

International Testing in Our Nations Schools: How to Sound Gifted in 150 + Languages.



Teddi Beam-Conroy
and Priscilla Ramirez Lurz
TAGT Conference Fall 2007

Monolingualism is the norm in the US

- Bernal: "Let them learn English first"
 - Bilingual education programs are viewed as compensatory.
 - Many GT programs do not consider students for placement until they can take assessments in English.
 - This is one factor in their underrepresentation in our gifted programs.

Bernal, 2003, p. 11

Monolingualism is the norm in the US

- Bilingualism is often regarded as inherently problematic instead of being viewed as an asset.
- Problem: How do we identify and service these children?
- Monolingual Myopia
 - Often testing is not valued unless it is done in English. Primary language assessment has not been considered a "coherent branch of educational measurement"
 - People whose primary language is something other than English are frequently viewed as having deficits.

Romaine, 1995, p. 6; Oller, 1991, p. 3

A bilingual individual is not two monolinguals in one

- Valdés and Figueroa: Research implies that the bilingual individual is “cognitively different from monolinguals.” More and different regions of the brain are used.
- The bilingual test taker cannot perform as a monolingual test taker in that having the use of more than one language complicate the student’s response.



Source: Valdés and Figueroa, 1994, pp. 69, 85

TX State Plan for the Education of Gifted/Talented Students

- 1.1.1.R. Nomination procedures and forms for assessment of gifted/talented students are communicated to families in a language and form that the families understand or a translator or interpreter is provided.



Student Assessment (TX State Plan)

- 1.5.1A Assessment of students for gifted/talented programs included measures collected from multiple sources for each area of giftedness served by the district.
- 1.5.2A Student are assessed in languages they understand or with non verbal based tests.
- 1.6.A Data and procedures assure that **all** populations of the district have assess to assessment, and if identified, services offered as part of the program for gifted students.

Student Assessment

- 1.6R Gains have been made over the 2 years toward having the population of the gifted program reflect the population of the district.
- 1.7A Final selection of students for services designed for gifted students is made by a committee of at least 3 local district or campus educators who have received training in the nature and needs of gifted students.

"No cut-off scores are required in order to be identified as gifted according to the state plan."

Some testing issues (or, a bilingual individual is not two monolinguals in one, part two)

- Tests, whether designed for bilingual individuals or not, often fail to include norming samples of a bilingual population.
- The test is asked to measure the abilities of a bilingual speaker when it was designed for monolinguals; the test-taker is being compared to someone who is unlike him/her.



Source: Valdés and Figueroa, 1994, pp. 87, 107

Other testing issues

- Even when bilinguals are included in the norming group, this does not guarantee a valid test
 - Translating a test does not resolve the language issue
 - The assumption is that the languages arrange themselves in the same order of difficulty
 - Some terms do not translate
 - Whose English? Whose Spanish? Whose Chinese? Whose....?



Source: Valdés and Figueroa, p. 104

Forms of Assessment

- Nominations
- Abilities Tests (IQ tests)
- Achievement Tests
- Creativity Measures
- Parent/Teacher Ratings

Nominations



- Teacher bias against culturally diverse and low-income students – evidence shows this remains an issue.
- Parent information – beyond the school: in other languages besides English, post information at local libraries, churches, community centers, etc.
- Provide staff development 'Nature and Needs' to District student placement staff.
- Sociocultural considerations – what does giftedness look and sound like in a certain culture? Help educate everyone on this matter.

Hispanic Gifted Bilingual Screening Instrument

- Designed as a prescreening instrument to assist teachers in nominating potentially gifted children for further testing
- Initiated in 1992
- Hispanic bilingual children grades K – 4
- Teacher administered (student is not involved in screening)
- Eleven clusters of attributes of potentially gifted Hispanic bilingual elementary school children derived from research and questionnaires of teachers, parents and community members
- Computer-scoring produces instant reports

Source: <http://teachstudents.com/>; Irby and Lara-Alecio, 1997, p. 20

Abilities Tests



BVAT

- o Designed to provide a measure of overall verbal ability with regard to the unique combination of cognitive/academic language abilities possessed by bilingual individuals
- o Published in 1998
- o Ages 2 and over
- o Individual administration
- o Jim Cummins is one co-author
- o Based on the English Woodcock-Johnson-Revised
- o Translated into 18 languages
- o Administered first in English, then in L1 for corrections to ceiling
- o Intended to capitalize on bilingual individual's Common Underlying Proficiency
- o Valdés and Figueroa: "Bilingualism is not the sum of a score in L1 plus the sum of a score in L2" – precisely how this test is scored
- o Some issues of cultural bias with certain items
- o Some items do not translate
- o Norming issues: small populations; norms represent BVAT performance for monolingual English speakers (although norms updated in 2005)
- o Most useful with examinees whose home language is not Spanish

Source: Garfinkel (2001); Stansfield (2001); Jean & Genest (2000)

Bateria Woodcock-Muñoz Revisada

- o Designed to measure achievement and cognitive abilities, scholastic aptitudes, and Spanish oral language
- o Published in 1996
- o Ages 2 and over
- o Individual administration
- o Verbal, quantitative, processing skills tested
- o Translation/adaptation of the English Woodcock-Johnson
- o Norming sample included Spanish-speakers from Peru, Mexico, Puerto Rico, Spain, Argentina, Costa Rica and Arizona, New York, California, Florida, and Texas
- o Validity correlations in the .60-.70 range (moderate)
- o One researcher cites "limited empirical evidence of the adequacy of the Bateria-R" (Frary)
- o Other reviewer cautions that "psychologists should not assume the appropriateness of results from using this test" ((Lintel)

Batería III

- Designed to assesses the cognitive abilities and achievement levels of Spanish-speaking individuals
- Published in 2005
- Ages 2 through 90
- Individual administration
- First Spanish-language cognitive assessment designed bottom-up for Spanish-speakers in the US
- 31 cognitive tests and 22 achievement tests organized into Brief scale, Full scale (all tests can stand alone)
- Includes verbal, quantitative, memory tasks
- Not a translation
- Gifted bilingual students included in norming sample
- Based on CHC Theory – includes measures of crystallized and fluid intelligence
- Small norming population in US consisting of monolingual or near-monolingual individuals (N=279)
- Larger norming population in Mexico, Puerto Rico, Spain, Argentina, Panama, Costa Rica and Colombia (N=1413)

Source: Riverside Publishers, Alvarado (2004 and 2005)

Cognitive Abilities Test (CogAT)

- Designed to assesses students' abilities in reasoning and problem solving using verbal, quantitative and nonverbal symbols
- Published in 2002 (Spanish version 1997)
- Grades K-12
- Group administration
- Items reviewed for bias by panels of classroom teachers and professional educators
- Norming process termed "carefully planned and executed" by one reviewer (Rogers)
- Large standardization sample (N=180,538) representative of 2000 Census data in terms of school type, economic status, and geographical region
- Hispanic students slightly underrepresented (DiPerna)
- ELLs represented 4-8% of sample in grades K-5; .5-2.5% in older grades
- Gender differences are minimal; however, no data reported on differences according to race, disability ELLs
- One reviewer reports "significant weakness in the reliability and validity evidence relative to the intended purposes of the test"; recommends that until empirical evidence is provided test scores should not be used to make instructional recommendations (DiPerna)
- Nonverbal battery has instructions and booklets in Spanish
- Publisher asserts that translation of the directions for the quantitative and nonverbal batteries into other languages will not affect the norms for those tests

Source: Riverside Publishing, DiPerna (n.d.) Rogers (n.d.)

Nonverbal Abilities Tests

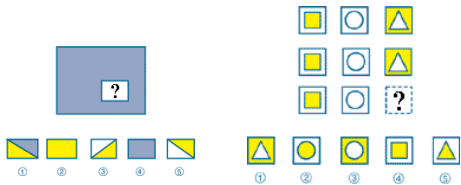
- "It is an illusion that we adjust to reality without language." Edward Sapir (1928)
- Language is what allows us to shape thought. One of the criticisms of nonverbal tests is that it is impossible to have a "language-free" testing environment, as all human action and interaction is mediated by language.

Nonverbal Abilities Tests

- Sternberg: “The kinds of geometric symbols that are touted by some to be the bases for the culture fair measurement of intelligence...are really not culture fair at all. Children who do not receive Western schooling are less likely to be familiar with such symbols than are children who do receive such schooling.”
- Another criticism is that NV tests favors those who are spatially oriented at the expense of those who are verbally/linguistically oriented.

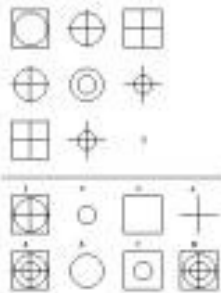
Sternberg, 2000, p. 120

Nonverbal Abilities Tests



Sample from Naglieri Non-Verbal Abilities Test

Nonverbal Abilities Tests



Sample from the Raven's Progressive Matrices

Kaufman Assessment Battery for Children (K-ABC)

- Designed to measure the processing and cognitive abilities of children and adolescents
- Published in 2004
- Ages 3-18
- Individual administration
- Norm sample mirrors 2001 U.S. Census data with respect to gender, ethnicity, parental education level, geographic region of the U.S., and educational and psychological classifications
- Gifted and talented students included in norming sample
- Data suggest that the goal of reducing ethnic group differences is not completely achieved (scale means for African American, American Indian, Asian American, Hispanic, and White examinees typically differed by one-half standard deviation, with Hispanics usually scoring lowest)
- Lack of evidence supporting the utility of test data for driving educational interventions.
- Test must be administered in English, but responses in Spanish are acceptable (guide provides correct Spanish-language responses)

Braden and Thorndike

Naglieri Nonverbal Abilities Test

- Designed to measure nonverbal reasoning and problem solving independent of educational curricula and cultural or language background
- Published in 1997 (Individual battery published in 2003)
- Grades K-12
- Individual or group administration
- An extension and revision of the Matrix Analogies Test
- Norming sample was stratified for geographic region, ethnicity, economic status, urbanicity, and public/private school attendance
- Norm sample mirrors NCES data "impressively" (Stinnett). Gifted and talented students included in norming sample
- ELLs included in norm sample, although in small numbers (103 Spanish-speakers and 18 speakers of Asian languages) compared to norm sample (N=89,600)
- Test directions also available in French, Italian, Russian and Spanish
- One of the least culturally biased assessments available today (Harcourt, 1996, p. 4), based upon reports that the NNAT standard scores differed only .25 of a standard deviation for the White and African American group, and that there were minimal differences for White and Hispanic examinees, and essentially no difference for White and Asian students
- Test items were created to be free of gender, cultural or ethnic bias (Harcourt, 1996, p. 4)
- One reviewer cautions that validity evidence is insufficient (Trevisan)
- Some districts use it universally.

Raven's Progressive Matrices

- Designed to assess nonverbal abilities at three levels
- Published in 1995
- Ages 5-11, plus mentally impaired adolescents
- Individual administration, although small groups are possible with children < age 8
- Size of norming sample varies by level
- Stratified for age only
- Reported fairness for Native American, African American and Hispanic children
- No available fairness reports for ELLs
- Spearman considered SPM to be best measure of "g"
- Moderate to good concurrent validity (majority in .70-.80)

Source: Center for Psychological Studies, Riverside Publishing, Harcourt Assessment

Test of Nonverbal Intelligence (TONI)

- o Designed to provide a "language-free measure of cognitive ability"
- o Published in 1997
- o Ages 6-89
- o Individual administration
- o One single score (Two forms, 45 items each)
- o VERY brief (10-15 minutes)
- o Abstract, figural content geared to problem solving
- o Norming sample chosen to represent US population geographically, by gender, community type, ethnicity and race, economic status, and disabling condition Concurrent validity termed "at best moderate" by one reviewer (Atlas)
- o African-Americans scored about one-third of a standard deviation below total sample
- o Test administration to gifted students termed "unnecessarily awkward" (Atlas)

Pearson Assessment Group

Universal Nonverbal Intelligence Test (UNIT)

- o Designed to provide a fair assessment of intelligence regardless of speech, language
- o Published in 1998
- o Ages 5-17
- o Individual administration
- o Three batteries, six subtests
- o Analogies, block designs, matrices, mazes
- o Although nonverbal, test offers some items amenable solution by verbal mediation (Trevisan)
- o Norming sample reflected the 1995 Census data for gender, race, Hispanic origin, region, and parental educational attainment
- o Validity evidence termed "impressive" by reviewer
- o No significant differences in performance based on gender or ethnic group
- o Equally reliable for members of all subgroups
- o School Psychology International recommends its use
- o Main criticism: not "user-friendly"

Bandalos, D.

Achievement Tests



APRENDA

- o Designed to measure the academic achievement of Spanish-speaking children
- o Published in 1998 (APRENDA 3 released in 2005)
- o Grades K5 – 12
- o Group administration
- o Reading, mathematics, language batteries based on a "consensus" curriculum (Lintel and Peterson)
- o Reading passages feature authentic literature by native language authors
- o Would provide inaccurate information on students in districts not offering Spanish-language instruction (Diaz)
- o Norms based on Spanish-speakers in the United States
- o Sample stratified for country of origin, but not for socioeconomic factors
- o Concern that unless examinees' CALP levels are known, test could measure proficiency instead of achievement (Lintel and Peterson)
- o Concern expressed over the length of the test
- o Reliability coefficients ranged from .73 to .97 (high moderate to good)
- o Some concern for construct validity (correlations between .36 to .89 between subtests at adjacent levels)

Diaz, J. O.

Logramos

- o Designed to measure the reading, math, and language achievement of Spanish-speaking children
- o Published in 2002 (currently being renamed)
- o Grades K5 – 12
- o Group administration
- o Reading, mathematics, language, listening comprehension batteries
- o Reading passages feature "varied cultural texture of the Hispanic world" (publisher)
- o Spanish common to Spanish-speaking children in the US (publisher)
- o Norms based on "large, representative national probability sample" (publisher)
- o Sample stratified for country of origin, but not for socioeconomic factors
- o Logramos, while designed to be the Spanish-language version of ITBS, is not exactly parallel to ITBS
- o More items per subtest than APRENDA
- o Reliability coefficients ranged from .73 to .97 (high moderate to good)
- o Some concern for construct validity (correlations between .36 to .89 between subtests at adjacent levels)

Riverside Publishing Co.

SABE

- o Designed to measure achievement in basic skills with children for whom Spanish is the language of instruction
- o Published in 1994
- o Grades 1-8
- o Group administration
- o Reading, mathematics, language, listening comprehension batteries
- o Reading passages feature "varied cultural texture of the Hispanic world" (publisher)
- o Norms based on Spanish-dominant sample of over 10,000 students in grades 1 through 12 included Mexican American, Puerto Rican, and Central and South American children
- o 142 schools in Arizona, California, Illinois, Massachusetts, Michigan, New Mexico, New York, Oregon, and Texas
- o Norming information regarding other variables not available
- o Items found to be biased in terms of gender, ethnicity, and age were eliminated
- o One reviewer noted a major weakness in the area of validity (López)
- o Predictive and construct validity data not provided, although content validity is strong
- o Urges more reliability and validity studies

Lintel, M. P. and Lopez, E.C.

SUPERA

- Designed as a comprehensive series offering multiple measures of Spanish-language student achievement.
- Published in 1994
- Grades 1-10 (English version is K-12)
- Equivalent to English Terra Nova
- Group administration
- Assesses reading/language arts and mathematics
- Reading passages feature original literature (publisher)
- Norms based on the performance of Spanish-speaking students throughout the United States (publisher)
- Norming information regarding other variables not available
- Most reviewer information based on the English version
- Differential item functioning found to be against Hispanic students by ratio of 13/45; African-American students 11/44 (English version)
- Unclear how much of assessment is a translation of the English version
- Predictive and construct validity data not provided for Spanish version

Source: CTB ,McGraw Hill Lintel (n.d.b) and López (n.d.)

Torrance Tests of Creative Thinking

- Designed to sample a variety of dimensions of creative thinking
- Published in 1998; norms update 2007
- Grades K- Graduate school
- Group administration
- Most commonly used tests are figural and verbal
- In wide use since the 1960s
- Respected for efforts to eliminate cultural bias, when scored as recommended
- Has been translated into over 35 languages (Kim)
- Excellent validity reported by the Center for Creative Learning
- Less biased for ELLs because figural test does not involve language (Kim)



Source: Center for Creative Thinking (2002), Kim (n.d.)

Parent and Educator Checklists



Frasier Talent Assessment Profile (F-TAP)

- Teacher completes
- Involves the gathering of standardized and nonstandardized information: Testing, rating scales and observation scales
- Designed to be useful in identifying underrepresented children
- Covers five areas:
 - Achievement
 - Intelligence
 - Critical Thinking
 - Creative Thinking
 - Observation/Performance



Frasier and Passow, 1994.

Kingore Observation Inventory

- Designed to aid teachers in the identification of gifted children
- Published in 2001 (second edition; first published in 1991)
- Teacher rating scale
- Two forms: grades K-3 and grades 4-8
- Also includes a parent information form
- Norming information not available
- Used in the US, Mexico, Canada, Australia
- Includes classroom activities to elicit observable behaviors



Source: Professional Associates

Renzulli

- Designed to aid teachers in the identification of gifted children
- Published in 1976
- Teacher/parent rating scale
- Developed from research with students grades 4-6.
- Norming information not available
- According to one reviewer, a primary weakness is the lack of demonstrated validity in enhancing the selection process (Rust)
- A shortened form may be used (instrument is long!)
- May be translated with author's consent
- Language very educator-oriented

Source: Rust (n.d.).

Testing and Policy Options

- Develop alternative approaches to testing and evaluation



Source: Valdés and Figueroa (1994), p. 172

Alternative Assessment

- Portfolios
 - Include work samples
 - Parent/student interviews
 - Observations
 - Outside interests
 - Review of classwork
 - Student-directed project/product
- Performance Assessments
 - See the "Discover" Model
 - Region 1
- Talent Pools
 - Bernal (2003)

DISCOVER Assessment

The DISCOVER Assessment is a measurement of abilities unlike any other, a research based instrument founded upon [multiple intelligence](#), [problem solving](#) and [diversity](#). Participants are guided through active, hands-on problem solving exercises (using toys and other age-appropriate materials) designed to measure problem-solving abilities. Problems & challenges presented range from simple and closed—requiring convergent thinking—to open-ended and complex—requiring divergent thinking. As participants solve the problems, certified observers take careful notes, recording which of approximately 120 "superior problem solving skills" are observed. This information is later compiled according to the respective intelligences and is used to create "[Strength Profiles](#)", reports that show the levels of strength for all the intelligences. Strength Profiles explain and expand upon the Assessment results, and can be used to improve learning or (for older students) guide selection of higher education topics and careers.

Discover Assessment

Modifications to the testing environment

- Romaine: All tests need to be interpreted in light of the effect of a particular child's minority background and knowledge about bilingualism
- Extra time
- Providing bilingual dictionaries
- Individual and small group testing
- Allowing "testing-the-limits" strategies (Castellano)
i.e. Teacher clarifies meaning or simplifies vocabulary for ELLs



Source: Romaine, 1995, p. 276

NISD International GT Testing

- Saturday International Testing
 - Hire Interpreters to act as ancillary test administrators.
 - Seek speakers of languages other than English from within District personnel.
 - Make interpreters available to parents for completing Parent Checklist.

Divergent views of giftedness

The overarching variable is **sociocultural**:
What is gifted in one cultural context is not necessarily gifted in another



References

- July, J. L. (2005) Pioneering definitions and theoretical positions in the field of gifted education. *Gifted Child Today*, 29(3), pp. 38-44.
- Kingma, B. (2001) *Sigona's Classification Inventory*.
http://www.professionalsassessmentspublishing.com/Merchant2/merchant.mvc?Screen=PROD&Store_Code=PA&Product_Code=BK-22&Category_Code=BK
- Kim, K.L. (n.d.) Why do we need the TCTET? Retrieved October 23, 2004 from <http://www.coe.edu/education/W/W/WhatTCTET.pdf>
- Lara-Alecio, R., Ivey, B., & Walker, M.V. (1997) Identification of Hispanic, Bilingual Gifted Students. *Tempo*, 16(2), pp. 20-25.
- Linke, M.P. (n.d.) Review of the Spanish Vocabulary-Matrix Revisited. In *Buros 13 Mental Measurements Yearbook*. (Electronic version). Retrieved November 20, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- Linke, M.P. (n.d.) Review of the Spanish Assessment of Basic Education, Second Edition. In *Buros 13 Mental Measurements Yearbook*. (Electronic version). Retrieved November 21, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- Linke, M.P. and Peterson, F.E. (n.d.) Review of APRENDA, la prueba de logros en español. In *Buros 13 Mental Measurements Yearbook*. (Electronic version). Retrieved November 20, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- López, E.C. (n.d.) Review of the Spanish Assessment of Basic Education, Second Edition. In *Buros 13 Mental Measurements Yearbook*. (Electronic version). Retrieved November 21, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- McDonough, M. (1999) Language attitudes, motivation, and standards. In S. McKay & N. Hornberger (Eds.), *Sociolinguistics and language teaching*. (pp. 3-46). Cambridge: Cambridge University Press.
- Naglieri, J.A., & Ronning, M.E. (2005) Comparison of White, African American, Hispanic and Asian Children on the Naglieri Nonverbal Ability Test. Washington, DC: Psychological Assessment, 17, 381-384.
- Oler, J. (1992) Language testing research: Lessons applied to LEP Students and Programs from Proceedings of the Second National Research Symposium on Limited English Proficient Students: Focus on Evaluation and Measurement. Washington: U.S. Department of Education, Office of Bilingual Education and Minority Languages Affairs. Proceedings. Retrieved September 6, 2004 from <http://www.ed.gov/olof/olof2002/olof2002.htm>
- Riverside Publishing <http://www.riverpub.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- Rogers, R.C. (n.d.) Review of the Cognitive Abilities Test, Form 6. In *Buros 13 Mental Measurements Yearbook*. (Electronic version). Retrieved October 28, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- Romano, S. (1965). *Bilingualism*. Oxford: Blackwell Publishing.
- Rust, J.D. (n.d.) Review of the Scales For Rating the Behavioral Characteristics of Superior Students. In *16 Mental Measurements Yearbook*. (Electronic version). Retrieved October 28, 2005 from <http://igwebps.utovid.com/ibweb/ib.usa.edu/ig13ovweb.asp>
- Spooley, B. (n.d.) Assessment and testing: Notes for the third millennium. Retrieved September 6, 2004 from <http://www.wednet.org/Topic/Sociology/WhatPaper.html>
- Stansfield, C.W. (2001) Test Review of Bilingual Verbal Ability Tests. From Plake, B. S., & Impara, J. C. *The Fourteenth Mental Measurements Yearbook*. (Electronic version). Retrieved March 28, 2003 from the Buros Institute's Test Reviews Online website.
- Impara, J. C. *The Fourteenth Mental Measurements Yearbook*. (Electronic version). Retrieved March 28, 2003 from the Buros Institute's Test Reviews Online website.
- Storfer, K. (2003) *Giftedness issues in the testing of ability and achievement*. In Esquivel, G.B., & Knutz, J.C. (Eds.) *Creativity and giftedness in culturally diverse students*. Clark, NJ: Hampton Press.
- Texas Department of Education (2005) List of Approved Tests for Assessment of Limited English Proficient Students. Retrieved November 21, 2005 from <http://www.ted.state.tx.us/assessment/testing/ApprovedTests0505.pdf>
- Trevisan, M. S. (2001) Test Review of Naglieri Nonverbal Ability Test. From Plake, B. S., & Impara, J. C. *The Fourteenth Mental Measurements Yearbook*. (Electronic version). Retrieved October 28, 2004 from the Buros Institute's Test Reviews Online website. <http://www.buros.org/olof/olof2002/olof2002.htm>
- Valdes, G. & Figueroa, R. (1994). *Bilingualism and testing: A special case of bias*. Westport, CT: Ablex Publishing.

Image Credits

- <http://www.agsnet.com/group.asp?nGroupInfolD=a2100>
- <http://www.riverpub.com/>
- <http://mypsychologist.com/nnat.htm>
- <http://www.gmu.edu/departments/ccd/index.html>
- <http://harcourtassessment.com/>
- www.harcourt-uk.com/
- <http://www.wilderdome.com/intelligence/IQExampleTests.html>
- http://www.cyber-test.co.kr/testinfo/images/testinfo_on02_img01.gif
- <http://www.bertiekingore.com/index.htm>
- <http://aqs.pearsonassessments.com/group.asp?nGroupInfolD=a19100>