International J. Soc. Sci. & Education
2014 Vol.4 Issue 3, ISSN: 2223-4934 E and 2227-393X Print

The Construct and Predictive Validity of Beggs's Developing Cognitive Abilities Test among Junior Secondary School Students

By

A. Adediwura and F. C. Ayeyemi
Department of Educational Foundations and Counseling
Faculty of Education, Obafemi Awolowo University, Ile-Ife

Abstract

The study determined the construct validity of DCAT and investigated the predictive ability of DCAT on students' performance in Mathematics and English Language with the aim of ascertaining the suitability of DCAT use among junior secondary school students in Nigeria. The study adopted descriptive survey design. The population for the study comprised all public junior secondary school students in Ondo state. The sample consisted of 1080 JSS III students randomly selected from 18 junior secondary schools. Developing Cognitive Ability Test and OSJSCE questions were adopted for data collection. Data were analysed using Pearson product moment correlation Coefficient (r) and regression analysis. The results showed that while the items of verbal ability subset of DCAT are measuring a construct not different from English Language subset of OSJSCE (r = 0.671), the reverse is the case for quantitative and spatial abilities subsets (r = -0.019 and r = -0.034) respectively. The result also showed that the items of quantitative and verbal abilities subset of DCAT converge with the items of mathematics subset of OSJSCE (r = 0.709 and r = 0.178) respectively. Furthermore, the results showed that while DCAT significantly predict students' performance in mathematics (β = 0.517), it does not significantly predict students' performance in English Language. However, the verbal ability subset of DCAT significantly predicts students' performance English Language. The study therefore concluded that only the verbal and the quantitative subsets of DCAT is suitable for use among the junior secondary school students.

1. Introduction

A test can be considered as technical operation that consists of determination of one or more characteristics of a given product, process or service according to a specified procedure (Godfrey, 1999). Often a test is part of an experiment. Tests are designed to measure various aspects of intellectual and emotional functioning, including personality traits, motivation, beliefs and attitudes, ability and intelligence, and various areas of emotional concern. All tests have a prescribed content, empirically defined measurement properties, and standardized methods of administration and interpretation.

According to Thissen and Wainer, (2001), a test score may be interpreted with regards to a norm or criterion, or occasionally both. The norm may be established independently, or by statistical analysis of a large number of participants. The test result can be qualitative (yes/no), categorical, or quantitative (a measured value). It can be a personal observation or the output of a precision measuring instrument. Usually the test result is the dependent variable, the measured response based on the particular conditions of the test or the level of the independent variable. Some tests, however, involve changing the independent variable to determine the level at which a certain response occurs: in this case, the test result is the independent variable.

A test may be developed and administered by an instructor, a clinician, a governing body, or a test provider. In some instances, the developer of the test may not be directly responsible for its administration. For example, Educational Testing Service (ETS), a non-profit educational testing and assessment organization, develops standardized tests such as the SAT but may not directly be involved in the administration or proctoring of these tests. As with the development and administration of educational tests, the format and level of difficulty of the tests themselves are highly variable and there is no general consensus or invariable standard for test formats and difficulty. Often, the format and difficulty of the test is dependent upon the educational philosophy of the instructor, subject matter, class size, policy of the educational institution, and requirements of accreditation or governing bodies. In general, tests developed and administered by individual instructors are non-standardized whereas tests developed by testing organizations are standardized. Psychological tests fall into a multitude of categories, with many different kinds of tests. These are tests that are typically found in a school setting, clinical settings and hospital settings, all of which measure different aptitudes. These tests are merely a way to examine and assess information.

Validity is often claimed to be the single most important consideration in assessment. The general concept of validity was traditionally defined as the degree to which a test measures what it claims, or purports, to be measuring. Validity is also conceptualized in relation to its use; in this sense, validity describes how appropriate or accurate are the interpretations made from the test scores in relation to its use. Thus, validity is concern with the specific use of the results. For example, a test may be constructed to measure student achievement, to predict the likelihood of successful completion of higher education in a discipline or to diagnose the problems that student have in reading comprehension. Even though, validation procedures are not matched to specific test uses on a one-to-one basis, it is important that the interpretation of the test outcome be meaningfully related to the purpose for which the test was developed. It is generally accepted that the concept of scientific validity addresses the nature of reality and as such is an epistemological and philosophical issue as well as a question of measurement. The use of the term in logic is narrower, relating to the truth of inferences made from premises. Validity is important because it can help determine what types of tests to use, and help to make sure researchers are using methods that are not only ethical, and cost-effective, but also a method that truly measures the idea or construct in question.

Validity is related to another concept, called reliability. Reliability refers to the consistency with which an instrument measures what it purports to measure. It deals with consistency of measurement scores. Validity of an assessment, on the other hand, is the degree to which it measures what it is supposed to measure. This is not the same as reliability, which is the extent to which a measurement gives results that are consistent. Within validity, the measurement does not always have to be similar, as it does in reliability. When a measure is both valid and reliable, the results will appear as in the image to the right. Though, just because a measure is reliable, it is not necessarily valid (and vice-versa). Validity is also dependent on the measurement measuring what it was designed to measure and not something else instead. (Kramer, Geoffrey, Douglas, Bernstein, and Phares. 2009). Validity (similar to reliability) is based on matters of degrees; validity is not an all or nothing idea. Construct and predictive validity are two of the many different types of validity.

Construct validity refers to the extent to which operationalizations of a construct (i.e., practical tests developed from a theory) do actually measure what the theory says they do. For example, to what extent is an IQ questionnaire actually measuring "intelligence"? In a concise form, construct validity is the degree to which a test or questionnaire score is a measure of the psychological characteristic of interest (Cronbach and Meehl, 1955). Construct validity evidence involves the empirical and theoretical support for the interpretation of the construct. Such lines of evidence include statistical analyses of the internal structure of the test including the relationships between responses to different test items. They also include relationships between the test and measures of other constructs.

Predictive validity refers to the degree to which the operationalization can predict (or correlate with) other measures of the same construct that are measured at some time in the future. Again, with the selection test example, this would mean that the tests are administered to applicants, all applicants are hired, their performance is reviewed at a later time, and then their scores on the two measures are correlated. In psychometrics predictive validity is the extent to which a score on a scale or test predicts scores on some criterion measure (Cronbach, & Meehl, 1955). For example, the validity of a cognitive test for job performance is the correlation between test scores and, for example, supervisor performance ratings. Such a cognitive test would have predictive validity if the observed correlation were statistically significant.

An early definition of test validity identified it with the degree of correlation between the test and a criterion. Under this definition, one can show that reliability of the test and the criterion places an upper limit on the possible correlation between them (the so-called validity coefficient). Intuitively, this reflects the fact that reliability involves freedom from random error and random errors do not correlate with one another. Thus, the less random error in the variables, the higher the possible correlation between them. Under these definitions, a test cannot have high validity unless it also has high reliability. However, the concept of validity has expanded substantially beyond this early definition and the classical relationship between reliability and validity need not hold for alternative conceptions of reliability and validity. Within classical test theory, predictive or concurrent validity (correlation between the predictor and the predicted) cannot exceed the square root of the correlation between two versions of the same measure — that is, reliability limits validity.

Cognitive abilities are the brain-based skills we need to carry out any task from the simplest to the most complex. They have more to do with the mechanisms of how we learn, remember, problem-solve, and pay attention rather than with any actual knowledge. Any task can be broken down into the different cognitive skills or functions needed to complete that task successfully. For instance, answering the telephone involves at least: perception (hearing the ring tone), decision taking (answering or not), motor skill (lifting the receiver), language skills (talking and understanding language), and social skills (interpreting tone of voice and interacting properly with another human being).

With age, some cognitive abilities decline, especially the executive functions. In addition, cognitive abilities that are not used regularly tend to diminish over time. This may happen at any age but is mostly observed in older age. Fortunately, these skills can also be improved at any age with regular practice. Cognitive development is the construction of thought processes, including remembering, problem solving, and decision-making, from childhood through adolescence to adulthood. It was once believed that infants lacked the ability to think or form complex ideas and remained without cognition until they learned language. It is now known that babies are aware of their surroundings and interested in exploration from the time they are born. From birth, babies begin to actively learn. They gather, sort, and process information from around them, using the data to develop perception and thinking skills.

Cognitive development refers to how a person perceives, thinks, and gains understanding of his or her world through the interaction of genetic and learned factors. Among the areas of cognitive development are information processing, intelligence, reasoning, language development, and memory.

The most well-known and influential theory of cognitive development is that of French psychologist Jean Piaget (1896–1980). According to Piaget, cognitive development involves an ongoing attempt to achieve a balance between assimilation and accommodation that he termed equilibration. At the center of Piaget's theory is the principle that cognitive development occurs in a series of four distinct, universal stages, each characterized by increasingly sophisticated and abstract levels of thought. These stages always occur in the same order, and each builds on what was learned in the previous stage. They are as follows: Sensorimotor stage (infancy), Pre-operational stage (toddlerhood and early childhood), Concrete operational stage (elementary and early adolescence) and Formal operational stage (adolescence and adulthood). The ability to learn and make sense of new information is crucial to successful learning… and that's why developing cognitive skills is so important.

Henry and Bardo (1987) asserted that the Developing Cognitive Abilities Test (Beggs and Mouw, 1980) was developed for the assessment of the aptitudes of students in the overall Comprehensive Assessment Program (CPA) test battery. The Developing Cognitive Abilities Test (DCAT) is similar in format to traditional intelligence tests but the DCAT was developed along a format that included both a cognitive taxonomy and content areas. The concept seems to be unique for tests designed to measure student aptitudes. The DCAT is part of the Comprehensive Assessment Program that also includes the National Achievement Test and the School Attitude Measure. All three measures were co-normed and nationally standardized from 1988–1989 using a stratified, multistage probability sample with 92,397 students. The DCAT "provides continuous measurement of student growth from grades 1 through 12" and is purported to be useful in educational decision making and student evaluation (Wick, 1990). Other uses of the DCAT include simultaneous use with achievement batteries and attitude measures in order to develop profiles and identify discrepancies, strengths, and weaknesses that could be used in determining areas for intervention (Wick, 1990). The DCAT is also used as a screening measure for identifying potentially gifted students in an objective manner (Wick, 1990). Perhaps the most important distinguishing characteristic of the DCAT is the link between specific items and Bloom's cognitive taxonomy. This study was concerned with the predictive validity of the DCAT.

Students' performance in internal and high stake examinations in Nigeria has been relatively poor over the years. Teaching and learning resources, even when they are available, have been observed not to be helping in solving this problem. Studies (Carretta & Ree, 1995, Schmidt & Hunter, 2004, Wick, Beggs, & Mouw, 1989) have shown that students' developing cognitive abilities measured using the Development Cognitive Abilities Test (DCAT) has effect on their performance. The DCAT provides continuous measurement of student growth and is purported to be useful in educational decision making and student evaluation. Thus, to improve students' performance, there is the need to improve their developing cognitive abilities. However, the suitability of the available Developing Cognitive Abilities Test (DCAT) in establishing the developing cognitive abilities among students in the study area has not been established. Thus, the study is specifically aimed at: determining the construct validity of DCAT as well as investigating the predictive ability of DCAT on students' performance in Mathematics and English Language.

Research Question: What is the construct validity of Developing Cognitive Abilities Test (DCAT) among Ondo State students?

Research Hypotheses

1. DCAT score cannot significantly predict students' performance in Mathematics.
2. DCAT score cannot significantly predict students' performance in English Language.

2. Method

The study adopted descriptive survey design. The population for the study comprised all public junior secondary school students in Ondo state. The study sample consisted of 1080 junior secondary three (JSS III) students. From each of the three senatorial district of Ondo State, three Local Government Areas (LGAs) were randomly selected and from each of the LGAs two junior secondary schools were randomly selected making a total of 18 schools. A total of 60 JSS III students were randomly selected from each school. JSS III students were used in this study based on their ability to read and interpret what they read.

Two instruments Developing Cognitive Ability Test (DCAT: Wick, Beggs, & Mouw, 1989) and 2008 Ondo state Junior Certificate Examination questions were adopted for collecting necessary data for the study. DCAT is a standardized test and a group administered measure of mental abilities for students in secondary school, it includes Verbal, Quantitative and Spatial subsets which combined to provide a total DCAT score. Each Subset comprises of 10 items.

The Ondo State Junior School Certificate Examination (OSJSCE) is standardized examination conducted for all the JSS III students in the State in both public and private schools as qualifying pre-requisite into senior secondary school. DCAT was first administered on the selected students and two weeks after the

OSJSCE was also administered on the students. The scripts were scored and collated for analyses. However, only 960 students seated for both tests. Data were analysed using Pearson product moment correlation Coefficient (r) and regression analysis

3. Result

Research Question 1: What is the construct validity of Developing Cognitive Abilities Test (DCAT) among Ondo State students?

To answer this question, the construct validity of the DCAT was examined through convergent validity. Students" score on each of the subsets of DCAT were correlated with their scores in mathematics, English Language and Integrated Science serving as the subsets of OSJSCE. Using convergent procedure of establishing the construct validity of DCAT it is expected that subset of DCAT and OSJSCE measuring similar constructs should have higher correlations than subtests measuring different constructs. Table 1 presents the correlation between the subsets of DCAT and subsets of JSCE.

Table 1: Correlation of between the subsets of DCAT and subsets of JSCE

		VA	QA	SPA	DCAT
English Language	Pearson Correlation	.671*	-.019	-.034	.006
	Sig. (2-tailed)	.000	.548	.300	.845
	N	956	956	956	956
Mathematics	Pearson Correlation	.178*	.709*	-.155*	.031
	Sig. (2-tailed)	.000	.000	.000	.342
	N	950	950	950	950
Integrated Science	Pearson Correlation	.040	.129*	.033	.105**
	Sig. (2-tailed)	.220	.000	.306	.001
	N	960	960	960	960
JSCE	Pearson Correlation	.183*	.053	-.078*	.070*
	Sig. (2-tailed)	.000	.103	.017	.031
	N	948	948	948	948

Table 1 showed that the correlation between the DCAT Verbal subtest and JSCE English Language subtest (r = .671, p <.05) was higher than the DCAT Quantitative subtest and JSCE English Language subtest, (r = -.019, p>.05), and higher than the correlation between the DCAT Spatial subtest and JSCE English Language subtest, (r = -.034, p>.05). The Table also showed that the correlation between the DCAT Quantitative subtest and JSCE Mathematics subtest, (r =.709, p<.05) was higher than the DCAT Verbal subtest and JSCE Mathematics subtest,(r = .178, p<.05); and higher than the DCAT Spatial subtest and JSCE Mathematics subtest correlation, (r = -.155, p<-05). Finally, the correlation between the DCAT quantitative subtest and JSCE Integrated Science subtest, (r = .129, p<.05), was higher than the DCAT verbal subtest and correlation, (r = .040, p >.05); and higher than the correlation between the DCAT Spatial subtest and JSCE Integrated Science subtest, (r = .033, p >.05).

Research Question 2: What is the consistency of DCAT?
To answer this question, two approaches (test-retest and split-half) of estimating the reliability of an instrument were employed. The use of split half involved dividing the items of the DCAT into two along odd and even number items, while the test-retest involved administering the instruments twice on the

students within a period of two weeks. Students scores on each of the two approaches were then correlated using Pearson Product Moment correlation Coefficient (r). The results are as presented in Table 2.

Table 2: DCAT Consistency estimate among Ondo State students

Reliability Estimate Method		N	$\overline{X}$	SD	r	p
Test-retest	DCAT$_1$	960	26.03	10.14	0.994	<.05
	DCAT$_2$	960	26.17	10.26		
Split Half	½DCAT$_1$	960	13.67	5.10	0.923	<.05
	½DCAT$_2$	960	12.36	5.04		

Table 2 that the test-retest reliability coefficient estimate of DCAT was r = 0.994, which implies that the use of the instrument will consistently produce stable result. The Table also showed that the DCAT is internally consistent with a split half reliability coefficient estimate r = 0.923.

Hypothesis 1: DCAT score cannot significantly predict students' performance in Mathematics.
To test this hypothesis, students score in JSCE Mathematics and the entire score on DCAT were compiled and analysed using simple regression statistics in order to determine the extent to which DCAT score as whole can predict students score in Mathematics. The result is as presented in Table 3.

Table 3: Beta coefficient and t-ratio for DCAT prediction of students' score in Mathematics

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	12.298	.511		24.073	.000
	DCAT	.517	.078	.451	3.947	.004

The first important thing to note in Table 3 is that the sign of the coefficient of ***DCAT*** is positive. This implies that students' performance in Mathematics increase with increase in DCAT score. Also, the probability (p = 0.04) as reported in Table 3 for DCAT implies that the slope (□ = 0.517) is statistically significant. The intercept (□ = 12.298, p <.05) as shown in the Table is also significant. Thus regression line *Y = 0.517X + 12.298 means that* when students score in DCAT increase by one unit (i.e. 1%), students' performance in Mathematics increase by significant 51.7%.

Further analysis was carried out using multiple regressions; in this case the subsets of DCAT (verbal, quantitative and spatial abilities) were used as independent to determine the extent to which each of the subsets can predict students' performance in Mathematics. Table 4 presents the results.

Table 4: Beta coefficient and t-ratio for relative contribution of DCAT subsets prediction of students' score in Mathematics

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	11.878	.506		23.460	.000
	Verbal Ability	.190	.036	.178	5.256	.000
	Quantitative Ability	.721	.025	.679	7.263	.000
	Spatial Ability	-.145	.029	-.159	-5.055	.000

Table 4 showed that the sign of the coefficient of verbal and quantitative subset of DCAT were positive while that of spatial was negative. This implies that students' performance in Mathematics increase with increase in verbal and quantitative abilities subsets of DCAT score. However, students' performance in mathematics decreases with increase in spatial ability. Also, the probability (p< .05) as reported in Table 4 for verbal, quantitative and spatial abilities subsets of DCAT implies that the slopes (□ = 0.190, □ =0.721 and □ = -.145) respectively are statistically significant. The intercept (□ = 11.878, p <.05) as shown in the Table is also significant. Thus regression line *Y = 0.190X + 0.721X – 0.145X +11.878 means that* when students score in verbal and quantitative abilities subset increase by one unit each (i.e. 1%), students' performance in Mathematics increase significantly by 19.0% and 72.1% respectively, while a unit increase in spatial ability reduce students' performance in Mathematics significantly by 14.5%

Hypothesis 2: Students' DCAT score cannot significantly predict students' performance in English Language.

To test this hypothesis, students score in JSCE English Language and the entire score on DCAT were compiled and analysed using simple regression statistics in order to determine the extent to which DCAT score as whole can predict students score in English Language. The result is as presented in Table 5

Table 5: Beta coefficient and t-ratio for DCAT prediction of students' score in English Language

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.528	.404		38.391	.000
	DCAT	.003	.014	.006	.195	.846

The first important thing to note in Table 5 is that the sign of the coefficient of ***DCAT*** is positive. This implies that students' performance in English Language increase with increase in DCAT score. However, the probability (p = 0.846) as reported in Table 5 for DCAT implies that the slope (□ = 0.003) is not statistically significant. Thus regression line *Y = 0.003X + 15.528 means that* when students score in DCAT increase by one unit (i.e. 1%), students' performance in English Language increase by non significant 0.3%.

Further analysis was carried out using multiple regressions; in this case the subsets of DCAT (verbal, quantitative and spatial abilities) were used as independent to determine the extent to which each of the subsets can predict students' performance in Mathematics. Table 4 presents the results.

Table 6: Beta coefficient and t-ratio for relative contribution of DCAT subsets prediction of students' score in English Language

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	15.295	.412		37.090	.000
	VA	.074	.029	.087	2.519	.012
	QA	-.055	.037	-.053	-1.517	.130
	SPA	-.027	.023	-.037	-1.155	.248

Table 6 showed that the sign of the coefficient of verbal subset is positive while quantitative and spatial abilities subsets of DCAT were negative while. This implies that students' performance in English Language increase with increase in verbal ability subsets of DCAT score, but decreases with increase in quantitative and spatial abilities. Also, the probability ($p< .05$) as reported in Table 6 for verbal ability and $p>.05$ for quantitative and spatial abilities subsets of DCAT implies that while the slope (□ = .074) of the verbal ability is significant, the slopes (□ = -0.55, and □ = -0.055) for quantitative and spatial abilities respectively are not statistically significant. With a significant intercept (□ = 15.295, $p <.05$) as shown in the Table the regression line $Y = 0.074X -0.055X - 0.027X +15.295$ *means that* when students score in verbal ability increase by one unit, students, performance in English Language increase significantly by 7.4% while a unit increase in each of quantitative and spatial abilities non-significantly reduces students' performance by 5.5% and 2.7% respectively.

4. Discussion

These results show that the DCAT Verbal subtest was associated with the verbally oriented OSJSCE subtest (English Language) to a greater extent than either the DCAT Quantitative or Spatial subtests that are not as verbally oriented. Likewise, results showed that the DCAT Quantitative subtest was associated with the quantitatively oriented OSJSCE subtest (Mathematics and Integrated Science) to a greater extent than either the DCAT Verbal or Spatial subtests. These findings are in agreement with Canivez (2000) finding that reported that DCAT Verbal subtest was associated with the verbally oriented ITBS subtests (Vocabulary, Reading, and Language Usage) to a greater extent than either the DCAT Quantitative or Spatial subtests that are not as verbally oriented and that the DCAT Quantitative subtest was associated with the quantitatively oriented ITBS subtest (Mathematics Problem Solving) to a greater extent than either the DCAT Verbal or Spatial subtests. The lowest correlations obtained were between the DCAT Spatial subtest and the three OSJSCE subtests, although it was significant for Mathematics. This was not an unexpected result as these subtests are not as theoretically related.

The present study showed that while DCAT as whole will significantly predict students' future performance in mathematics; it does not significantly predict students' future performance in English Language. However, an increase in verbal ability subset of DCAT brings about significant improvement in students performance in OSJSCE Mathematics and English Language. But while an increase in quantitative ability brings about significant increase in students performance in Mathematics the reverse is the case for English Language. These results are similar to the results obtained by Henry and Bardo (1987) in their study of the DCAT standardization sample and Canivez (2000). The Verbal and Quantitative subtests of the DCAT and the subtests of the OSJSCE are most certainly *crystallized* (*Gc*) abilities whereas the Spatial subtest could be considered a type of *fluid* (*Gf*) ability (Cattell, 1971; Gustafsson, 1988) or *visualization* (*Gv*) ability (Carrol, 1993; Gustafsson, 1988). DCAT Spatial items are a mixture of tasks requiring recognition and reasoning of objects' size, shape, symmetry, and pattern; identification of objects changes in location or position; and mental rotation, folding, or identifying divisions of objects; and mechanical principles. As such, these items are associated with a broad array of major spatial factors (Lohman, 1988) that may very well overlap and be related to both *Gf and Gv* (Horn) in what may be referred to as *Gfv* (Snow & Lohman, 1984).

The present study provided ample evidence of the short term predictive and construct validity of the DCAT (Level H) with a heterogeneous sample of JSS III students. Additional validity studies are needed to see if these findings replicate. Additional research should investigate psychometric characteristics of the DCAT with different populations. Specifically, studies of differential predictive validity should be conducted to help determine the presence or absence of predictive validity bias. If future studies replicate the present findings the DCAT may become a more frequently used group measure of cognitive abilities among junior secondary school students. Limitations of the present study include examining the predictive validity of only one level of the DCAT (Level H). Thus, results may not generalize to other age groups or levels of the DCAT. Future studies should attempt to utilize more representative samples in order to generalize to the larger population. These findings, however, are certainly encouraging.

References

Beggs, D.L., & Mouw, J.T. (1980). *Developing Cognitive Abilities Test*. Carbondale, IL: Southern Illinois University Press.

Canivez, G. L. (2000). Predictive and construct validity of the Developing Cognitive Abilities Test: Relations with the Iowa tests of basic skills. *Psychology in the schools, 37 (2)* 107-113.

Carretta, T. R., & Ree, M. J. (1995). Near identity of cognitive structure in sex and ethnic groups. *Personality and Individual Differences, 19,* 149-155.

Carmines, E., and Zeller, R., 1979. *Reliability and Validity Assessment*. Sage Publications, Beverly Hills, California.

Carrol, J. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge, England: Cambridge University Press.

Cattell, R.B. (1971). *Abilities: Their structure, growth, and action.* Boston: Houghton-Mifflin.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity for psychological tests. *Psychological Bulletin*, *52*, 281-302.

Cronbach, L. J. (1969). Validation of educational measures. *Proceedings of the 1969 Invitational Conference on Testing Problems. Princeton*, NJ: Educational Testing Service, 35-52.

Godfrey, A. B. (1999). *"Juran's Quality Handbook"*, 1999, ISBN 007034003

Gregory, R.J. (1996). *Psychological testing: History, principles, and applications* (2nd ed.). Boston, MA: Allyn & Bacon.

Gustafsson, J.-E. (1988). Hierarchical models of individual differences in cognitive abilities. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 35–71). Hillsdale, NJ: Lawrence Erlbaum Associates. 112 *Canivez*

Henry, P., & Bardo, H.R. (1987). The predictive validity of the Developing Cognitive Abilities Test. *Educational and Psychological Measurement*, 47, 207–214.

Henry, P., & Bardo, H.R. (1990). Relationship between scores on Developing Cognitive Abilities Test and scores on Medical College Admissions Test for non-traditional premedical students. *Psychological Reports*, 67, 55–63.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. *Journal of Vocational Behavior, 29*(3), 340-362.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96,* 72-98.

Kramer, Geoffrey P., Douglas A. Bernstein, and Vicky Phares, (2009). *Introduction to clinical psychology.* 7th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2009. Print.

Lohman, D.F. (1988). Spatial abilities as traits, processes, and knowledge. In R.J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 4, pp. 181–248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality & Social Psychology, 86*(1), 162-173.

Snow, R.E., & Lohman, D.F. (1984). Toward a theory of cognitive aptitude for learning and instruction. *Journal of Educational Psychology*, 76, 347–376.

Thissen, D., & Wainer, H. (2001). *Test Scoring*. Mahwah, NJ: Erlbaum. *Page 1, sentence 1*.

Wick, J.W., Beggs, D.L., & Mouw, J.T. (1989). *Developing Cognitive Abilities Test* (2nd ed.). Chicago, IL: American Testronics. *Predictive and Construct Validity* 113.

Wick, J.W. (1990). *Developing Cognitive Abilities Test*: Technical Manual-Prepublication Edition. Chicago, IL: American Testronics.