or book learning“ and „Information based on everyday experience. By analyzing them separately according to the Rasch model he succeeded in producing a fit - after slight corrections of item allotment. For the second subtest, he distinguished between items whose missing part may be found by an analogy, on the one hand, and items which require specific experiences for identifying the missing part, on the other hand. The latter succeeded in fitting the model after the deletion of just a few items.

Consequently, the AID 2 subtest *Everyday Knowledge* focuses on information based on everyday experience, only. And indeed, this subtest from AID 2 stood the test so as to conform with the Rasch model and therefore can be said to guarantee uni-dimensional, fair scores. The AID 2 subtest *Competence in Realism* has been conceptualized so as to require specific experiences for identifying the missing part: There is always a detail missing which is essential for the object to work in an appropriate manner. Again, this subtest from AID 2 fits the Rasch model.

Any other subtest that looks similar to subtests from other test-batteries can be said to at least have the indicated guarantee, this being that every AID 2 subtest stood the test as concerns psychometric appropriateness: Eleven subtests fit the Rasch model and the subtests *Immediately Reproducing* (of numerical stimuli), *Coding and Associating*, and *Anticipating and Combining* refer either to measurements of physics (the longest remembered number of digits; speed) or stood an analogous test according to a multi-categorical generalization of the Rasch model.

There are additionally several item conceptions as to avoid interfering abilities or improper moderator variables.

As concerns all but one subtest of AID 2, there is neither a stressing time limit for item solution nor is there any benefit from speed for the solution of the items. That is to say, the abilities of speed and power are not fatally confounded as often elsewhere.

As already mentioned, *Applied Computing* primarily concerns reasoning rather than computing school techniques. There are almost no unhandsome calculations that have to be done. What is required is rather that a problem be solved logically. In other words, numerical reasoning and calculation excellence are not confounded as often in similar tests. In addition, a particular phenomenon has been discovered and disarmed in AID 2: Psychometric analyses have exposed that items misfit the solutions which are of numerous simplicity (e.g. „2“) even though the solution is hard to find by reasoning. This is due to testees, who are not so gifted, guessing and scoring by chance.

Social and Material Sequencing proves that sequencing need not be measured separately with respect to social actions, on the one hand, and to material actions, on the other hand. Bear in mind, that this fact is not too likely *a priori*.

Immediately Reproducing (numerical stimuli) aims at two different kinds of performances: In the first run, the testee has to repeat digits in the same sequence as they have been presented by the psychologist (“forwards”), in the second run, however, the testee has to repeat the presented digits “backwards”. While this also occurs in some other tests as well, AID 2 offers two different scores due to the fact that these scores do not correlate high enough to be summed up – psychometric presuppositions would be much more severe in the case of summing them up than in the case of using two scores on their own. As chance effects, black-outs, and vigilance disturbances are quite likely to occur by repeating a sequence of digits, AID 2 offers a second and a third trial for every sequence length, should the testee fail in the first trial. Bear in mind that if this is the case then it is a matter of the longest ever repeated length score and not (essentially) the number of trials that counts!

Producing Synonyms aims to measure a testee’s vocabulary, as is with many other tests, however takes care not to confound the ability to understand the concept of a word with the ability of verbal expression; hence, just a single word determines the solution. Furthermore, no superior

concept is required in order to describe the meaning of a word because this would most likely interfere with a hierarchical higher cognitive process of abstracting that is tested in AID 2 in a subtest of its own.

In contrast to similar tests, *Coding and Associating* also offers two separate scores. Not only the speed of information processing by a testee but also his/her incidental learning ability are utilized in this subtest. In other words, the testee does not, of course, have to learn the codes by heart, but simply may remember them “by the way”. Conceptualization of this particular subtest resulted from practical experience that has proven that certain testees code many more symbols at the end of such a test than at the beginning, while other testees do not follow this pattern. The signs that are used and which have to be reproduced are simple rather than any typical school character symbols, which could possibly handicap someone or produce a reactivity effect – it should be added that due to the immense simplicity of the signs, the motoric skills shown here are probably not of great importance.

Anticipating and Combining is the single AID 2 subtest where quick solutions can be of any benefit to the scoring. As a matter of fact, scoring either only certain correct pieces or every single correctly puzzled pieces did not succeed in fitting a multi-categorical generalization of the Rasch model, however scoring the items with regard to “no solution”, “slow solution”, and “quick solution” does fit this generalization. Psychometric analyses have also disclosed that the matter of quick solutions being of benefit to a testees score should not be applied for the two easiest items. In order to avoid solutions by chance and failures by misfortune, respectively, because of a lucky or unlucky decision by the testee for whichever two pieces he attempts first, an anchor piece is provided to which every other piece borders. It is also important that any cut of a figure happens in a most pregnant way, that being according to some functional unit of the given everyday object; this is in order to provoke a content-based cognitive process rather than a trial and error procedure.

In contrast to other similar tests, the subtest *Abstracting* is restricted to common functionalisms of two given objects but does not refer to the ability to produce any kind of superior concept.

Since the cubes’ surfaces in *Analyzing and Synthesizing* are variously colored, the given patterns represent highly complex figures whose decomposition into several single cubes is not at all obvious. To be more explicit, red cubic surfaces often border with other red cubic surfaces and white cubic surfaces often border with other white cubic surfaces. Hence, the items demand for a lot of field independency which seems to be a more analytic ability rather than a mere perception of the lines that are quite obviously cut. Bear in mind that psychometric analyses have proven that any partial solution, for instance, with regard to the number of correctly composed cubes, does not serve as a fair scoring.

Social Understanding and Material Reflection proves that comprehension of pertinent relationships and coherences

need not be measured separately with respect to social relations, on the one hand, and with regard to material relations, on the other hand. Of course, this fact is not too likely a priori!

Immediately Reproducing (figural stimuli) consists of abstract figures as well as of pictures of everyday objects; with regard to psychometric analyses these different kinds need not be measured separately. According to psychometric analyses this subtest also proves that a second or third trial does not pay off because any benefit scoring of a correct reproduction then is not at all fair in comparison to the case where a testee fails a certain sequence of figures three times. Bear in mind that the items are conceptualized in such a way that any significant arrangement of a sequence's figures would not privilege testees with any spatial superiority.

The subtest *Learning and Remembering* proves that senseless stimuli need to be presented twice because after being presented only once chance effects, black-outs, and vigilance disturbances might prevent the scoring being fair.

Recognition of (figural) Structures refers to the analyzing process in the subtest *Analyzing and Synthesizing*; however this particular subtest excludes any support or handicap which would come from the actual handling of the cubes and focuses instead on the field independency: The given patterns differ, in contrast to *Analyzing and Synthesizing*, with respect to the scales of the cubes being reduced.

A Non-Verbal Instruction

Several subtests within AID 2 offer a non-verbal instruction for testing children and juveniles who suffer from handicaps, this being, above all, testees who have a foreign

language as their native tongue and those who are socially handicapped – although the (verbal instruction) AID 2 norms have proven that there are no significant differences between testees who stem from different social groups (with regard to the father's profession), it might be advisable, though, not to risk any kind of unfair testing. The utilization of a non-verbal instruction for testees with a foreign language as their native tongue or with auditive deficiencies is obvious. These AID 2 non-verbal instructions are standardized just as the verbal instructions are (see for example the non-verbal instruction of the subtest *Anticipating and Combining* in Figure 2).

Of course, Competence in Realism, Social and Material Sequencing, Coding and Associating, Anticipating and Combining, Analyzing and Synthesizing, Learning and Remembering, and Recognition of (figural) Structures enable such a non-verbal instruction. If only these subtests would be used this pool of subtests represent, according to a factor analysis, two of the established four AID 2 factors (see below); that is to say, they serve to provide a wide spread impression of a testee's abilities so that serious consulting is to some extent still possible.

Koller (2001) proved, however, that the subtest's *Competence in Realism* non-verbal-instruction is not suitable because from 62 eight to twelve years old subjects without any confirmed deficiencies as yet, only 34 were able to understand the non-verbal instruction; the other testees needed to be additionally advised as per the verbal instruction. Hence, until this point it is not advisable to use this subtest in its non-verbal instruction form. However this does not apply to other subtests. As a matter of fact, no significant differences between verbal and non-verbal instruction were found for the other subtests in Koller's experiment.

The pieces of the first figure (item) have to be placed in front of the testee, with the anchor piece (the middle piece) directly in front of him/her.

The examiner then puts his/her own left hand onto the anchor piece and rotates it a bit to the left and to the right (2 - 3 times) and finally brings it back into its original position again. While rotating the piece left and right, the examiner also holds up the right palm and shakes his/her wrist as a sign for "no". Then, he/she lifts up his/her left hand about 2 inches, only to press it down onto the anchor piece immediately again; this lifting and pressing down of the left hand is to be repeated 2-3 times.

While continuing to hold the anchor piece with the left hand, he/she now taps the remaining pieces of the first figure, one after the other, with the index finger of the right hand, then with the same finger, he/she taps the table around his/her left hand 3- 5 times. After that, he/she again taps at any piece, lifts the index finger from that piece with deliberate slowness, draws a noticeable arrow into the air with the index finger towards the left hand, and finally taps the table around the left hand 3-5 times. This procedure should also be repeated for any other piece.

Now , the examiner lifts his/her left hand from the anchor piece and invites the child with the open palm directed towards the pieces to begin with the task.

Figure 2. Standardized non-verbal instruction of AID 2 subtest *Anticipating and Combining*.

A Pool of Optional Subtests

The subtests *Immediately Reproducing* (figural stimuli), *Learning and Remembering*, and *Recognition of (figural) Structures* are optional since they indicate certain specific and basic neurofunctional disorders (or exclude them) rather than measure cognitive abilities due to the pragmatic approach of testing intelligence as sketched above.

Immediately Reproducing (figural stimuli) was added to AID 2 since the subtest *Immediately Reproducing* (numerical stimuli) demonstrates the capacity of short term memory with respect to consecutively presented information, however a low performance in the latter may simply be the result of the testee's difficulties with numbers. Hence, should the testee display a low performance in *Immediately Reproducing* (numerical stimuli) the psychologist is advised to then administer *Immediately Reproducing* (figural stimuli). If, again, a low performance occurs then a general deficiency in (short term) sequencing memory is most likely to be the cause.

However, low performances in both, *Immediately Reproducing* (figural stimuli) and *Immediately Reproducing*

(figural stimuli), does not necessarily have to be the result of a general short term memory disorder. It is possible that it is a matter of a disorder of the ability to perceive or remember or reproduce stimuli in a given sequence. In order to prove this psychologists are advised to additionally apply the subtest *Learning and Remembering*. The sequence is not of any relevance in this particular subtest but rather the testees have the chance to use their memory capacity store since many people simply need to learn a material twice.

Finally, *Recognition of (figural) Structures* serves to discriminate between low performances in the subtest *Analyzing and Synthesizing* that are caused by the testee's inability to decompose abstract figures into given units or low performances that are caused merely by the testee's problem with composing such figures by handling the pieces manually. In case a testee also has a low score in the subtest *Recognition of (figural) Structures* the disorder to be established may be even more severe because the ability to analyze seems to be a prerequisite for synthesizing.

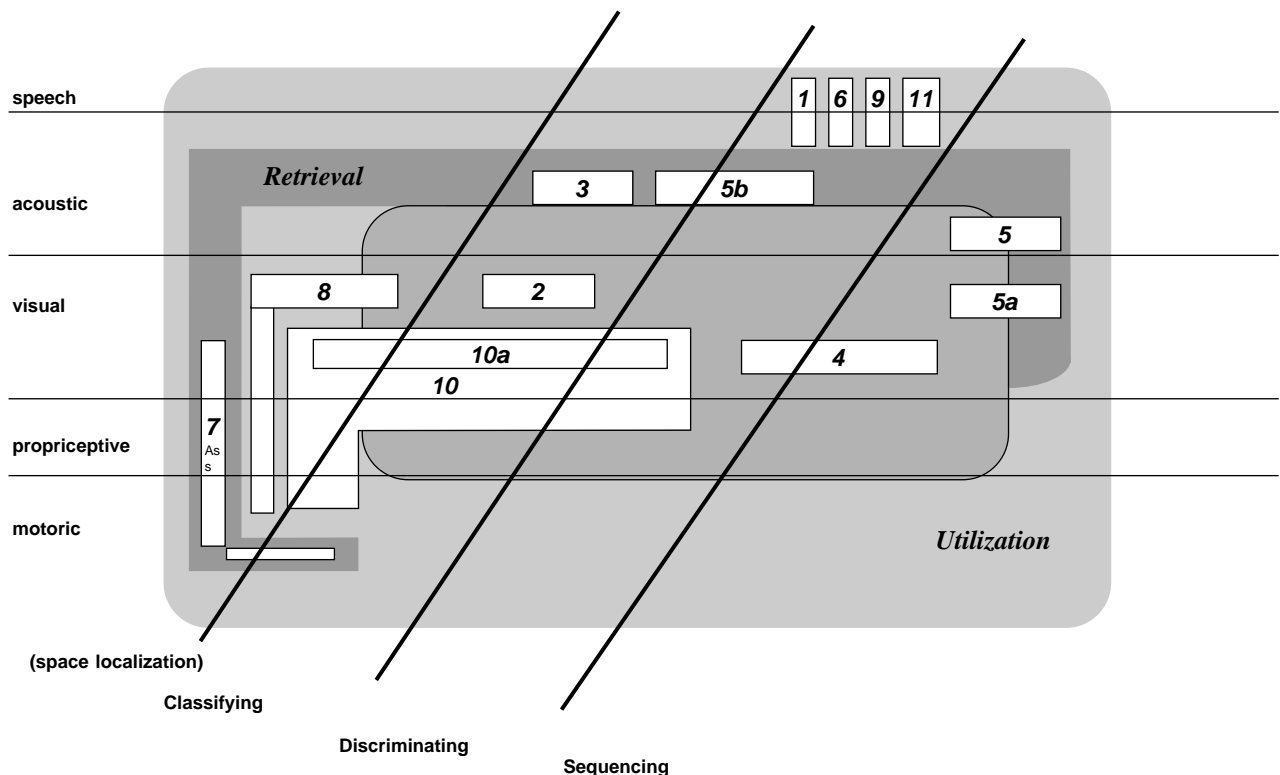


Figure 3. The AID 2 model of developmental disorders or learning disabilities. It distinguishes, primarily, between perception, retrieval, and utilization. The numbers in the given diagram correspond to the AID 2 subtests as follows: 1- Everyday Knowledge; 2- Competence in Realism; 3- Applied Computing; 4- Social and Material Sequencing; 5- Immediately Reproducing (of numerical stimuli); 6- Producing Synonyms; 7- Coding and Associating- Numbers of Associations; 8- Anticipating and Combining; 9- Verbal Abstraction; 10- Analyzing and Synthesizing; 11- Social Understanding and Material Reflection; 5a- Immediately Reproducing (figural stimuli); 5b- Learning and Remembering; 10a- Recognition of (figural) Structures.

A Survey to Identify Specific Developmental Disorders or Learning Disabilities

Enriched with these optional subtests, AID 2 offers a survey for the case that specific developmental disorders or learning disorders are suspected. AID 2, however, does not proclaim to serve such a psychological investigation with respect to every thinkable disorder or disability. Nevertheless, AID 2 does proclaim a screening approach for indicating what kind of deficiencies, learning disorders or basic neuro-functional disorders, handicap an individual's performance. A particular model of AID 2 (cf. Fig. 3) is the result of a post-hoc conceptualization: the authors managed to breakdown the abilities that are necessary to solve the respective subtests.

The authors focused on the basic neuropsychological skills, if given in AID 2, which have been commonly discussed in literature on learning disorders. This model is again a process oriented one. According to some neuro-functional hierarchy, it distinguishes between perception, retrieval, and utilization. Perception applies to several modalities, those being visual, acoustic, and proprioceptive activities – visual and acoustic modalities can be tested with lexical or numeric content; visual modalities can additionally be tested with figural content. Perceiving concerns classifying and discriminating (space localization included), on the one hand, and sequencing on the other hand. Sequencing refers to the ability to handle consecutive stimuli in an appropriate manner. Retrieval also refers to the same modalities and contents, but only functions well if perception has already been mastered. Bear in mind that retrieval however does not necessarily have to function

well even though perception may already have been mastered. Besides the modality and general content, as well as modality and content specific disorders of the ability “retrieval”, there may also be general or specific deficiencies particularly with regard to sequencing. Utilization also involves the same modalities and contents but frequently also incorporates intermodal transfers with motor aspects, at least with regard to speaking. Inter-modal and motor, particularly visuo-motor, abilities can be compromised even though perception and retrieval function well.

Obviously, by identifying any such dysfunction practitioners are able to decide on an appropriate treatment.

Some Discriminative Indicators for Intellectual Neglect vs. Intellectual Advancement

In accordance with Cattell's polarization of crystallized intelligence, on the one hand, and fluid intelligence, on the other hand, consulting psychology often looks for indicators as to whether a child or even juvenile has lived in an (intellectually) stimulating environment or rather in a more (intellectually) deprived environment. Of course, once intelligence deficiencies are established, it does make a difference whether a testee comes from one of the above described environments.

In the AID 2 manual empirical evidence is given for the fact that – at least for children up to an age of ten – the subtests *Everyday Knowledge*, *Producing Synonyms*, and *Social Understanding and Material Reflection* do indeed

Table 1

Loadings of the four factors solution of 16 scores (T-values) of AID 2 (662 children and juveniles); the highest loading of any score is fat printed. Communalities (h^2) of the four factors solution and eigenvalues of the complete solution are given as well

	1	2	3	4	h^2	Eigenvalues
Everyday Knowledge	.70	.26	.12	.23	.61	5.40
Competence in Realism	.40	.28	.09	.12	.27	1.44
Applied Computing	.56	.48	.28	.12	.60	1.34
Social and Material Sequencing	.37	.26	.13	.16	.29	1.07
Immediately Reproducing (of numerical stimuli)						
Forwards	.17	.16	.83	.03	.75	0.89
Backwards	.16	.11	.85	.06	.76	0.82
Producing Synonyms	.81	.16	.16	.08	.66	0.79
Coding and Associating (number of)						
Coded Symbols	.13	.11	.09	.83	.73	0.73
Associations	.13	.13	–.02	.84	.73	0.66
Anticipating and Combining (the parts of a cut up “Gestalt”)	.28	.63	.03	.11	.52	0.55
Abstracting (the common functionalisms of things)	.81	.18	.07	–.01	.67	0.46
Analyzing and Synthesizing (abstract figures)	.31	.76	.01	.05	.65	0.45
Social Understanding and Material Reflection	.81	.18	.07	.08	.70	0.42
Immediately Reproducing (of figural stimuli)	–.03	.43	.31	.31	.33	0.36
Learning and Remembering	.35	–.06	.31	.34	.36	0.33
Recognition of (figural) Structures	.19	.71	.04	.02	.52	0.30

provide such an indicator: If a child has, relatively speaking, top scores in all these subtests but poor scores in the other subtests then one can suppose that this is related to the child coming from an (intellectually) stimulating environment. On the other hand, if a child has poor scores in the above mentioned subtests but has high scores at least in some other subtests then one can suspect that the child lacks intellectual stimulation. These conclusions are based on two studies which put AID 2 scores in relation to a questionnaire score, which inquired into the intellectual environment of the child and respective activities undertaken within the family.

Abandoning the IQ

The result of a factor analysis of the AID 2 subtests (principal component analysis, Kaiser's Varimax criterion of factor rotation) are four factors (see Table 1 for the factor loadings in detail). These are: Information processing in the given social surroundings, information processing of new contents, capacity of intellectual apprehension, and ability to (re-) produce by structuralization (cf. the model by Roth, Oswald, and Daumenlang, 1980, cited above).

At any rate, these four factors definitively exclude the existence of any Spearman-like general factor model. Hence, the authors do not recommend that one use any kind of global IQ by computing the average of the subtest scores.

This test battery is of particular use if psychologists are looking for a test-battery that indicates directly specific preventive or rehabilitative treatments, however it does not serve merely as an assessment of a testee's level of so-called "general intelligence". In other words AID 2 focuses on each and every indication in order to decide what treatment is to be best implemented. For this purpose, the profile of the subtest scores is of major interest and importance.

So, if any global index is to be of diagnostic relevance it would be the level of the lowest subtest score in AID 2, the so-called "intelligence quantity", on the one hand, and the range of the subtest scores, on the other hand. This conceptualization comes from the idea that any chain is just as strong as its weakest link. On the other hand, practitioners' experience discloses that too much heterogeneity in a testee's intelligence profile indicates a certain risk of its own, just as both underachievement as well as over-achievement may irritate the person him/herself as well as his/her parents and teachers. An AID 2 range of (13) subtest scores that is larger than 40 *T*-scores is found within 10 percent of the population, at the most. The (lowest level of) intelligence quantity amounts to a mean *T*-score of 36 (given 13 subtest scores). A quarter of the population has an intelligence quantity of lower than 28. In order to reduce any kind of chance effects in AID 2, the second lowest score has also been normed; in the case that the second lowest score is not low in comparison to the lowest *T*-score, this would indicate a

particular deficiency rather than a general deficiency. Therefore, a first step towards a global interpretation can be taken by using only these three indices, that is to say, before a detailed test profile interpretation may take place.

A Schedule of Behavioral Misfits

In accordance with the proclaimed significant diagnostic relevance and practical importance of stimulating a testee's interaction with various materials as well as with the psychologist himself, AID 2 offers a standardized behavioral check list which may be used after test administration. The psychologist is given a schedule which serves as a retrograde observation support of remarkable behavioral aptitudes. Of course, misfits are of primary interest.

No doubt, psychologists who have a great amount of experience with intelligence testing might not need such a schedule but this would rather be of use to inexperienced psychologists. The schedule helps the psychologist to systematize and declare his/her implicit impressions. This evaluation schedule of a testee's attitude and willingness to work as well as of his/her communication virtues refers overall to his/her achievement motivation (apathetic, disinterested *vs.* keen, pleasure at performing *vs.* over exaggerated achievement orientation), alertness (fluctuating, easily distracted *vs.* well concentrated *vs.* restricted, rigid), and the like.

Further Advantages Because of an Adaptive Testing Item Pool

Moreover, AID 2 offers certain additional administration procedures. For instance, by implementing a kind of branched testing, testing might even cease after the second subset of five items. This might particularly be the case when the testee masters two or three items in the second subset – by then the testee's achievement level should have already been sufficiently determined.

One should take note that it is also possible to administer AID 2 conventionally. Any psychologist may choose three subsets that he administers to every testee irrespective the testee's age and achievement level. Even a conventional two-subset administration is imaginable. Of course, the error of estimation would become much greater in these cases.

The given item pool of 60 homogeneous items within each of the subtests spoken about enables one to establish a parallel test concept for retesting a testee. The AID 2 subtests, with the possibility of parallel tests, that consist explicitly of items that have not been administered before to the testee under consideration – although the testee's first time achievement level is taken into account – are particularly useful for consulting practices where retesting is quite the routine affair.

Last but not least, it should be mentioned that a computer program is also available at the psychologist's disposal which not only calculates *T*-scores and percentile norms but also offers confidence intervals for every subtest. This being for the purpose of establishing whether a significant test profile exists at all. This decision is of vital importance because in the case that significant deeps – in contrast to any tops – are established then proper performance training programs might be decided upon. One should also bear in mind that this significance is however more likely to be established in AID 2 than elsewhere due to the minor errors of ability parameter estimation that result from adaptive testing.

Conclusion

The test-battery AID 2 illustrates that a practitioner's need for a more precise intelligence test is, indeed, feasible.

For commonplace consulting, adaptive testing even works without the online use of a computer. In particular, it does not just work out to be more economical than conventional testing but also offers other advantages: According to empirical experience, the psychological benefits of adaptive testing are, on the one hand, a severe reduction of frustration due to failure and, on the other hand, one avoids deflating the achievement motivation of the testee due to the fact that no challenge is presented. Furthermore, the administration of AID 2 does not usually require any longer than 60 minutes, and it offers, nevertheless, a wide-spread test profile.

Moreover, AID 2 demonstrates developments of content which guarantee for high quality consulting. First of all, the items have been conceptualized so as to avoid any interfering abilities or improper moderator variables which have not yet been realized. AID 2 can also be said to be treatment centered, in other words, AID 2 offers a survey for identifying specific developmental disorders or learning

disabilities. AID 2 tries to avoid any handicaps in the scoring by additionally offering a non-verbal instruction for item presentation. And one should finally note that, in comparison to other test-batteries, AID 2 takes into consideration various discriminative indicators regarding either an intellectually neglected testee or an intellectually advanced one.

References

- Gittler, G. (1986). Inhaltliche Aspekte bei der Itemselektion nach dem Modell von Rasch [Rasch model-based item selection and the matter of content]. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 33, 386-412.
- Koller, M. (2001). *Die Angemessenheit der Normen im AID 2 für die nonverbale Instruktion* [The appropriateness of AID 2 norms for using the non-verbal instruction]. Unpublished master's thesis, University of Vienna, Austria.
- Kubinger, K.D. (1998). Psychometric shortcomings of Wechsler's intelligence scales - Results on the German WISC, conclusions for the WISC-R. *Review of Psychology*, 5, 3-12.
- Kubinger, K.D. (2004). Towards economic Wechsler-like testing: Adaptive Intelligence Diagnosticum (AID 2). Manuscript submitted for publication.
- Kubinger, K.D., & Wurst, E.(2000). *Adaptives Intelligenz Diagnostikum (AID 2)*. [Adaptive intelligence diagnosticum.] Göttingen: Beltz.
- Roth, E., Oswald, W.D., & Daumenlang, K. (1980). *Intelligenz*. [Intelligence.] Stuttgart: Kohlhammer.
- Turnbull, W.W. (1979). Intelligence testing in the year 2000. *Intelligence*, 3, 275-282.

Received August, 2, 2004

Revision received September, 10, 2004

Accepted, September, 15, 2004