Available online at www.sciencedirect.com

SciVerse ScienceDirect

Procedia - Social and Behavioral Sciences 46 (2012) 2104 – 2110

Procedia
Social and Behavioral Sciences

WCES 2012

An examination and comparison of the revisions of the Wechsler Intelligence Scale for Children

Fatih Kezer[a] *, Recep Serkan Arik[b]

[a]Ankara University, Faculty of Educational Sciences, Ankara, 06590, Turkey
[b]Ankara University, Faculty of Educational Sciences, Ankara, 06590, Turkey

Abstract

One of the most frequently used tools to measure intelligence, which has been accepted to be measurable, is the Wechsler Intelligence Scales. The first of the scales which were prepared for three different age groups is "Wechsler Intelligence Scale for Children" the first form of which was brought forth by Wechsler in 1949. There are four revisions of this 1949 test. This study examines the changes the Wechsler Intelligence Scale for Children has undergone throughout this process, what kind of changes has taken place in its number and structure of sub-scales, the changes in the score types and scoring, and it also examines the techniques used in putting forth its psychometric qualities and the data related to the studies.

Keywords: Intelligence, Wechsler, revisions;

1. Introduction

Individuals are different in terms of their skills of comprehending the concepts, adapting to the environment, developing new ways of cogitation, and overcoming the obstacles they face. When the interaction between the human being and the environment is taken into consideration, it is obvious that the behavior of an individual is shaped according to their heritage, quality of the stimulants in their environment, and the effect of these environmental stimulants on the individual; and these form individual differences (Özgüven, 2007). Comprehending human behavior is possible only when one understands what the quality of these differences is, and when one understands the level of these differences (Tyler, 1965). Determining the characteristics of an individual, from his physiological qualities to his psycho-motor skills, from knowledge and skills to personal characteristics, is important in order to understand human behavior, and in relation, to construct and form educational programs in accordance with these individual differences. One of the most important and least understood characteristic of an individual is intelligence (Özgüven, 2004).

Several views have been put forth arguing that heritage or environmental factors are effective in the emergence of intelligence, and researches have been made. John Locke brought forth, for the first time, the idea that environmental factors are effective over intelligence, and the idea that heritage is effective over intelligence was expressed by Jean Jacques Rousseau, for the first time. Findings obtained from researches have shown that both

* Fatih Kezer. Tel.: +9-312-363-3350
E-mail address: fatihkezer@yahoo.com

1877-0428 © 2012 Published by Elsevier Ltd. Selection and/or peer review under responsibility of Prof. Dr. Hüseyin Uzunboylu
doi:10.1016/j.sbspro.2012.05.436

heritage and environmental factors have an effect on intelligence with varying percentages (Liebert, Ritan & Kail, 1986; Bartholomew, 2005).

Throughout ages, several theories have been developed about intelligence (Sperman Two-Factor Theory of Intelligence; Thordike Multi-Factor Theory of Intelligence; Group-Factor Theory; Guildford's Three-Dimensional Intellectual Structure Model; Stenberg's Theory of Hierarchy; Piaget's Theory; Gardner's Multiple Intelligence Theory, etc), and the aspects forming and affecting intelligence have been examined from various angles (Toker, 1968). Although scientists are all in agreement about the measurability of intelligence, they cannot come up with a single, unified definition of intelligence (Walsh & Betz, 1995). Despite the differences in definitions, a consensus was reached about the basic things intelligence tests measure. One of them is one's "ability to learn" the topics which are part of the curricula of schools. However, this pretty much limits the meaning of intelligence. Another issue is that intelligence can be measured indirectly. According to some psychologists, intelligence can be defined as the ability to adapt to one's environment, according to some others, it is the ability to learn, and according to some, it is the ability of abstract thinking (Flanagan & Kaufman, 2004). Starting from this, Binet (1905) defines intelligence as "the ability of good reasoning, of good judgment, and having a critical mind" (Anastasi, 1988). Since Alfred Binet's successful design of the first intelligence test, which was designed to differentiate children with mental deficiency from children with behavior disorder, psychometric tools have played an important role in the lives of especially the Americans and the Europeans. The tests have been used for various ends, such as recognition, selection, and evaluation. Although the frequently used tests have not been used in order to measure intelligence, they have been used in order to measure structures such as social competence, educative abilities, special skills, etc, which are closely related to intelligence. The forms of intelligence tests are highly different from one another. While a singular question or article type is used in some tests ("Raven's Progressive Matrices" and "The Peabody Picture Vocabulary Test"), in some of them (such as the Wechsler test), different article or question types are used which are designed in order to measure some special purposes in addition to measuring general intelligence. Moreover, some tests consist only of verbal articles while some of them are constructed by using non-verbalized articles. Scores obtained after the application of the intelligence tests are interpreted usually by turning them into standard scores the general average which is 100, and the standard deviation of which is 15.

In 1930s, the Stanford-Binet Intelligence Test was used as individual intelligence test. This test was subject to criticism by many people including Wechsler. One of the most obvious of these critiques indicated that the Stanford-Binet Test was developed for children and that the article choice was done accordingly, and that it was not appropriate for adults (Özgüven, 2004). Wechsler defined intelligence in his way, and initiated designing an intelligence test. Wechsler defined intelligence as "the general capacity of an individual as a whole to act purposefully, to reason, and to affect his environment." The first Wechsler Individual Intelligence Scale was published as a scale prepared for adults in 1939 under the name "Wechsler Bellevue Form 1 (W-B 1), two years after Stanford-Binet's 1937 revision. The scale was prepared in order to apply to individuals who are in the age range of 10 and 60. These years saw the beginning of the Second World War, and they were marked with a high patient population in military hospitals. Diagnosis gained especial important during these years. Due to the high demand during the war years, Wechsler developed a parallel special second set of the "Wechsler Adult Intelligence Test." The two sets were not really that different from each other save for some questions and explanations (Kaufman, Flanagan, Alfonso & Mascolo, 2006). Wechsler Intelligence Scales used today are of three sets:

1. (Wechsler Adult Intelligence Scale) (WAIS)
2. (Wechsler Intelligence Scale for Children) (WISC)
3. (Wechsler Preschool and Primary Scale of Intelligence) (WPPSI)

Within the scope of this study, the revisions of the Wechsler Intelligence Scale for Children from the beginning to the present, how sub-tests have differentiated from its first form to the present, and how the psychometric qualities have been put forth in the revisions were examined. Three revisions were made from the first form to the present. With small changes, a new form called WISC-IV-Integrated, which is quite close to the WISC-IV form, was launched. The WISC-IV-Integrated form was left out of this study.

- 1949 - WISC
- 1974 - WISC-R
- 1991 - WISC-III
- 2003 - WISC-IV

WISC (Wechsler Intelligence Scale for Children) was constructed with the adaptation of the Wechsler-Bellevue Intelligence Scale (1939) and the design of new sub-tests. The form was developed for the 5-15 age range. Sub-tests were addressed in two sections, namely, verbal and performance. Number Sequence and Labyrinths (sometimes Code instead of Labyrinths) sub-sections are auxiliaries. After the Verbal section, the Performance section is applied (Wechsler, 1949). The application of the sub-tests is as follows:

1.General Information	7.Picture Completion
2.General Comprehension	8.Picture Arrengement
3.Arithmetic	9.Block Design
4.Similarities	10.Object Assmebly
5.Vocabulary	11.Coding
6.Digit Span	12.Mazes

Three different score are obtained from the test. These are the Verbal IQ (VIQ), Performance IQ (PIQ), and Full Scale IQ (FSIQ). The WISC was standardizes on a sample of 100 boys and 100 girls from each from the ages five through fifteen. There were a total of 1100 boys and 1100 girls in eleven age groups. In selecting this sample, Areas, Urban-Rural and Parental Occupation criteria were taken into consideration (Wechsler, 1949). For the representative ages for the 5-15 age range, which are 7.5, 10.5 and 13.5, the inter-correlation of the sub-tests were interrogated. For 7.5, the inter-correlations among the 12 sub-tests were between .12 and .55; for 10.5, the inter-correlations were between .23 and .75; and for 13.5, the inter-correlations were between .59 and .89. The correlation between the Verbal IQ and the Full Scale was determined as .90 for 7.5; .93 for 10.5; and .89 for 13.5. Similarly, the correlation between the Performance IQ and the Full Scale was determined as .89 for 7.5; .90 for 10.5; and .87 for 13.5. Likewise, the two halves reliability for the 7.5, 10.5, and 13.5 age groups were checked, and it was determined that the reliability coefficient for the three scores was between .86 and .95.

The first form of the WISC-R (Wechsler Intelligence Scale for Children-Revised) was revised in 1974 as a result of an approximately 25 year-old evaluation. The age range was expanded to 6-16 as opposed to the previously used 5-15. Several changes were made related to the application and scoring ways of the sub-tests (Cronbach, 1990, Savaşır & Şahin, 1995). Number sequence, labyrinths and codes were kept the same. Picture Competion sub-test was completely altered. The changes were thought necessary in order to be attention-grabbing during the application of the test, although the same sub-tests were used.

1.General Information	7.Vocabulary
2.Picture Completion	8.Object Assmebly
3.Similarities	9.General Comprehension
4.Picture Arrangement	10.Coding
5.Arithmetic	11.Digit Span
6.Block Design	12.Mazes

The standardization of the test was made through 2200 children from different parts of the USA, which were divided into 11 age groups between 6 and 16 years and each group was comprised of 100 boys and 100 girls. As for the reliability of the test, split-half reliability and test –retest reliability was calculated. The results of the split-half reliability which was calculated separately for each age range were .94 for verbal IQ, .90 for performance IQ and .96 for total IQ. 303 children at the age groups of 6.6 - 7.6; 10.6 – 11.6; 14.6 – 15.6 were tested for test-retest

reliability and the reliability coefficients were .87-.95 for verbal IQ, .88-.89 for performance IQ and .92-.95 for full scale IQ.

The standard error value for verbal, performance and full scale IQ sections were calculated as 3.60, 4.66 and 3.19 respectively. Inter-correlations between subtests were calculated and the correlation of the whole test ranged from .38 to .74. WISC and WISC-R were applied on 50 children at the age of 6 and the correlation between them was determined as 80 for verbal and performance sections and 82 for full scale IQ section. WİSC-R and WAİS were applied to 40 children at the age of 16 and the correlation between them was determined as .96 for verbal section, .89 for performance section and .95 for full scale IQ section. WISC-R and Stanford Binet Intelligence Scale L-M form were applied to 118 children and correlation coefficients between the subtests ranged from .51 to .77. Just like in WISC, three different scores are obtained from the test. These are Verbal IQ, Performance IQ and Full Scale IQ.

The third revision which Wechsler made in 1991 (WİSC-III) can be applied to an age range from 6 years 0 months to 16 years 11 months just like the previous form (Braden, 1995; Sandoval,1995). The WISC-III promised to improve the WISC-R through contemporary and representative norms, better floors and ceilings for subtests, new artwork and items sensitive to multicultural and gender concerns, and improved clarity of factor structure—all while maintaining "the basic structure and content of the WISC-R (Wechsler, 1974). WISC-III and WISC-R coincide with each other approximately up to 72%. With this revision, items have become more colourful and more sensitive to gender and racial discrimination. New items have been added to subtests, some of the items were revised, and besides, a new subtest has been added (symbol search). Although many of its properties are the same with those of WISC-R, some changes have been made with regard to certain matters such as the management of the test. The test begins with the subtest of picture completion. The verbal section contains 6 tests and the performance section contains 7 tests. Out of these tests, the subtests of number sequence, symbol search and mazes are alternate tests.

Verbal		Performance	
1.	Information	1.	Picture Completion
2.	Smilarities	2.	Coding
3.	Arithmetic	3.	Picture Arrangement
4.	Vocabulary	4.	Block Design
5.	Comprehension	5.	Object Assembly
6.	Digit Span	6.	Symbol Search
		7.	Mazes

The main purpose of handling WISC-III is to reconstitute the norms. Along with that, strengthening the items, improving the content and subtests, re-handling the factor structure are the other purposes. For the age group of 11, the standardization of test was carried out through 2200 children from 31 different regions. Except the subtests of coding and symbol search which can be characterized as speed tests, the coefficient of split-half reliability of other subtests were calculated to be from .66 to .89. When the factor was examined, a new dimension occurred, which is different from WISC-R. The sub-dimensions of Verbal Comprehension (VO), Perceptual Organization (PO), and Freedom from Distractibility (FFD) are the same as the ones in the previous versions (Kaufman, 1979; Sattler, 2001). The new Symbol Search subtest (along with Coding) created a fourth factor, named Processing Speed (PS). Norms and procedures for organizing 12 WISC-III subtests into four index scores (VC, PO, FFD, and PS) were described in the WISC-III manual. The Mazes subtest contributed only to the calculation of IQs as an alternate test; none of the four factors include it.

For the group of age 8, the correlations between subtests were .66 for verbal and performance IQs, .92 for verbal and full scale IQs, , and .90 for performance and full scale IQs). Respectively, convergent and discriminant validity are both evidenced in the relationships between the verbal and performance subtests. Construct-related and criterion-related validity are discusses in terms of the WİSC-R because of the similarities between the WISC-III and the WISC-R and because of abundance of research already accumulated on the WISC-R. Inter validty is specifially discussed in relation to the WISC-III's intercorrelations of the subtests, Verbal and Performance scales and factor-based scales.

Just like the former versions, WISC-IV was revised in 2003 for the age range from 6 years 0 moths to 16 years 11 months. With this revision, full scale IQ and four indexes have started to be calculated instead of the previously-used verbal and numeric scores.

1. Verbal Comprehension (VCI)
2. Perceptual Reasoning) (PRI)
3. Working Memory (WMI)
4. Processing Speed (PSI)

Except for the Object Assembly, Picture Arrangement and Mazes available in the WISC-III, the other 10 subtests have also been included in this revision. These three subtests were excluded from the test in the forth revision. In addition to 10 subtests, 5 new subtests were included into the test and thus, the number of subtests increased to 15 (10 core, 5 Supplement) (Maller, 2004; Thompson, 2004). Picture Arrangement, Object Assembly, and Mazes were deleted, all of which were associated with Wechsler's Performance Scale and measured problem-solving ability. Picture Arrangement and Object Assembly were heavily dependent on bonus points for quick, perfect performance and presumably were eliminated to reduce emphasis on response time. Mazes had notably weak reliability, stability, and validity (Kaufman, 1994). Information and Arithmetic were moved to supplemental status. This change reduces emphasis on school achievement; success on the WISC-IV VCI is less influenced by knowledge of facts than previous Verbal scales on Wechsler's tests, and success on the WISCIV. WMI is minimally influenced by math achievement, relative to previous Wechsler composites that included Arithmetic (e.g., Verbal IQ, Freedom From Distractibility Index). Five new subtests were added (Word Reasoning, Matrix Reasoning, Picture Concepts, Letter-Number Sequencing, and Cancellation), all of which place emphasis on fluid reasoning, working memory, or both, and none of which traces its heritage to Wechsler's original sources Matrix Reasoning and Picture Concepts are good measures of fluid reasoning, an important ability that has consistently been underrepresented on Wechsler's Performance Scale and perceptual indexes. Although the subtests these new tasks replaced (i.e., Picture Arrangement and Object Assembly) also measured problem-solving ability, they did so with too great an emphasis on processing speed, visualization, and crystallized abilities (Maller, 2004; Thompson, 2004).

Subtests are as follows:

VCI subtests		PRI subtests	
1.	Similarities	1.	Block Design
2.	Vocabulary	2.	Picture Concepts
3.	Comprehension	3.	Matrix Reasoning
4.	Information	4.	Picture Completion
5.	Word Reasoning		
WMI subtests		PSI subtests	
1.	Digit Span	1.	Coding
2.	Letter-Number Sequencing	2.	Symbol Search
3.	Arithmetic	3.	Cancellation

As distinguished from other revisions, in WISC–IV, various renewals and updates were made related to topics of improved assessment of Fluid Reasoning, Working Memory, and Processing Speed, Enhanced clinical validity, Decreased emphasis on time with fewer time bonuses, Improved reliabilities and validities, Improved floors and ceilings on all subtests, updated norm groups. census data, Replacement of outdated items, Including a chapter on interpretation, Mazes, Object Assembly, and Picture Arrangement have been dropped, Information,Word, Reasoning, Picture Completion, Arithmetic, and Cancellation are supplemental subtests, Culturally fair, Reduced weight and increased portability.

The WISC–IV sample consisted of 2,200 children between the ages of 6 and 16 years. 200 children were selected for each of the 11 age groups. The sample was stratified on age, sex, parent education level, region, and race/ethnicity. Split-half reliability calculated with regard to subtests changes between .70 - .90. Test-retest reliability was checked through 243 children and 32 days apart and the test-retest reliability of the subtests were found out to range between .72 - .93. The reliability of the WISC-IV is the average internal consistency coefficients are .94 for VCI, .92 for PRI, .92 for WMI, .88 for PSI, and .97 for FS-IQ. Internal consistency values for individual subtests across all ages ranged from .72 for Coding (for ages 6 and 7) to .94 for Vocabulary (for age 15). The median internal consistency values for the individual subtests ranged from .79 (Symbol Search and Cancellation) to .90 (Letter-Number Sequencing). Like other major intelligence batteries, the WISC-IV's total test score (FS-IQ) and lower order composite (indexes) reliabilities are generally high (□□.90□) across the age range whereas its subtest reliabilities are generally medium (.80 to .89). Likewise, test-retest (mean interval □□32 days) reliability coefficients for the FS-IQ and indexes for a sample of 243 children ages 6 to 16 years were high to medium for the five age groups studied. The WISC-IV is a stable instrument with average test-retest coefficients (corrected for variability of the sample) of .93, .89, .86, and .93 for the VCI, PRI, WMI, PSI, and FS-IQ, respectively. One-month practice effects (gains from test to retest) for the WISC-IV indexes and FS-IQ for three separate age groups (i.e., 6 to 7, 8 to 11, and 12 to 16) and the overall sample are reported in the WISC-IV. In general, practice effects are largest for ages 6 to 7 and become smaller with increasing age. Average FS-IQ gains dropped from about 8 points (ages 6 to 7) to 6 points (ages 8 to 11) to 4 points (ages 12 to 16). Certain WISC-IV subtests demonstrated relatively large gains from test to retest. Specifically, for ages 6 to 7, Coding and Symbol Search showed the largest gains; for ages 8 to 16, Picture Completion showed the largest gains.

2. Conclusion

Like every psychometric instrument which raise question marks about their qualities as the time lapses, Wechsler Intelligence Scale for Children also required certain revisions after its initial form as the time passed and new forms were put forth. The major reasons for those revisions include the testing of factor structures, re-determination of norms, detection of failing items and outdated subtests and replacing them with new ones, etc. With each new revision, it was aimed to obtain a more valid and reliable instrument than the previous one. Among the four revisions, more studies were conducted on third and forth revisions and these studies were in more detail. In addition to that, particular importance was given to construct validity and criterion-referenced validity. While reliability was more emphasized in the initial scales, the relationship between the subscales and the total scale was taken into account in the third and forth revisions and as a result, it was reported that reliabilities were more prioritized.

The Wechsler scales continue to be the most widely used intelligence batteries. The concepts, methods, and procedures inherent in the design of the Wechsler scales have been so influential that they have guided most of the test development and research in the field for more than a half century (Flanagan, McGrew, & Ortiz, 2000). Virtually every reviewer of these scales, including those who have voiced significant concerns about them, have acknowledged the monumental impact that they have had on scientific inquiry into the nature of human intelligence and the structure of cognitive abilities. Kaufman's (1993) review of the third edition of the Wechsler Intelligence Scale for Children (WISC-III), "King WISC the Third Assumes the Throne," is a good example of the Wechsler scales' position of authority and dominance in the field (Flanagan et al., 2000). Although the strengths of the Wechsler scales have always outweighed their weaknesses, critics have identified some salient limitations of these instruments; in particular, they lack a contemporary theory and research base (e.g., Braden, 1995; Burns & O'Leary, 2004; Flanagan & Kaufman, 2004; Keith, Fine, Taub, Reynolds, & Kranzler, 2006; Little, 1992; McGrew, 1994; Shaw, Swerdlik, & Laurent, 1993; Sternberg, 1993; Witt & Gresham, 1985). Nevertheless, when viewed from an historical perspective, the importance, influence, and contribution of David Wechsler's scales to the science of intellectual assessment are both obvious and profound.

References

Anastasia, A. (1988). *Psychological Testing (6.th ed.)*. New York: Macmillan.

Bartholomew, J. D. (2005). *Measuring Intelligence: Facts and Fallacies,* USA: Cambridge University Press.

Braden, J. P. (1995). Review of the Wechsler Intelligence Scale for Children-third edition. In J. V. Mitchell (Ed.), *The Tenth Mental Measurements Yearbook* (Vol. 1, pp. 1098-1103). Lincoln, NE: Buros Institute of Mental Measurement.

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

Cronbach, L.J. (1990). *Essentials of Psychological Testing*. New York. Harper&Row Publishers.

Flanagan, D. P., & Kaufman, A. S. (2004). *Essentials of WISC-IV Assessment*. New York: John Wiley.

Flanagan, D. P., McGrew, K. S., & Ortiz, S. O. (2000). *The Wechsler Intelligence Scales And Gf-Gc Theory: A Contemporary Approach To Interpretation*. Boston: Allyn & Bacon.

Kaufman, A. S. (1994). *Intelligent Testing With The WISC-III*. New York: John Wiley.

Kaufman, A. S. (1979). *Intelligent Testing With The WISC-R*. New York: John Wiley.

Kaufman, A.S., Flanagan, D.P., Alfonso, V.C., & Mascolo, J.T. (2006). Test review: Wechsler Intelligence Scale for Children, Fourth Edition (WISC-IV). *Journal of Psychoeducational Assessment, 24(3),* 278-295.

Keith, T. Z., Fine, J. G., Taub, G. E., Reynolds, M. R., & Kranzler, J. H. (2006). Higher Order, Multi-Sample,Confirmatory Factor Analysis of The Wechsler Intelligence Scale For Children–Fourth Edition: What Does İt Measure*? School Psychology Review, 35*, 108-127.

Liebert, M. R., Rita, N., & Kail, V.R. (1986). *Psychology*, USA: Prentice Hall.

Little, S. G. (1992). The WISC-III: Everything Old Is New Again. *School Psychology Quarterly*, *72*, 148-154.

Maller, S. J. (2004). Review of the Wechsler Intelligence Scale for Children – Fourth Edition. *Mental Measurements Yearbook, 16*.

McGrew, K. S. (1994). *Clinical interpretation of the Woodcock-Johnson Tests of Cognitive Ability–Revised*. Boston: Allyn & Bacon.

Özgüven, İ.E. (2004). *Psikolojik Testler.*Ankara: PDREM Yayınları

Özgüven, İ.E. (2007).*Bireyi Tanıma Teknikleri*. Ankara: PDREM Yayınları

Sandoval, J. (1995). Test Review of the Wechsler Intelligence Scale for Children, Third Edition. In J. C. Conoley & J. C. Impara (Eds.), The twelfth mental measurements yearbook [Electronic version].

Sattler, J.M. (2001). Assesment of Children: Cognitive Applications. San Diego,CA: Author.

Savaşır, I., & Şahin, N. (1995). *Wechsler Çocuklar İçin Zeka Ölçeği(WISC-R).*Türk Psikologlar Derneği.

Shaw, S. E., Swerdlik, M. E., & Laurent, J. (1993). Review of the WISC-III [WISC-III Monograph]. *Journal of Psychoeducational Assessment*, 151-160.

Sternberg, R. J. (1993). Rocky's back again: A review of the WISC-III [WISC-III Monograph]. *Journal of Psychoeducational Assessment*, 161-164.

Thompson, B. (2004). Review of the Wechsler Intelligence Scale for Children – Fourth Edition. *The Mental Measurements Yearbook, 16*.

Toker, F. (1968). *Zeka Kuramları*. Ankara. Milli Eğitim Bakanlığı Talim ve Terbiye Dairesi Araştırma ve Değerlendirme Bürosu.

Tyler, E.L. (1965). *The Psychology of Human Differences.* The Century Psychology Series. New York: Meredidth Publishing Company.

Walsh W. B., & Betz, N. E. (1995). *Test and Assessment.* New Jersey. Prentice-Hall, Englewood Cliffs.

Wechsler, D. (1949). *Wechsler Intelligence Scale for Children*. New York: The Psychological Corporation.

Wechsler, D. (1974). *Wechsler intelligence scale for children-Revised edition*. San Antonio, TX: The Psychological Corporation.

Witt, J. C., & Gresham, F. M. (1985). Review of the Wechsler Intelligence Scale for Children–Third Edition. In J. V. Mitchell (Ed.), *The tenth mental measurements yearbook* (Vol. 2, pp. 1716-1719). Lincoln, NE: Buros Institute of Mental Measurement.