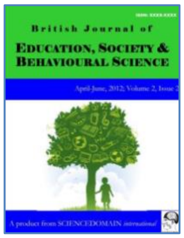


*British Journal of Education, Society & Behavioural Science*
*4(7): 869-882, 2014*

**SCIENCEDOMAIN** *international*
*www.sciencedomain.org*

# Psychometric Properties of the Intelligence Scale for Chinese Adults

**Xi-rong Tang[1], Li Jiang[1], Hui Zeng[1], Xiong-zhao Zhu[1], Jin-yao Yi[1] and Shu-qiao Yao[1*]**

[1]*Medical Psychological Institute, Second Xiangya Hospital, Central South University, #139 Ren-Min Road, Changsha, 410011, China*.

***Authors' contributions***

*This work was carried out in collaboration between all authors. Authors SY and XZ designed the study. Authors LJ and HZ were responsible for sampling. Authors XT and JY performed the statistical analysis, wrote the first draft of the manuscript and corrected it. All authors read and approved the final manuscript.*

***Original Research Article***

***Received 5th December 2013***
***Accepted 23rd February 2014***
***Published 19th March 2014***

## ABSTRACT

**Aims:** The Intelligence Scale for Chinese Adults (ISCA) has been recently developed to address the pressing need from assessment of adult intelligence in China. The purpose of this study is to test its psychometric properties in order to facilitate its utilization.
**Study Design:** This is an analytic study.
**Place and Duration of the Study**: The data of study were collected from 22 provinces of China within 8 months.
**Methodology:** The standardization sample of the ISCA (*n*=2035) was adopted to validate the construct validity and to estimate the internal consistency reliability. A sample of 101 participants tested the scale twice with an interval of 14-28 days to estimate test-retest reliability. A sample of 55 adults tested both the ISCA and the Chinese version of WAIS- III for Taiwan (WAIS- III-CT) in counterbalance order to estimate the concurrent validity.
**Results:** Age-group showed significant effects on raw scores for every subtest ($\eta^2$: 0.082-0.665). Educational levels had statistically significant influence on the age-corrected scaled scores, intelligence indices and full scale intelligence quotient. Effect sizes of educational levels on scaled scores of the subtests in the Crystallized Subscale ($\eta^2$: 0.144-0.202) were obviously larger than those in the Fluid Subscale ($\eta^2$: 0.054-0.128). Results

**Corresponding author: Email: shuqiaoyao@163.com;*

from a series of confirmatory factor analyses indicated that every subtest was unidimensional and the hierarchical structure of the full scale was acceptable. The internal consistency reliability of each core subtest was above 0.90, so was that of the full scale. The test-retest reliability of the full scale intelligence quotient was 0.94. Performance on the ISCA was correlated significantly with that on the WAIS- III-CT (*r*=0.90).
**Conclusions:** The ISCA has satisfactory levels of validity. The full scale and its subtests have sufficient internal consistence reliability and test-rest reliability. The ISCA is both valid and reliable as an instrument for the assessment of adult intelligence in China.

*Keywords: Intelligence scale; standardization sample; reliability; validity.*

## 1. INTRODUCTION

Adult intelligence scales are widely used in the medical and psychological practice, such as for diagnoses of dementia and researches regarding the mild cognitive impairment, although the concepts of intelligence and the intelligence quotient remain controversial [1]. In the United States, several new scales were developed while traditional major intelligence scales were revised in the last two decades [2-7]. However, the situation is distinct in mainland China, where there was no intelligence test available when the restoration of mental measurement was initiated in 1979 [8]. During the early 1980s, Yao-xian Gong revised the Wechsler adult intelligence scale (WAIS) [9] into the WAIS-revised-in-China (WAIS-RC) [10]. As the only adult intelligence battery available in China for more than two decades, the WAIS-RC was popularized and utilized widely in both medical and psychological practice and researches [11]. According to a survey by Dai et al. it used to be the most frequently used psychological test in China [12].

Notwithstanding, the need to replace the WAIS-RC has become increasingly obvious for three reasons: a) some items in the WAIS-RC have become dated as a result of social change. For example, an item regarding the abacus in the Picture Completion has become dated , as the abacus, which used to be a commonly used instrument in China, has almost disappeared from daily life after the arrival of the computer age; b) as it was revised from the WAIS, which was released in 1955, the WAIS-RC has not include the progress in the fields of cognitive psychology and intelligent assessments since that date and c) the discovery and establishment of the Flynn effect, which describes a pervasive increase over time in the performance on intelligence quotient (IQ) test, suggests a risk of misleading from utilization of an obsolete norm in research and clinical practice, in addition to the major problem of causal explanation [13,14]. An instrument for adult intelligence assessments with updated norms, but without outdated items, would be advantageous in China, a country that has a population of more than 1.3 billion and rapidly developing psychology [15,16].

With an appreciation of the urgent need for a new intelligence scale to assess adult intelligence in China, Yao et al. constructed a new intelligence scale, namely the Intelligence Scale for Chinese Adults (ISCA) in mid-2000 [17]. This scale taps the general cognitive and adaptive abilities of adults with a hierarchical structure, which consists of a full scale, three subscales and 12 subtests to measure general ability, some broad abilities and narrow abilities, respectively. This structure resembles the three strata of the Cattell-Horn-Carroll model, which synthesizes the Cattell and Horn's Gf-Gc Model and the Carroll's Three-Stratum Hierarchy model [18,19]. In a preliminary study, the preparatory version of the scale was administered to 400 adults stratified according to demographic variables. The results of the study suggested that the factorial structure consisted of three correlated factors (the Crystallized intelligence, Fluid intelligence, Attention and memory ability) fit the data well [17]. The Cronbach's alpha of the full scale was 0.84. The test-retest correlation of the full scale IQ was 0.97. The standardization of the scale was accomplished in 2006 [20,21].

Validity and reliability are indispensable psychometric properties of a measurement instrument and they underlie test interpretation. The purpose of the present study is to test validity and reliability of the ISCA, especially those based on the standardization data, to facilitate utilization of the battery in medical and psychological practice. Concretely, this study will investigate: a) influence of age and education on individuals' performance which can be regarded as evidence for validity of intelligence and cognitive tests, b) appropriateness of the factorial structure of the ISCA and its subtests, c) the internal consistency reliability, test-retest reliability and concurrent validity of the scale and its components.

## 2. METHODOLOGY

### 2.1 Participants

The standardization sample for the ISCA consists of 2,035 individuals aged 41.4±20.7 years old (range, 16-92) with educational attainment of 8.9±4.00 years (range, 0-19 yr) who were recruited as a nationally representative sample of individuals aged 16 years old or older. Other demographic characteristics of the sample are presented in Table 1. A retest sample of 101 individuals aged 53.5±21.9 years old (range, 16-83 yr) with educational attainment of 8.8±5.7 years (range, 0-19 yr) tested the ISCA twice. A validation sample of 55 adults with age of 32.8±8.7 years old (range, 20-55 yr) and educational attainment of 9.7±2.7 years (range, 6-16 yr) tested both ISCA and the Chinese version WAIS-III for Taiwan (WAIS-III-CT) [22]. The demographic characteristics of the retest sample and validation sample are also presented in Table 1.

**Table 1. Demographic characteristics of participants**

| Demographic Variables | Standardization sample(*N*=2,035) | | Retest sample (*N*=101) | | Validation sample (*N*=55) | |
|---|---|---|---|---|---|---|
| | Counts | Percents | Counts | Percents | Counts | Percents |
| Gender | | | | | | |
| Male | 1008 | 49.5 | 58 | 57.4 | 19 | 35 |
| Female | 1027 | 50.5 | 43 | 42.6 | 36 | 65 |
| Schooling Years | | | | | | |
| ≤ 3 years | 219 | 10.8 | 10 | 9.9 | | |
| 4-6 years | 311 | 15.3 | 30 | 29.7 | 7 | 13 |
| 7-9 years | 626 | 30.8 | 12 | 11.9 | 26 | 47 |
| 10-12 years | 611 | 30.0 | 19 | 18.8 | 15 | 27 |
| ＞12 years | 268 | 13.2 | 30 | 29.7 | 7 | 13 |

### 2.2 Procedure

The standardization sample was stratified in three stages according to age, gender and educational level. The stratified sampling plan was made in line with the fifth Chinese census data [23]. According to the plan, 100 males and 100 females should be sampled for each of the following age-groups: 16-17 yr, 18-19 yr, 20-24 yr, 25-29 yr, 30-34 yr, 35-44 yr, 45-54 yr and 55-64 yr, 150 individuals should be sampled in each age-group: 65-69 yr, 70-74 yr, 75-79 yr and ≥80 yr, among which, females should be more than males to correspond with the ratio in the general population. Educational level was defined by the years of schooling completed, as follows: ≤3 years (illiterate in China), 4-6 years (primary education), 7-9 years (secondary education), 10-12 years (high school), >12 years (advanced education). The proportion of each educational level in every age-group was in line with the census data. Although the protocol have been implemented successfully in adults aged 16-64 yr, the sampling plan has not been implemented completely in those aged >65 yr.

Sampling plan was executed simultaneously in 22 provinces distributed over seven geographic regions of the mainland China, namely the Northeast, North, Northwest, East, South, Southwest and Central regions. The proportion of samples from each region was set to correspond approximately with the population ratio in each region. All the examiners were trained and qualified to administer the battery. Participants were recruited from the community. All could communicate in the Mandarin Chinese. Most of participants were tested in clinical psychological department of medical college or hospital. A brief structural clinical interview was carried out before administration of the battery to exclude individuals with physical or mental conditions that affected their test performance, including hearing or visual impairments, affective disturbance, psychoactive substance addiction or treatment with psychotropic substances.

Participants in the retest sample tested the ISCA twice with an interval of 14 through 28 days (median: 18 days). 55 participants of the validation sample were administered both the ISCA and WAIS-III-CT in counterbalance order.

All samples were collected during November 2005 through June 2006. Participants were all well informed and signed consent before administration of the battery. Each participant received incentive equivalent to half a day's wage.

## 2.3 Measurements

### 2.3.1 ISCA

The ISCA is an individually administered intelligence battery that encompasses three subscales and 12 subsets [17]. The Crystallized Intelligence Scale (CIS) consists of three core subtests, namely the Vocabulary (VOC), Similarity (SIM) and Information (Info), in addition to a supplemental subtest, namely the Life Wisdom (LW). The Vocabulary contains 24 words of which oral definitions are demanded. This subtest measures word knowledge and expressive ability, and requires long-term memory also. The Similarities contains 15 pairs of words. Examinees are required to explain how they are alike. This subtest measures abstract verbal thinking and the ability to distinguish between non-essential and essential features. The Information encompasses 21 questions designed to test factual knowledge and long-term memory on a broad range of general knowledge. The Life Wisdom contains 14 questions about problems from daily life or social practices to measure practical knowledge, social judgment, common sense, long-term memory and verbal expression. All items within the subtests are untimed and each item yields 0-2 points according to accuracy and integrality of the answers.

The Fluid Intelligence Subscale (FIS) contains four core subtests, namely the Picture Completion (PC), Block Design (BD), Figural Reasoning (FR) and Picture Arrangement (PA), in addition to a supplemental subtest, namely the Object Assembly (OA). The Picture Completion contains 20 pictures with a missing part in each one. Examinees are required to identify the missing parts. The subtest measures visual perception and attention to visual detail. The Block Design contains 12 pictured designs. Examinees are required to reproduce the designs with special designed blocks. It measures visual perception and organization, and visual-motor coordination. The Figural Reasoning contains 16 figural matrices of which a missing section is required to identify from the response choices. It measures perceptual organization and non-verbal reasoning. The Picture Arrangement contains 10 items. Each item consists of a set of picture cards presented in error sequence. Examinees are required to rearrange the cards in proper sequence. The subtest measures perceptual organization and simultaneous processing. The Object Assembly contains five object assembly puzzles designed to measure visuospatial ability, analysis and synthesis and visual-motor coordination. All items in this subscale were timed and additional time bonuses were awarded to half items in Block Design and all items in Object Assembly.

The Attention and Memory Subscale (AMS) consists of three subtests, namely the Coding (COD), Arithmetic (AR) and Digit Span (DS). The Coding requires examinees to copy simple symbols paired with number 1-9 respectively as quickly as possible within a 120-second time limit. It measures visuomotor processing speed, short-term visual memory and concentration. The Arithmetic contains 17 arithmetical problems which require examinees to solve mentally within a given time limit. The subtest measures working memory and numerical reasoning. The Digit Span required examinees to reproduce several series of number sequences (2 digits to 12 digits) forward or backward. It measure memory, attention and auditory processing.

Each subtest yields an age-corrected scaled score with mean of 10 and standard deviation of three in addition to the raw score. The sum of scaled scores for the core subtests subsumed under the subscale is used to derive the corresponding index score, i.e., the Crystallized Intelligence Index (CII), Fluid Intelligence Index (FII) and Attention and Memory Index (AMI). The full scale intelligence quotient (FSIQ) is derived from the sum of scaled scores for ten core subtests. All of the composite scores, including the three index scores and the FSIQ, are scaled to a metric with mean of 100 and standard deviation of 15. A supplemental subtest is adopted to compute the intelligence index and quotient when a core subtest is either not suitable to the examinee or spoiled during its administration.

It will take about 80 minutes to administer the battery successively as follows: the Picture Completion, Vocabulary, Coding, Similarity, Block Design, Arithmetic, Figural Reasoning, Digit Span, Information, Life Wisdom and Picture Arrangement. Administration of each subtest will discontinue when the examinee meet criterion for discontinuation of the subtest.

#### 2.3.2 WAIS-III-CT

The WAIS is a well-known intelligence battery. The WAIS-III was released in 1997 [2,3]. The Chinese version of the WAIS-III for Taiwan (WAIS-III-CT) was standardized and released in Taiwan in 2002 [22]. The four-factor structure of the WAIS-III-CT has been evaluated in mainland China [24].

## 2.4 Data Analyses

Performance of the standardization sample on the ISCA was described and the effects of age group and educational level on performance were explored using the SPSS v15.0 (SPSS Inc; 2006). To explore the effect of age-group, analyses of variance were implemented with dependent variable of raw score for each subtest and fixed factor of age-group. In analyses of variance to explore the effects of educational level, age-corrected scaled scores served as dependent variables and educational level served as fixed factor. The effect sizes of demographic variables were estimated by the eta squared values ($\eta^2$).

To validate the structure of the scale and its subtests, a series of confirmatory factor analyses (CFAs) with the standardization sample data were conducted using the EQS 6.1 program [25]. Firstly, the unidimensionality of each subtest was tested seriatim with the exception of the Coding and Digit Span. With the exception of Object Assembly, one-factor model was confirmed for every subtest based on the polychoric correlation matrix of its items with difficulty of 0.01-0.99, and correspondingly the robust maximum likelihood method (RML) was adopted which was appropriate for categorical data [26]. As for the Object Assembly, in which all items were awarded five or more points and scores of all the five items followed the multivariate normal distribution, so the CFA was executed based on covariance matrix of its items and the maximum likelihood estimator (ML) was adopted. Secondly the hierarchical factor structure of the scale (see Fig. 1) was confirmed, in which scaled scores of the subtests were introduced as interval indicators and the ML was adopted.

In all the CFAs mentioned above, correlation between every pair of residuals was set as zero and technical parameters were set as the default. To judge models, chi-square (while ML adopted) or Satorra-Bentler scaled chi-square (while RML adopted) were computed and several fit indices frequently recommended were employed[26-28], including the comparative fit index (*CFI*), Bentler-Bonett normed fit index (*NFI*) and root mean square error of approximation (*RMSEA*). The *CFI* and *NFI* ranges between 0 and 1, with larger values indicating better fit. The *RMSEA* has a minimum value of 0, with smaller values indicating a better fit. According to suggestions of some authors, *CFI* values >0.95, *NFI* >0.97, *RMSEA* <0.05 indicate a good fit, while *CFI* >0.90, *NFI* >0.95, *RMSEA* <0.08 indicate an acceptable fit [28].

To check the internal consistency of the scale and its subtests, during the CFAs procedure mentioned above, the Cronbach's alpha and the reliability coefficient rho were computed. The alpha coefficient has been well documented and frequently used to valuate internal consistency [29,30]. The reliability coefficient rho is computed based on the latent factor model being tested [26,31]. When the model is a one-factor model, the reliability coefficient rho is the same as McDonald's omega coefficient ($\omega_h$), which is an indicator of how well a test measures a single construct [32].

To explore stability of the scale, the Spearman's rho for 101 pairs of CIIs, FIIs, MAIs and FSIQs in the retested sample were computed. To validate concurrent validity of the ISCA, the Spearman's rho for 55 pairs of FSIQs (for the ISCA and WAIS-III-CT, respectively) in the validation sample was computed.

**Table 2. Mean raw scores of subtests by age group**

| | **16-17 yr (*n*=200)** | **18-19 yr (*n*=200)** | **20-24 yr (*n*=200)** | **25-29 yr (*n*=200)** | **30-34 yr (*n*=200)** | **35-44 yr (*n*=200)** | **45-54 yr (*n*=200)** | **55-64 yr (*n*=200)** | **65-69 yr (*n*=170)** | **70-74 yr (*n*=120)** | **75-79 yr (*n*=95)** | **80- yr (*n*=50)** | **$F(11,2023)^a$** | **$\eta^2$** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| VOC | 30.68 | 30.63 | 29.51 | 29.32 | 29.41 | 28.88 | 25.62 | 24.49 | 22.37 | 20.68 | 18.68 | 15.88 | 36.61 | .166 |
| SIM | 20.96 | 21.23 | 21.00 | 20.99 | 19.98 | 20.16 | 18.21 | 16.63 | 15.51 | 14.63 | 14.13 | 11.66 | 44.47 | .195 |
| Info | 30.95 | 31.25 | 30.69 | 30.11 | 29.26 | 29.88 | 26.64 | 25.46 | 23.94 | 22.78 | 23.01 | 18.16 | 33.25 | .153 |
| LW | 16.53 | 16.95 | 16.83 | 17.30 | 16.58 | 16.80 | 15.65 | 15.14 | 14.58 | 13.33 | 13.57 | 11.04 | 16.46 | .082 |
| PC | 13.96 | 13.94 | 13.47 | 13.38 | 13.02 | 12.89 | 11.44 | 10.29 | 10.11 | 9.05 | 8.94 | 7.34 | 59.86 | .246 |
| BD | 26.78 | 27.11 | 26.12 | 25.05 | 23.97 | 22.75 | 19.22 | 17.47 | 15.74 | 14.32 | 13.92 | 11.94 | 87.89 | .323 |
| FR | 10.49 | 10.58 | 10.20 | 9.91 | 9.54 | 8.86 | 7.25 | 5.94 | 5.12 | 4.54 | 4.32 | 3.30 | 114.28 | .383 |
| PA | 14.20 | 14.10 | 13.56 | 13.44 | 12.78 | 12.12 | 10.50 | 9.27 | 7.80 | 6.39 | 6.25 | 3.48 | 114.92 | .385 |
| OA | 28.47 | 29.31 | 28.25 | 26.87 | 26.33 | 25.99 | 22.49 | 20.05 | 17.41 | 15.74 | 14.91 | 12.46 | 47.53 | .205 |
| COD | 83.73 | 81.20 | 79.46 | 73.96 | 70.09 | 62.83 | 48.22 | 38.77 | 29.06 | 22.82 | 21.71 | 14.74 | 364.53 | .665 |
| AR | 13.60 | 13.61 | 13.31 | 13.31 | 12.93 | 12.74 | 11.49 | 11.19 | 10.05 | 9.69 | 9.72 | 8.58 | 51.46 | .219 |
| DS | 22.88 | 22.52 | 22.03 | 20.16 | 19.99 | 19.17 | 17.07 | 15.43 | 14.53 | 13.31 | 13.77 | 12.18 | 104.54 | .362 |

*VOC=Vocabulary; SIM=Similarity; Info=Information; LW=Life Wisdom; PC=Picture Complement; BD=Block Design; FR=Figure Reason; PA=Picture Arrangement; OA=Object Assembly; COD=Coding; AR=Arithmetic; DS=Digit Span.*

*Note a. all $P < 0.01$*

## 3. RESULTS AND DISCUSSION

### 3.1 Performance of the Standardization Sample on the ISCA

The mean raw scores for 12 subtests of each age-group of the standardization sample are presented in above Table 2. All apexes of the mean raw scores appeared in the 16-24 age-band and decreased from the age of 25-29 yr, as described by David Wechsler [33]. There was a statistically significant difference ($p < 0.01$) among the 12 age-groups in every subtest. The effect sizes of the age-group ranged from 0.082 to 0.665. The mean scores for Digit Span, Coding and subtests in the Fluid Intelligence Subscale (FIS) decreased more obviously and earlier than those for subtests in the Crystallized Intelligence Subscale (CIS). And the age-group showed relatively great effects on subtests subsumed under the FIS. These manifestations were accordant with the theory regarding crystallized intelligence and fluid intelligence proposed by John Horn [34].Scaled scores of all subtests, from which the effect of age was eliminated, increased with years of educational attainment (Table 3), especially those for the Vocabulary, Similarity, Information and Coding. Correspondingly, the three intelligence indices and FSIQ also increased with educational level. The difference among five educational levels in the scaled scores, intelligence indices and FSIQ were all significant ($p < 0.01$). The effect sizes of educational level ranged from 0.054 to 0.259 (Table 3). Among them, eta squared values for scaled scores of subtests in the CIS, the Crystallized Intelligence Index and FSIQ were relatively large. Meanwhile, educational levels had only medium or low effect sizes on scaled scores for subtests in the FIS. This observation mirrored the consensus regarding the relationships among intelligence, crystallized intelligence, fluid intelligence and education [35]. All these findings were evidence for the validity of ISCA as intelligence scale.

**Table 3. Mean Scaled Scores of Subtests, Intelligence Indices and FSIQ by Educational Level**

| | ≤3 yrs (*n*=219) | 4-6 yrs (*n*=311) | 7-9 yrs (*n*=626) | 10-12yrs (*n*=611) | ＞12 yrs (*n*=268) | $F(4,2030)^a$ | $\eta^2$ |
|---|---|---|---|---|---|---|---|
| VOC | 7.89 | 8.67 | 9.62 | 10.83 | 12.35 | 125.17 | .198 |
| SIM | 8.12 | 8.66 | 9.42 | 10.85 | 12.37 | 119.34 | .190 |
| Info | 7.94 | 8.63 | 9.52 | 10.94 | 12.37 | 128.23 | .202 |
| LW | 8.21 | 9.00 | 9.61 | 10.61 | 12.13 | 85.40 | .144 |
| PC | 8.45 | 8.78 | 9.64 | 10.70 | 11.88 | 74.68 | .128 |
| BD | 8.64 | 8.90 | 9.62 | 10.65 | 11.80 | 64.44 | .113 |
| FR | 8.54 | 8.79 | 9.42 | 10.51 | 11.84 | 68.77 | .119 |
| PA | 8.65 | 9.02 | 9.71 | 10.67 | 11.60 | 55.68 | .099 |
| OA | 9.14 | 9.08 | 9.74 | 10.53 | 11.13 | 29.14 | .054 |
| COD | 7.95 | 8.76 | 9.62 | 10.83 | 12.00 | 98.24 | .162 |
| AR | 8.09 | 8.87 | 9.56 | 10.60 | 12.14 | 92.22 | .154 |
| DS | 8.41 | 9.12 | 9.59 | 10.68 | 11.62 | 58.57 | .103 |
| CII | 88.64 | 92.26 | 97.09 | 104.72 | 113.57 | 172.58 | .254 |
| FII | 91.25 | 93.22 | 97.65 | 104.05 | 111.19 | 106.25 | .173 |
| AMI | 88.48 | 93.28 | 97.68 | 104.57 | 112.16 | 143.51 | .220 |
| FSIQ | 88.33 | 92.09 | 97.17 | 104.92 | 113.71 | 177.61 | .259 |

*VOC=Vocabulary; SIM=Similarity; Info=Information; LW=Life Wisdom; PC=Picture Complement; BD=Block Design; FR=Figure Reason; PA=Picture Arrangement; OA=Object Assembly; COD=Coding; AR=Arithmetic; DS=Digit Span; CII= Crystallized Intelligence Index; FII= Fluid Intelligence Index; AMI= Attention and Memory Ability Index; FSIQ=full scale intelligence quotient Note a. all $p < 0.01$*

## 3.2 Construct Validity of the Scale and Its Components

In all the CFAs of subtests, estimates of free parameters were all substantial and significant and there were no improper solutions. The chi-squares and fit indices are reported in Table 4. Chi-squares of the models were all statistically significant. Despite that, fit indices indicated an excellent fit between the data and model for every subtest except the Object Assembly. These results suggested that the one-factor model fitted the Data well. In other words, it was acceptable that items in each subtest were homogenous. As for the Object Assembly, the *NFI* and *CFI* suggested a good fit, while the *RMSEA* was 0.085 with 90% confidence intervals of 0.069-0.102, a range overlapping reasonable fit, mediocre fit and poor fit [36]. As a whole, the one factor model for the Object Assembly should be accepted. Even if the model was refused, validity of most examinees' measurement would not be compromised since the Object Assembly is a supplemental subtest and therefore most examinees won't test it.

**Table 4. CFAs and internal consistence of the full scale and its subtests**

| | *k* | $\chi^{2\,a}$ | *df* | $\chi^2/df$ | *NFI* | *CFI* | *RMSEA* | *90% CI* | *Alpha* | *Rho* |
|---|---|---|---|---|---|---|---|---|---|---|
| VOC | 24 | 682.76 | 276 | 2.47 | .983 | .989 | .029 | .026-.032 | .935 | .937 |
| SIM | 15 | 576.82 | 90 | 6.41 | .971 | .975 | .052 | .048-.056 | .932 | .934 |
| Info | 16 | 266.10 | 104 | 2.56 | .988 | .993 | .028 | .024-.032 | .937 | .939 |
| LW | 12 | 127.35 | 54 | 2.36 | .980 | .988 | .026 | .020-.032 | .836 | .838 |
| PC | 17 | 273.32 | 119 | 2.30 | .980 | .988 | .025 | .021-.029 | .926 | .911 |
| BD | 8 | 259.18 | 27 | 9.60 | .978 | .980 | .065 | .058-.072 | .934 | .931 |
| FR | 15 | 187.63 | 90 | 2.08 | .990 | .995 | .023 | .018-.028 | .943 | .943 |
| PA | 10 | 65.56 | 35 | 1.87 | .991 | .996 | .021 | .013-.028 | .916 | .918 |
| OA | 5 | 78.68 | 5 | 15.74 | .961 | .963 | .085 | .069-.102 | .713 | .736 |
| AR; | 12 | 275.48 | 65 | 4.24 | .984 | .988 | .040 | .035-.045 | .960 | .960 |
| FS 10 | 10 | 225.01 | 32 | 7.03 | .978 | .981 | .054 | .048-.061 | .909 | .916 |
| FS 12 | 12 | 368.05 | 51 | 7.22 | .971 | .975 | .055 | .050-.061 | .920 | .928 |

*k= number of items or subtests in CFA; VOC=Vocabulary; SIM=Similarity; Info=Information; LW=Life Wisdom; PC=Picture Complement; BD=Block Design; FR=Figure Reason; PA=Picture Arrangement; OA=Object Assembly; COD=Coding; AR=Arithmetic; DS=Digit SpanPC=Picture Complement PA=Picture Arrangement; OA=Object Assembly; FS 10=the full scale of 10 core subtests; FS 12=the full scale of 10 core subtests and 2 supplemental subtests; NFI=Bentler-Bonett Normed fit index; CFI=Comparative fit index; RMSEA=Root mean square error of approximation; 90%CI=90% confidence interval of RMSEA.*

Note a. In CFA of Object Assembly, subscales and the full scale, Chi-squares were reported in column below $\chi^2$, otherwise Satorra-Bentler scaled chi-squares were reported. All the $\chi^2$ were statistically significant.

The standardized solutions of the hierarchical factor model are shown in Fig. 1. The loading of subtests on the first-order intelligence factor were 0.60 through 0.83 and loadings of the first-order factors on the general intelligence factor were all more than 0.90. Fit indices shown in Table 4 suggested a good fit between the model and empirical data, with either 10 core subtests or with all 12 subtests. These findings suggested that description and interpretation of performance of examinees on the ISCA in three strata, namely the scaled scores of the subtests, intelligence index scores of the subscales and the intelligence quotient of the full scale, were rational and acceptable.

## 3.3 Internal Consistency of the ISCA

The Cronbach's alpha and reliability coefficient rho for each subtest and the full scale were equal or very close, respectively (Table 4). The two coefficients of eight core subtests were all >0.90, whereas those of two supplemental subtests were slightly lower, 0.713 and 0.736 for Object Assembly, respectively. The alpha and rho for the full scale, whether for ten core subtests or all 12 subtests, were both more than 0.90 (Table 4). The two indicators for the Crystallized Intelligence Subscale were 0.85 and 0.85, the Fluid Intelligence Subscale 0.82 and 0.83, the Attention and Memory Subscale 0.69 and 0.70.

Cronbach's alpha has been a widely used estimator of the reliability of tests [30,37]. The reliability coefficient rho was deemed to be superior to Cronbach's alpha [38] and was proposed to improve the routine reporting of psychometric internal consistency [39]. Therefore, values of the two coefficients for the composite reliability of subtests, subscales and the full scale have been reported, although they were either equal to or very close to their counterparts. Values of the two reliability coefficients for eight core subtests were all above 0.90. These values suggested that test scores of these subtests were sufficiently reliable [30]. Since the reliability coefficient rho of each subtest was estimated based on the one-factor model in CFA of each subtest, it was the same as McDonald's omega coefficient and could serve as a measure of homogeneity [25]. Therefore, the values of the reliability coefficient rho also indicated that items in each subtest were highly homogeneous. Values for the alpha and rho of the two supplemental subtests were slightly lower, but they were acceptable (>0.70) [37].

The Cronbach's alpha and reliability coefficient rho of the three subscales indicated that subtests in each subscale were substantively homogeneous and summary of subtests score were rational, interpretable and reliable [30*]. It also should be mentioned that values of the* indicators for the subtest were less than those for core subtests. This suggested that each subtest subsumed under a subscale owned specific functioning and therefore was irreplaceable.

The Cronbach's alpha and reliability coefficient rho of the full scale, whether ten core subtests or all 12 subtests were included, were both more than 0.90. This indicated that the performance of the ISCA was a sufficiently reliable base for decisions about individuals [30].

## 3.4 Stability of the ISCA

The test-retest correlation coefficients for the CII, FII, AMI and FSIQ were 0.87, 0.85, 0.85 and 0.94. They were all significant ($p < 0.01$) and substantive. The test-retest reliability of the FSIQ indicated that ISCA was reliable enough to evaluate individuals [30].

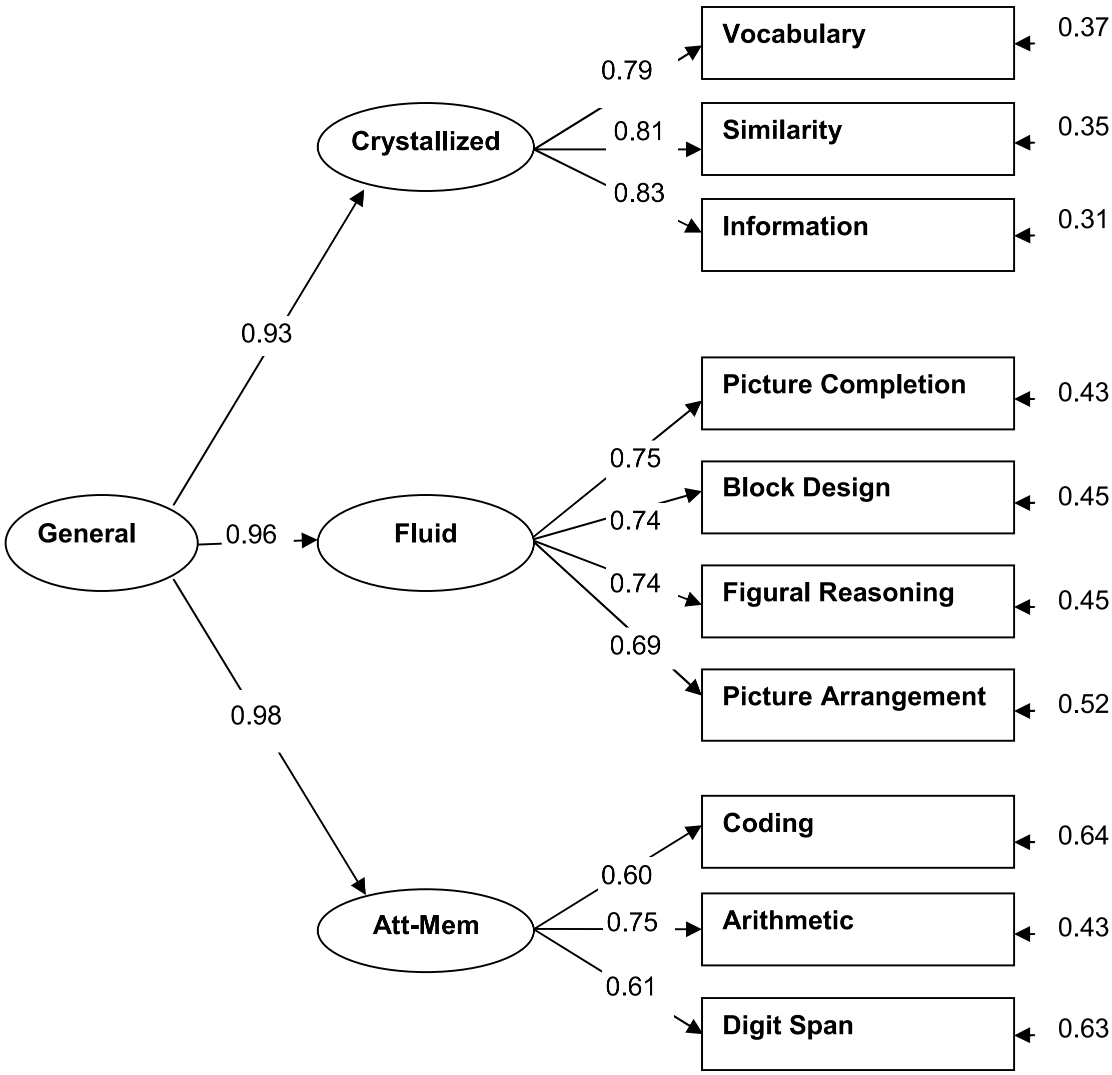


**Fig. 1. Hierarchical Factor Structure of ISCA ( 10 core subtest )**

*General = General intelligence factor; Crystallized = Crystallized intelligence factor; Fluid = Fluid intelligence factor; Att-Mem = Attention and memory factor*

## 3.5 Concurrent validity of the ISCA

Correlation coefficient between two series of FSIQs ($n$ = 55) yielded from the ISCA and WAIS-III-TC was 0.90 ($p$ < 0.01). WAIS is well-known and widely used intelligence scale, and was regarded as "synonymous with adult intelligence" [40] and gold standard for intelligence measurement [41]. The significant and substantive relationship suggested that the ISCA was a valid instrument to assess adult intelligence. Yet as the scale was developed to assess adults form varied subpopulations, especially those with neuro-cognitive disorder, further validation with a larger and more representative sample is still wanted to provide information for decision of consequence.

## 4. CONCLUSION

The ISCA was an intelligence battery recently developed in China to address the pressing need for an assessment tool in medical and psychological practices. Performance of participants on the battery revealed effect of age-group and educational level and proved the ISCA valid. CFAs based on the standardization data of ISCA indicated that the scale had satisfactory construct validity since all subtests were unidimensional and the hierarchical factor structure of ISCA was acceptable. The ISCA also had excellent concurrent validity. The scale, subscale and its subtests had sufficient levels of internal consistency reliability and test-retest reliability. In conclusion, the ISCA was both valid and reliable as an instrument to assess adult intelligence in China.

## COMPETING INTERESTS

The authors have declared that no competing interests exist.

## REFERENCES

1. Deary IJ. Intelligence. Annu Rev Psychol. 2012;63:453-82.
2. Wechsler D. Wechsler adult intelligence scale–third edition and wechsler memory scale–third edition: administration and scoring manual. San Antonio, TX: Psychological Corporation; 1997.
3. Wechsler D. Wechsler adult intelligence scale–third edition and wechsler memory scale–third edition: technical manual. San Antonio, TX: The Psychological Corporation; 1997.
4. Wechsler D. Wechsler Adult Intelligence Scale–fourth edition. San Antonio, TX: Pearson; 2008.
5. Woodcock RW, McGrew KS, Mather N. Woodcock-Johnson III tests of cognitive abilities. Rolling Meadows, IL: Riverside Pub; 2001.
6. Roid GH. Stanford Binet intelligence scales (5th ed.): technical manual. Itasca, IL: Riverside; 2003.
7. Kaufman AS, Kaufman NL. Manual for the Kaufman adolescent and adult intelligence test (KAIT). Circle Pines, MN: American Guidance Service; 1993.
8. Zhang H: Psychological measurement in China. International Journal of Psychology. 1988;23(1-6):101-117.
9. Wechsler D. Manual for the wechsler adult intelligence scale. San Antonio, TX: Psychological Corp; 1955.
10. Dai X, Gong Y, Zhong L. Factor analysis of the mainland Chinese version of the wechsler adult intelligence scale. Psychological Assessment: A Journal of Consulting and Clinical Psychology. 1990;2(1):31-34.
11. Chen Z, Zhen H. A survey of popularization and application of wechsler adult intelligence scale in China. Chinese Mental Health Journal. 2000;14(04):235. Chinese.

12. Dai X, Zhen L, Ryan JJ, Paolo AM. A survey of psychological tests used in clinical psychological practical of China and its comparison with the data of the United States. Chinese Journal of Clinical Psychology. 1993;1(01):47-50. Chinese.
13. Flynn JR. The mean IQ of Americans: Massive gains 1932 to 1978. Psychological Bulletin. 1984;95(1):29-51.
14. Hiscock M. The Flynn effect and its relevance to neuropsychology. Journal of Clinical and Experimental Neuropsychology. 2007;29(5):514-529.
15. Jing Q. Development of psychology in China. Inter J Psychol. 199429(6):667-675.
16. Higgins LT, Zheng M. An introduction to Chinese psychology: its historical roots until the present day. J Psychol. 2002;136(2):225-239.
17. Yao S, Zhou Y, Jiang L, Xiong Y. Development of the intelligence scale for chinese adults: construction of the scale and tryout. Chinese Journal of Clinical Psychology. 2006;14(5):441-445. Chinese.
18. Schneider WJ, McGrew KS. The Cattell-Horn-Carroll model of intelligence. In: Flanagan DP, Harrison PL. editors. New York Contemporary intellectual assessment: Theories, tests, and issues. 3rd ed. Guilford Press. 2012;99-139.
19. Alfonso VC, Flangan DP, Radwan S. The impact of the Cattell-Horn-Carroll theory on test development and interpretation of cognitive and academic abilities. In: Flanagan DP, Harrison PL. editors. Contemporary Intellectual Assessment: Theories, Tests, and Issues. 2nd ed. New York. Guilford Press. 2005:185-202.
20. Yao S, Jiang L, Zhou Y, Xiong Y, Chen H, Tang X. The intelligence scale for chinese adults: development of national norms with urban samples. Chinese Mental Health Journal. 2007;21(1):32-35. Chinese.
21. Zeng H, Yao S, Jiang L, Tang X. The intelligence scale for chinese adults: reliability, validity and national norms for the elderly. Chinese Journal of Clinical Psychology. 2009;17(5):521-525. Chinese.
22. Chen Y, Chen H. Wechsler Adult intelligence scale-third edition (WAIS-III) manual for Taiwan. Taipei, Taiwan: The Chinese Behavioral Science Corporation; 2002. Chinese.
23. The National Census Office. Chinese Census Data of 2000. Beijing, China: Chinese Statistics Press. 2002;537-578, 603-621.Chinese
24. Yao S, Chen H, Jiang L, Tam WC. Replication of factor structure of wechsler adult intelligence scale-iii chinese version in chinese mainland non-clinical and schizophrenia samples. Psychiatry and Clinical Neuroscience. 2007;61(4):379-84.
25. Bentler PM. EQS 6 structural equations program manual. Encino, CA: Multivariate Software, Inc; 2006.
26. Guo B, Aveyard P, Dai X. The chinese intelligence scale for young children：testing factor structure and measurement invariance using the framework of the wechsler intelligence tests. Educational and Psychological Measurement. 2009;69(3):459-474.
27. Hu L. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling. 1999;6(1):1-55.
28. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Research Online. 2003;8(2):23-74.
29. Cronbach L. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334.
30. Webb NM, Shavelson RJ, Haertel EH. Reliability coefficients and generalizability theory. In: Rao CR, Sinharay S. editors. handbook of statistics, Psychometrics. Amsterdam, Netherlands. North-Holland. 2007;26:81-124.
31. Raykov T. Estimation of composite reliability for congeneric measures. Applied Psychological Measurement Applied Psychological Measurement. 1997;21(2):173-184.

32. McDonald RP. The theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. British Journal of Mathematical and Statistical Psychology. 1970;23(1):1-21.
33. Wechsler D. The measurement and appraisal of adult intelligence. 4th ed. Baltimore, MD: Williams & Wilkins Co. 1958;135-143.
34. Horn JL, Blankson AN. Foundations for understanding of cognitive abilities. In: Flanagan DP, Patti LH. Editors. Contemporary intellectual assessment: Theories, tests, and issues. 3rd ed. New York. The Guilford Press. 2012;73-98.
35. Neisser U, Boodoo G, Bouchard TJ, Boykin AW, Brody N, Ceci SJ, et al. Intelligence: knowns and unknowns. American Psychologist. 1996;51(2):77-101.
36. Byrne BM. Structural equation modeling With EQS: Basic concepts, applications and programming. 2nd ed. Mahwah, New Jersey. Lawrence Erlbaum Associates. 2006;98-102.
37. Cortina JM. What is coefficient alpha? An examination of theory and applications. Journal of applied psychology. 1993;78(1):98-98.
38. Green S, Yang Y. Commentary on coefficient alpha: A cautionary tale. Psychometrika. 2009;74(1):121-135.
39. Bentler PM. Alpha, Dimension-free and model-based internal consistency reliability. Psychometrika. 2009;74(1):137-143.
40. Kaufman AS: Forward. In: Weiss LG, Saklofske DH, Coalson D, Raiford SE. editors. WAIS-IV Clinical Use and Interpretation. San Diego: Academic Press; 2010.
41. Hartman DE. Wechsler Adult Intelligence scale IV (WAIS IV): return of the gold standard. Appl Neuropsychol. 2009;16(1):85-7.


---



*Peer-review history:*
*The peer review history for this paper can be accessed here:*
*http://www.sciencedomain.org/review-history.php?iid=462&id=21&aid=4031*