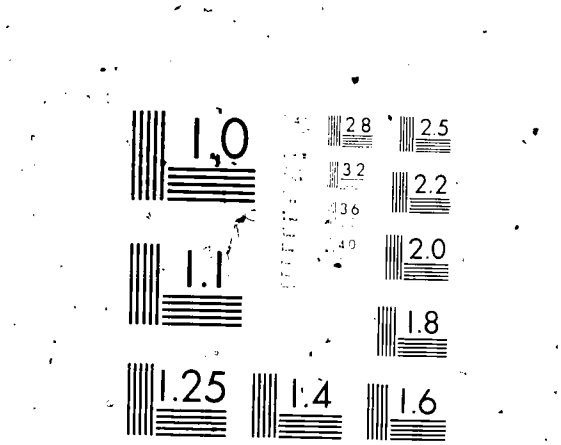

DOCUMENT RESUME

ED 132 173 TM 005 883

AUTHOR Meyers, Edna O.
TITLE The Revised WISC: Does It Serve Inner City Children?
PUB DATE [Mar 76]
NOTE 21p.; Speech given before the Annual Convention of the American Orthopsychiatric Association (53rd, Atlanta, Georgia, March 4-5, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS *Comparative Analysis; Culturally Disadvantaged; *Disadvantaged Youth; *Educable Mentally Handicapped; Elementary Education; *Intelligence Tests; Minority Group Children; Standardized Tests; Statistical Analysis; *Test Bias
IDENTIFIERS *Wechsler Intelligence Scale for Children; *Wechsler Intelligence Scale for Children Revised

ABSTRACT

In order to initiate appropriate programs for screening, early intervention, and meaningful psycho-educational programs, a children's community mental health center in Harlem, where over 95 percent of its clientele is black and/or Spanish speaking, requires, for all its referrals, an accurate assessment of intellectual strengths via an IQ battery. This investigation compared scores achieved by approximately 150 youngsters residing in East and Central Harlem on the Revised version of the Wechsler Intelligence Scale for Children (WISC) with scores achieved by a comparable population on the 1949 version of the WISC. Contrary to expectations, the revised version (supposedly standardized, unlike its predecessor, to reflect the 1970 census for the critical variables of race, occupational head of household, geographic region and the urban rural distribution) produced significantly lower scaled scores on several subtests which, in turn, significantly lowered full scale IQ scores. This fact seemingly penalized the very population which, it was indicated, was most likely to gain. Appropriate statistical measures were applied to determine the true significance of these results, taking into account sex and age of subjects. Implications for continued use of this new instrument were discussed and conclusions drawn. (Author/MV)

THE REVISED WISC: DOES IT SERVE INNER CITY CHILDREN?

Edna O. Meyers

THE CITY COLLEGE, CUNY
(Asst. Professor)

and

THE NORTHSIDE CENTER FOR CHILD DEVELOPMENT
(Chief Psychologist)

Prepared for delivery at panel: Reducing Transcultural Distortions in Clinical Practice. 53rd Annual Meeting of the American Orthopsychiatric Association, Atlanta, Georgia, 4, 5 March, 1976.

With the appearance of the revised version of the WISC in 1974, Northside Center, an accredited children's psychiatric clinic in harlem, immediately purchased this instrument, grateful in advance that here might be a more appropriate assessment device for our inner city children. It had been standardized, so we read in the advance publicity, to include our population - non-white, poor.

Throughout the year from July 1974 through June 1975, our staff, composed of two black and two white psychologists utilized the revised WISC. By the end of that first year, however, we reverted back to the 1949 version of the WISC - and therein lies the reason for this paper.

Let me give you some background information about our clinic. The Northside Center for Child Development was founded three decades ago by Kenneth and Mamie Clark, primarily to rescue from out of the CRMD (children with retarded mental development) classes in that section of the inner city the hundreds of youngsters placed there by the standardized administration of group IQ tests to elementary grade pupils.

The Center, as you can deduce from its very name, was basically concerned with child development, with a stated focus on contributing towards the healthy growth of children in the black community rather than ministering to and treating as uppermost their pathological components.

One pathological component, ego devaluation, was being fostered and perpetuated not only by the economic disadvantages which all people who live in poverty and frustration experience,

but also and equally by the side effect of minimal educational standards set for the schools of this community. The justification for this tragic directive was the so-called "inferior" IQ's of its children.

Because of the importance of challenging the results of group IQ tests which, in the inner city community, provided a high population for the CRMD classes and no population at all for the IGC (intellectually gifted children) classes, Northside gave and still gives each child referred to it an individual IQ test.

With an average of approximately 150 IQ administrations a year, over 2,000 test protocols since 1960 were overviewed. The conclusions all our psychologists arrived at during the period preceding the WISC-R were 1. many of the children's responses, scored as wrong according to the 1949 manual, were the result of forethought, foreknowledge and experience, and that these therefore should have been given credit; and 2. both psychometrician and child were being kept in a bind. Because the agency was challenging the assessments given by the schools, it had to guarantee its results of IQ testing as beyond question or reproach; it had to administer a standardized test, and score according to Hoyle.

Our staff was sufficiently concerned both about our testing procedures and our recommendations to overview one typical year's WISC results (1972-1973). We wanted data on what real live children from the real live Harlem community were telling us about their own lively intelligence, not only where their IQ scores fell.

It is important to emphasize, at this point, that the youngsters being assessed had been referred to the Center for a multitude of so-called clinical reasons. Presenting complaints were behavioral deviations and/or gross school failure of all varieties, compounded by teacher and/or parental helplessness to deal appropriately with either problem.

Table 1 does indicate that our full scale IQ means corroborated earlier statistical results which promulgated that the inner city minority child scored, as a general rule, within this range. But, remember, this was a "pathological," not a "normal" population, and the question we asked ourselves was whether, if we were testing an equal number of so-called "normal" inner city children, the results would have been identical? Is not the likelihood more possible that the scores could have been closer to the accepted mean?

TABLE I

MEANS, STANDARD DEVIATIONS AND RANGE OF WISC SCORES FOR TOTAL CLINIC POPULATION (during one year) AND ALSO FOR POPULATION WITH IQ SCORES OF 89 AND BELOW

	VERBAL	PERFORMANCE	FULL SCALE
Total Group N-131			
Mean	86.82	91.47	88.02
SD	14.75	16.32	15.23
Range	58-126	53-139	51-136
Boys N-93			
Mean	87.3	92.4	88.7
SD	14.66	16.58	15.43
Range	58-123	53-135	51-125
Girls N-38			
Mean	85.5	89.1	86.3
SD	14.87	15.41	14.58
Range	60-126	57-139	54-136
Select Group N-75* (IQ's = 89 or below)			
Mean	78.7	81.7	78.2
SD	9.99	10.41	8.64
Range	58-100	53-104	51-89

*Boys (N-53) and Girls (N=22) combined.

But what concerned us particularly was the scoring procedures we religiously abided by with that percentage of our youngsters categorized objectively as of borderline intelligence or mentally retarded (i.e. at 89 IQ or below).

Our data for 1972-1973 revealed that of our total N of 131, over fifty percent were in this category and, within this lower IQ group, there were twice as many boys as girls. Table I provides this information as well.

We then studied four of the verbal subtests - subtests more often yielding alternative answers which, by our lights, although scored as incorrect, reflected the manual's rather than the subject's deficiencies. Let me give you some examples.

On the Information subtest, the child was asked, "What does the stomach do?" Well, said our children, it growls, it rumbles, it grumbles, it aches, it feels empty, it feels full, it holds the food. Zero score for all of these! (Only the final answer is now also considered acceptable in the 1974 revised WISC). Ask yourself whether, without any scientific or physiological introduction to the question offered by the examiner, a child giving any of those various answers would not be supplying information based on real experience?

Where is Chile? we asked the child. He answered at the North Pole or where the Eskimo lives or when it's cold outside or on a sandwich or something you eat that's hot. Any one of these answers could have legitimately reflected acquisition of information, but not the one the manual asked for, so our child was denied credit.

Or again: When asked what C.O.D. means, if an inner city child answered a fish, or someone you prayed to, you might fault his spelling in the second response, but was he wrong? Especially since he didn't see the punctuation which indicated an abbreviation? (In the revised WISC, the wording of the former question has been altered to remove the ambiguity, and the latter question has been eliminated. But, for a quarter

of a century; how many kids have lost credit which, in some way or other in the area of this particular subtest, had some effect on the final IQ score?)

The question about what you would do if you were sent to purchase a loaf of bread and the grocer said he didn't have any raises an alternative answer- and not necessarily a reflection of dependency needs. Certainly there has been enough criminal molestation of children in the inner city to justify enjoining a child not to go too far afield nor stay out too long on an errand.

The Similarities subtest provided our investigation with an interesting phenomenon. Remember, our studied population were all children scoring borderline or retarded in intelligence. Yet, on Table 2 it can be observed that this subtest yielded the highest mean score, and this is the subtest which Wechsler himself once considered truly indicative of intellectual potential. In this area, also, we challenged some of the questions.

It was open to debate whether the seasonal and often highly priced peach and plum should have been the opening gambit. (As a matter of fact, the revised WISC changed the fruits to apple and banana). And when was the last time our city kids had a real experience with coal, or thought of paper as anything other than what you did your homework on or used in the bathroom? What English-speaking kid would offer that they were both carbons when carbon is that piece of dark blue paper which make copies for you if you don't use a photocopier?

That paper and coal are both fuel seems to me more the product of acquired information than high level conceptualization.

Or take Vocabulary! All of the children we questioned who failed "cushion" knew what pillow was. (The WISC-R has eliminated that one). As for "fur" - was it a tree? a number? or a skin that kept you warm? Are we dealing with auditory perception or intelligence when a child says that "spade" is what the astronauts travel in?

When we came to "sword" we were again presented with an auditory perceptual ambiguity. Was it to fly like Jonathan Seagull? Was it what cuts wood: Was it when you saw someone climb up a ladder? But wouldn't you agree that "sword" was not necessarily a useful weapon for assaying intelligence?

Hero was a sandwich! Wrong! To join, said a youngster was "like to join a club." Yes, but what does "to join" mean, you repeated, like a good psychometrician. Since the child didn't know he was missing the point, he wondered why you were badgering him. "You know," he answered, "like you join the boy scouts. Say, lady, don't you understand English?" So we went on to the next word....

All of these "wrong" answers came from the protocols of children with IQ's of 89 or below. What we therefore did with our data was rescore some of these responses and, when we felt that forethought and foreknowledge and experience invested the given answer with intelligence, we upgraded the scores for the purpose of this presentation.*

* Sections of this paper were also reported at a paper session at the American Psychological Convention in New Orleans, August-September, 1974.

Eleven youngsters of the 75, we concluded, were slotted into lower categories than were warranted by the substance of their responses: three who scored 69 or below moved into the borderline category; four moved from borderline into dull normal and four out of dull normal into the average category. Table 2 summarizes this information. Now, this may not have had statistical significance, but it certainly could have had practical and human significance for the children involved.

TABLE 2

MEANS AND STANDARD DEVIATIONS BEFORE AND AFTER REASSESSMENT OR FOUR VERBAL SUBTESTS OF THE WISC AND ITS EFFECT ON FULL SCALE IQ OF CLINIC POPULATION WHOSE IQ's WERE 89 OR BELOW

N=75	Original	Reassessed
General Information (scaled scores)		
Mean	5.96	6.36
SD	1.91	1.98
Comprehension		
Mean	6.73	7.15
SD	2.14	2.23
Similarities		
Mean	7.64	8.52
SD	2.35	2.73
Vocabulary		
Mean	5.92	6.83
SD	2.06	2.36
Verbal IQ Score		
Mean	78.7	81.68
SD	9.99	10.75
Full Scale IQ Score		
Mean	78.2	80.31
SD	8.64	8.87

Since, in New York, children with IQ's of 75 or below are placed in CRMD classes, out of the 20 youngsters who scored as being eligible, only 15 would have been retained in that questionable category if we had scored with more flexibility and insight. And I repeat, we overviewed only four of the 10 subtests to arrive at this estimate.

Remember, also, these WISC results represent one year's overview in one small clinic in Harlem. The WISC has been in use since 1949! How many children throughout the world, in the past quarter of the century, subjected to a formal administration of this battery, have had their intellectual potential misjudged and their intelligence falsely assessed? (At this point, it must be reemphasized that our total assessment of the child never deferred arbitrarily to the numerical IQ score; intra- and intertest variability and numerous other observations, both standardized and subjective, pervaded our final recommendations).

Nonetheless, you can understand why we were among the first purchasers of the revised version of the WISC. Why, then, did we discontinue its use after one year?

Several months into the year of the revised version's administration, we began to develop uneasy feelings that, rather than providing an equivalence, or an improvement, in scores, our children seemed to be doing poorly. Had our population undergone change in the severity of their pathology? On the basis of our psychiatric diagnoses, such was not discernible! Were our examiners resistant to the demands of a

new instrument and running interference when congruence was demanded? We had no hard evidence to recommend this possibility. They had accommodated themselves happily to the McCarthy Scales introduced that same year. Further, the eagerness with which they awaited the revised WISC and studied its administration seemed to controvert that possibility.

We discussed our suspicions with school psychologists working in the inner city community. They were experiencing the same reservations! A communique (January 27, 1976) to the author from Dolores Morris, Ph.D., a supervisor of school psychologists, said in part:

> The supervisors and staff members of the Bureau of Child Guidance have been concerned about the lower scores that were being obtained on the WISC-R. In June, 1975, a committee of supervisors were attempting to determine whether a comprehensive study should be undertake to compare the WISC with the WISC-R. They developed a questionnaire to assist them in this decision. Out of a total psychology staff of 348 for the school year 1974-1975, 239 psychologists were directly involved in examining students in the age range that the WISC and the WISC-R would cover. A questionnaire was distributed to the psychologists and 104 responded...You might be interested in the following still unanalyzed results (due to budget cutbacks and limited staff, the study was not done):
>
> 1. If you use the WISC-R, compare it with the WISC by giving your impressions (of the obtained scores - ed. note)
>
> a. Full Scale Higher 3% Lower 29% Same 8% No response 61%
> b. Verbal Scale Higher 3% Lower 29% Same 7% No response 62%
> c. Performance Scale Higher 4% Lower 27% Same 8% No response 61%
>
> 2. Comment on ease of administration of WISC-R as compared to WISC
> WISC-R 14% WISC 22% No response 44% Same 1%
> Don't know 2%
>
> 3. Comment on your preference for the WISC-R or WISC
> WISC-R 29% WISC 1% No response 47% Same 3%
> Don't know 4%

As can be observed, even though psychologists admitted that the WISC-R provided lower Verbal, Performance and Full Scale IQ scores, it nevertheless (by a much smaller margin), was preferred by the psychometricians and this was so even though they found the WISC easier to administer. On the basis of these data, dare we accept the hypothetical assumption that, for the Bureau of Child Guidance, the WISC-R is a preferred evaluative device and is a fairer test of minority children's intelligence precisely because its scores are lower?

More objective data were available, however. A carefully designed study, initiated by the Psychological Corporation and executed by Jerome D. Doppelt and Alan S. Kaufman, estimated the magnitude of the differences between IQ's obtained from the WISC-R and those from the original WISC. Their conclusions confirmed our still unsubstantiated feeling: to wit, that if the 1974 WISC-R sample had been tested with the WISC, the contemporary WISC-R sample would have obtained higher scores on the earlier battery. The WISC-R was producing lower scores!

One more individual incident propelled us into action! A guidance counselor in a prestigious east side school in New York asked us to test a young black boy whom she suspected of being retarded and whom she wanted placed, therefore, in a special class (or, preferably, taken out of that school). Our test scores on the revised WISC seemed to confirm low borderline intelligence, but did not fully support the counselor's suspicions. A school psychologist, exactly four months later,

administered the 1949 WISC, unaware the child had been tested. The full scale IQ score was 89 - 17 points higher. Practice effect could not possibly have accounted for this increase in scores, since these were verbal subtest scores not susceptible to variability. The WISC and WISC-R were not equivalent for individual subjects.

For all these reasons, we initiated our study. We used data from the WISC scores our population produced in 1973-1974 and the WISC-R scores elicited from our subjects in 1974-1975. The population, of course, was different, but in terms of diagnoses, referral complaints and socio-economic status, comparable. The subjects were 137 boys (too few female subjects were available for inclusion). The means, standard deviations and values for all subtests, Verbal, Performance and Full Scale scores were computed. Additionally, the Ss, on the basis of the chronological break from elementary to intermediate school, were divided into two groups: those under eleven years of age and those over age eleven.

As can be perceived from Table 3, means for three of the subtests in the verbal area for boys under 11 were lower on the WISC-R than on the WISC, and this was also apparent in two of the subtest means for boys over eleven, the largest difference occurring in the Comprehension subtest with the younger group and in Similarities with the older group.

TABLE 3

MEANS, STANDARD DEVIATIONS, AND t VALUES FOR WISC AND WISC-R SUBTESTS (Verbal scores)

	BOYS UNDER 11 YEARS			BOYS OVER 11 YEARS		
	WISC	WISC-R	t	WISC	WISC-R	t
Information						
N	32	30		36	41	
Mean	8.15	7.67	0.59	6.28	7.15	0.71
SD	2.74	3.59		2.33	3.45	
Comprehension						
N	32	30		36	40	
Mean	10.00	9.06	1.30	8.31	8.02	0.40
SD	2.45	3.19		3.00	3.27	
Similarities						
N	32	30		36	41	
Mean	9.46	8.63	0.95	8.69	7.00	1.61
SD	3.45	3.36		3.26	3.37	
Arithmetic						
N	30	30		32	41	
Mean	8.87	8.90	0.04	7.03	7.22	0.29
SD	2.51	2.88		3.20	2.40	
Vocabulary						
N	32	30		36	40	
Mean	9.00	9.63	0.84	8.14	8.40	0.35
SD	2.82	2.93		3.19	3.18	

Table 4 tells a different story. For both the older and younger groups, Coding on the WISC-R produced significantly lower scaled scores. For the younger group a significant difference also was obtained in the Object Assembly subtest, while, for the older group, statistical significance occurred in Block Design.

TABLE 4

MEANS, STANDARD DEVIATIONS, AND t VALUES FOR WISC AND WISC-R SUBTESTS (Performance scores)

	UNDER 11			OVER 11		
	WISC	WISC-R	t	WISC	WISC-R	t
PC						
N	32	30		36	40	
Mean	8.81	9.03	0.34	8.36	9.15	1.16
SD	2.36	2.58		3.14	2.66	
PA						
N	32	30		36	40	
Mean	9.53	8.93	0.73	9.17	9.12	0.07
SD	2.63	3.62		3.14	3.16	
Bl Des						
N	32	30		35	40	
Mean	10.13	9.10	1.23	8.57	6.85	2.00***
SD	2.80	3.61		3.23	4.16	
Obj. Ass.						
N	29	30		31	39	
Mean	10.82	9.00	2.18*	9.87	9.18	0.83
SD	3.15	3.19		3.67	3.41	
Coding						
N	31	30		34	38	
Mean	10.45	8.00	3.00**	8.68	5.36	5.01****
SD	2.68	3.56		2.59	2.64	

* $p < .015$ (one-tail test)

** $p < .001$ (one-tail test)

*** $p < .023$ (one-tail test)

**** $p < .0001$ (one-tail test)

Table 5, presenting Verbal, Performance and Full Scale scores, reflects a significant difference between the means of the younger boys in their Performance IQ's, with lower scores produced by the revised WISC. The difference in the older group, while also large, was not significant.

TABLE 5

MEANS, STANDARD DEVIATIONS AND t VALUES FOR WISC AND WISC-R VERBAL, PERFORMANCE AND FULL SCALE SCORES

	BOYS UNDER 11 YEARS			BOYS OVER 11 YEARS		
IQ Scores	WISC	WISC-R	t	WISC	WISC-R	t
Verbal						
N	32	30		36	39	
Mean	94.34	92.17	0.57	85.42	85.69	0.08
SD	13.65	15.31		16.30	13.44	
Performance						
N	32	30		36	39	
Mean	99.28	91.83	1.96*	92.06	84.87	1.59
SD	12.68	16.03		18.60	20.01	
Full Scale						
N	32	30		36	39	
Mean	96.37	91.17	1.44	87.67	83.38	1.04
SD	12.38	15.34		16.54	18.59	

*$p < .025$ (one-tail test)

It is interesting to note, although peripheral to the focus of this paper, that from 1973 on the younger group, on the average, not only has moved into the normal IQ range, but also has IQ's substantially higher in all areas than the older group, both in the WISC and WISC-R means, although WISC means remain higher than WISC-R means. Can this signify that, as the youngsters in the inner city advance to higher grades, they shift from the category of normal intelligence into the borderline

category? Why? What is the educational environment in the inner city doing to children as they continue their schooling?

For all intents and purposes, the revised version of the WISC was modified to replace items that were ambiguous, obsolete or unfair to minority groups. The revised instrument also proclaimed a more equitable distribution of its standardization sample, including non-whites and the poor, setting up five occupational groups - the lowest being those categorized as "laborers, farm laborers and farm foremen." (p. 21 WISC-R manual).

Now, our inner city population, based on statistics which we acquire for each of our subjects, offers some - perhaps critical-exceptions. Fully 90 percent of our referred children are on Medicaid, and the majority of them come from families who have been on welfare into the second and third generation. Does the revised WISC consider such subjects and such families non-persons? They certainly are not included in the standardization! Are female headed households included in the sample? This is not indicated.

Overviewing our data, we see that scores are lower in most subtests on the WISC-R although not statistically so in all areas, than those obtained on the WISC. Where scores are statistically lower, these occur in the performance area, and most consistently in the Coding subtest which - if it pretends to any input into intellectual functioning, serves merely to deal with an ability to do exactly what you're told to do, using symbols. Coding is usually the final subtest administered; it requires the appearance of a red and blue

printed booklet form (an added expense to the purchaser) and a red pencil sans eraser. Is it a fatigue factor, a distractibility factor, or just plain boredom which so significantly lowers the scores of our population?

To sum up, those of us who work with minority groups have not as yet observed that the WISC-R has provided this sector of the population with a fairer or better assessment device.

One question still looms large in the psychological field. Can intelligence tests, such as the Wechsler, in the hands of most of the psychometricians academically trained to administer them, truly assess the intellectual potential of a minority group? The very administration of these batteries is most often required as the result of a situation which has already raised questions about a child's intelligence. What distortion occurs as the result of a devalued intelligence in the first place (all other variables being equal)?

If we want to eliminate transcultural distortion in clinical practice, the diagnostic approach to pathology (a hangover from the medical model) must undergo radical surgery. A reformulation of testing procedures must be instituted. Norms must be viewed suspiciously or meaningfully disregarded. Have we the temerity, in short, to transform the concept of the intelligence test into a teaching procedure, while using it as only one aspect of the total evaluation of the child's potential?

As serious psychologists, we cannot permanently advocate the infallibility of our instruments, nor the expertise of our

practitioners. Nor are we brash enough to suggest that we are privy to all the special coping mechanisms of the inner city community.

There is really only one asset we may be able to lay claim to and that is that we have a moral, social and community commitment to help inner city children develop towards their full potential. Where the school fails, or the environment defaults on it obligations, we have the professional responsibility not only to predict, but more so, to actualize potential.

If we are to tackle the question of eliminating, or at least minimizing, transcultural distortions in clinical practice three radical assumptions may provide the rationale for altering our testing procedures.

The first of these stems from the remarks that Professor Alexander Luria made at the 37th Annual Meeting of the American Orthopsychiatric Association in Chicago in 1960. Questioned about the Soviet Union's limited use of the IQ test as an objective method for assessing intelligence, he responded by stating that IQ tests were used in his country, but not for the purpose of <u>establishing</u> the child's IQ; rather for the purpose of <u>elevating</u> it. After the testing, he explained, the child was taught what he did not know, the instruction based on giving him an understanding of the cognitive skills employed in developing the appropriate "set" for answering the question. If, at a later date, retesting indicated the subject had retained this knowledge, hadn't he

proven himself more intelligent? Luria said, in effect, "We value our people! We want them to be as intelligent as possible, so our psychologists help them." Can we accept this approach?

The second assumption derives from the first. In 1962, the Massachusetts Institute of Technology published the translation of Lev Vigotsky's "Thought and Language," with an introduction by Jerome Bruner. There the definition of intelligence was offered as the "ability to benefit from instruction." This gives us our cue! Can we develop a paradigm for instructional procedures which concretely aims at developing the abstract attitute in young children.

The third assumption flows from the second. The act of thought itself, the engagement in problem-solving, once it is removed from the imperatives of school-oriented materials, separated from the performance of text-book chores, is an engaging and enjoyable process. It is a brain-storming experience which children, parents, yes, and especially teachers in the inner community, have lost sight of. The solution of problems provides the solver with emotional as well as cognitive pleasure. We have deprived our children - our adults - of this pleasure and offered them expensive hard- and soft-ware pap instead. Let's put the enjoyment of thinking back into the curriculum of life! And who can do it better, perhaps, than those in the mental health field whose goal must become the prevention of pathology rather than its measurement?

# # #