

## **Realability and Validity Study of the School and College Ability Test (SCAT) Advanced Form**

**Dr. Hayfa Tayseer Elbokai**  
Al-Balqaa' Applied University  
Department of Speical Education  
Salt- Jordan

### **Abstract**

*The aim of this study is to study the psychometric characteristic of the School and College Ability Test (SCAT) advanced form.*

*First, the translation of the items into Arabic is discussed; second, the validation and the reliability studies are examined.*

*The random sample was selected for this purpose, which consists of 299 students from the first and second year of secondary schools, also this sample selected from gifted students to support the validity by contrasted group's method. We use also the Standard Progressive Matrices (SPM) and the scholastic achievement for this purpose. The finding showed that there is a significant correlation between the verbal and quantitative sections for the SCAT, also between the total score of the SCAT and the scholastic achievement. Besides, the finding showed the significant correlation between the verbal, quantitative, total score of the SCAT and the SPM.*

*On the other hand, the finding showed significant differences between the mean scores of normal students and gifted students for the SCAT.*

*For the reliability, the finding shows the Alpha coefficient is 0.88 for the total score. And the split-half coefficient is 0.65.*

*Finally, the correlation between the two sections' scores and the total score is significant.*

**Key words:** Reliability, Validity, School & College Ability Test (SCAT), Standard Progressive Matrices (SPM), scholastic achievement.

### **1. Introduction**

The School and College Ability Tests (SCAT) Series III were developed in 1980 by Educational Testing Service (ETS). All rights to SCAT Series III were purchased from ETS in 1996 by the Johns Hopkins University Center for Talented Youth (CTY). At that time, the SCAT was revised to eliminate or change any items that contained gender, ethnic, or culturally biased wording. CTY has used the SCAT since 1985 to identify highly able students who may need special education programs such as summer courses or distance learning opportunities, as well as to plan appropriate education plans for individual students. CTY's Elementary Student Talent Search established in 1997 uses newly constructed and computer-administered forms of the SCAT to identify academically talented students nationwide.

The development of the SCAT was requested by Educational Records Bureau(ERB) to replace the Academic Council on Educational Psychological Examination. The new test was to be one that would furnish reliable and valid measurement of the scholastic aptitude of pre-college students.

The first step in the development of the test for ERB involved a review of the item types that might be used to provide an indication of the verbal and quantitative abilities of students as well as an estimate of total academic aptitude. An experimental program carried out prior to the development of the original SCAT series in 1957 tested the efficacy of nine item types in the measurement of the school-developed abilities. The results indicated that the verbal analogies were among the four best predictors of academic success. Because of this finding, and because there was good precedent for basing all or a major portion of an academic aptitude test on that item type, it was decided that the verbal part of the SCAT would be composed of analogies.

The decision regarding the item type to be used in the measurement of school-developed abilities was based on other considerations as well. In the years following the experimental program mentioned above, ETS had developed considerable effort to investigating the effectiveness of the quantitative-comparison type of item as a measure of quantitative ability and predictor of academic success. Experience showed that this item type is an effective measure and a valid predictor and possesses at least two other desirable characteristics: 1)the amount of reading required is kept to a minimum and 2)comparisons can often be made by estimation or by concept-quantitative items. Hence, the reliability of a test or test part can be improved by including more items in the same testing time.

The above factors led to a decision that SCAT would consist of a verbal part and a quantitative part, each containing 50 four-option items, and that two forms of the test would be created. The result of pretesting indicated that the fourth answer choice of the quantitative items, "Not enough information is given to decide," was a source of confusion for pupils in the lower grades (4-6), so that choice was eliminated from quantitative items at the lowest level.

SCAT III, developed in 1980, contains many from previous versions of the SCAT and some newly developed ones. Between 5 to 10 percent of the previous items were revised for SCAT III on the basis of changes in curricula that had made some items less relevant to the measurement of academic ability than they had been when originally developed. Minor changes in the verbal section of the SCAT were made in 1996 when the Johns Hopkins University Center for Talented Youth purchased the rights to the test. These changes were suggested by ETS to replace outdated language or word usage, or to eliminate any gender or ethnic bias in the verbal analogy items. The revisions for SCAT III do not affect the concept being measured by the items, or the published norms (CTY, 1999, p.p.1-3).

## **2. Literature Review**

### **2.1. School and College Ability Test (SCAT) Validation in Spain: Overview of the Process and Some Results**

In this study, we shall briefly cover some of the relevant aspects of the process of the School College Ability Test (SCAT) validation in Navarra, Spain, paying special attention to certain methodological questions of interest. The paper will begin by describing the tests to be validated and the methodology that has been followed during the process. Having established these points, we will present the results obtained, with the goal of establishing the norms of these tests to be used in Spain for the detection of verbally and mathematically talented pupils, applying the model developed for the Study of Mathematically Precocious Youth (SMPY) in 1971. Since 1979, the Center for Talented Youth (CTY) at Johns Hopkins University has continued using this model in Talent Searches, and other universities have adopted it as well, with history and results published in numerous studies (Tourón, J. 2000b).

### **2.2.Intelligence, creativity, and achievement in a nonselective public junior high school**

2 creativity tests from the Torrance battery were administered to 181 9th-grade students along with the School and College Achievement Test (SCAT) and Sequential Tests of Educational Progress (STEP) batteries. A high-SCAT group consisting of Ss scoring in the upper 3rd on SCAT but not on creativity was compared with a high-creativity group consisting of Ss scoring in the upper 3rd on creativity but not on SCAT. The high-SCAT group was superior on both school grade-point average and STEP scores. To test Torrance's threshold hypothesis, a twice talented group, high on both SCAT and creativity, was compared with the high-SCAT group. These 2 groups did not differ in STEP scores, but the twice talented group was significantly lower than the high-SCAT group on grade point average

(Payne & Leona 1965).

### **2.3. Differential characteristics of Japanese-American and Chinese-American college students.**

Asked all entering freshmen at a large northern California university to participate in a testing program that included the School and College Ability Test, the SVIB, and the Omnibus Personality Inventory. Results from 106 Japanese-Americans, 236 Chinese-Americans, and all other students were analyzed and compared to 1 another. Findings support the hypothesis that these 2 ethnic groups differ from all other students in the same direction because of similar cultural values.

However, Japanese-American Ss consistently occupied an intermediate position between the 2 other groups on all 3 measures. Results suggest some form of differential acculturation due to historical and political circumstances for the 2 ethnic groups (Derald & Barbara, 1973).

#### 2.4 Psychological characteristics of Chinese-American students.

Administered the School and College Ability Test, the Omnibus Personality Inventory, and SVIB to 3,053 University of California entering freshmen. 236 Chinese-American Ss were identified and compared to all other Ss. Results indicate that Chinese-American students (a) scored higher on quantitative and lower on verbal and combined sections of the ability test; (b) tended to be more interested in physical sciences, applied technical fields, and business occupations, and less interested in social sciences, aesthetic-cultural fields, and verbal-linguistic vocations; and (c) were more conforming, less socially extroverted, preferred concrete-tangible approaches to life, and experienced greater emotional distress than other students. These differences are discussed in relationship to the subcultural values of the Chinese-American (Derald & Barbara, 1972).

#### 2.5. The Implementation of the Talent Search Concept in Spain.

In this study, we will present some of the results we have obtained in validating the SCAT (School and College Ability Test) in Navarra, Spain. The data demonstrate the psychometric properties of this test with which for the first time in Spain.

For the reliability, the results shows vales Cronbach's of between 0.89 and 0.91 for the total score, of between 0.78 and 0.82 for the verbal section, and for the quantitative section of between 0.87 and 0.88.

For the difficulty (p) and discrimination (rb) The mean difficulty for the advanced level presented difficulties of between 0.53 and 0.65 for the verbal section and between 0.56 and 0.70 for the quantitative part. The mean discrimination for the advanced level are somewhat less, above all in the verbal section (0.41 to 0.43), given that those for the quantitative section are 0.49 or 0.50 (Tourón, J. 2000).

### 3. Methods

#### 3.1 Participants

The sample of this study consists of the (299) students were selected randomly from four secondary public schools in Damascus-Syria, besides to (153) students from Al Basel gifted school in Damascus.

Table 1 describes the characteristic of validity and reliability Sample.

	Grade			Total
	10	11 literature	11 Science	
Female	37	19	69	125
Male	67	22	85	174
Total	104	41	154	299

**Table 1 Characteristic of Validity and Reliability Sample.**

#### 3.2 Material

##### 3.2.1. School and Collage Ability Tests (Advanced Form)

The SCAT measures verbal and quantitative abilities of students in grades 3-12. It is useful in comparing students in class, comparing performance on the verbal and quantitative parts, estimating growth in abilities over time, and predicting success in related achievement areas. SCAT tests measure a student's specific skill or content taught in specific grades. To measure a student's specific skill levels or content knowledge, an achievement test should be used.

SCAT series III has two separately timed parts. Part I, Verbal, uses verbal analog items to measure a student's understanding of words. Part II, Quantitative, tests a student's understanding of fundamental number operations through quantitative comparison items. These items require the student to compare the magnitude of two mathematical quantities and make decisions concerning them. The questions place minimum emphasis on reading and require more resourcefulness and insight than traditional computation items. The tests yield three scores: Verbal, Quantitative, and total. Working time for the entire test is 40 minutes- 20 minutes for each part.

SCAT series III includes tests at three levels of difficulty. Alternate forms, which are comparable in content coverage and difficulty, are available at each level. The SCAT was designed for measuring the verbal and quantitative abilities of students who are in grades 3-12. It can also be used "above level" to assess the full extent of abilities for students who are exceptionally able and are, therefore, "hitting the ceiling" of a grade-level test. The forms and grades for which levels were developed and to be used when administering the SCAT at grade level are:

LEVELS	FORMS	GRADES
Elementary	X and Y	3 spring through 6 fall
Intermediate	X and Y	6 spring through 9 fall
Advanced	X and Y	9 spring through 12 spring

SCAT III has several features worth noting:

- The timing and directions for administering the tests are the same for all levels. This makes it possible to tailor testing to the needs of individual students. No grade designations appear on the test booklets, so students in a class may be given different levels of the test at the same time.
- Raw scores for all levels of the test are converted to a single standard score scale. It is possible, therefore, to compare a student's performance with that of students who have taken different test levels and forms and to trace the progress of an individual over a period of several years.
- The tests are long enough to yield very reliable estimate of the standing or progress of individuals and groups.
- The tests were designed to be power rather than speed tests.
- The series was standardized on a carefully selected sample of students, chosen to be representative of the national student population in grades 3 through 12.
- Provision is made for the interpretation of scores in terms of percentile ranks to use when points standings are imperative and stanines to use for student grouping and placement.

This study used the Form x of the SCAT Advanced form.

### 3.2.2. Standard Progressive Matrices (SPM):

The Raven's Progressive Matrices can be as "tests of observation and clear thinking". Each problem in the test is really the "mother" or "source" of a system of thought, while the order in which the problem are presented provides training in the method of working. Hence the name "progressive Matrices".

The Standard Progressive Matrices (SPM) is divided into five sets of twelve problems (sets A, B, C, D and E). Each set starts with a problem which is, as far as possible, self-evident, and develops a theme in the course of which the problems build on the argument of what has gone before and thus become progressively more difficult. This procedure provides the respondent with five opportunities to become familiar with the field and method of thought required to solve the problems. Administered in the standard way, the test therefore provides a built-in training programme and indexes the ability to learn from experience or "learning potential". Sometimes it is important to know a person's speed of accurate intellectual work, as distinct from the total capacity for orderly thinking. As the SPM is arranged into five sets, each of which begins with simple problems and grows increasingly difficult, a person's speed of intellectual work cannot be measured from the number of problems solved in affixed time. Use of the SPM with an overall time limit results in an uneven and invalid distribution of scores because some people devote a great deal of time attempting later problems of, say, Set D while others skip over them and greatly enhance their scores by correctly solving the easier items of Set E. This problem can be overcome by timing each Set separately. Until recently, this was the normal way of administering the test in Australia (Raven & Court, 1998, PP.G57-G58).

All respondents, whatever their age, are given exactly the same series of problems in the same order and asked to work at their own speed, without interruption, from the beginning to the end of the test. As the order of the problems provides the standard training in the method of working, the test can be given as an individual, a self-administered, or a group test. A person's total score provides an index of intellectual capacity. The consistency between the contributions which each of the five sets makes to the total indicates the reliability of the estimate. If this is not satisfactory, tracing the source of inconsistencies provides a means of exploring the psychological significance of the discrepancies (Raven & Court, 1998, P.SPM3).

Rahmmah's study of the validity and reliability of the SPM consist of /289/ subjects who selected randomly from primary and secondary school aged range 7-18 years. The finding shows that there is a significant correlation between the SPM and the scholastic achievement. The correlation coefficient for age 16 is 0.28. Besides, the finding shows the significant correlation between Mill Hill Vocabulary Scale, Culture Fair Intelligence Test and the SPM. The correlation coefficient are 0.37, 0.48 respectively (Rahmmah, 2004, PP.84-86).

3.2.3. Academic achievement: This measure includes of school grades, achievement tests scores, promotion and graduation records (Anastasi & Urbina, 1997, P. 120).

### **3.3. Procedures**

The SCAT advanced form was administered to student's grades ten and eleven in public secondary schools by a group of competent and trained testers. The testers in the boys' schools were male and female in girls' schools. In every class, testing was carried out by two testers and an assistant. Testing was carried out in whole classes of 30-35 students. Verbal instructions were given to the students on how to do the test. The test was given with time limits, 20minutes for every section. The testing was carried out in the year 2010. The raw data of the completed answer sheets were scored.

The same procedure did it for the SPM except for the time. Which given for the SPM without time limits and the raw data of the completed answer sheets were scored by hand.

## **4. Results and Discussion**

### **4.1 Translation**

Translation of the items is considered to be the most important and complicated aspects of cross-cultural research with objective personality inventories (Butcher & Pancheri, 1976).

The two sections of the SCAT Advanced form translated by the author into Arabic Language. Then, bilingual psychologists were asked to evaluate the translation, especially the verbal section. After that, the test was administered to /50/subjects from first and second year of secondary level. Subjects gave some notice regarding the instruction and for some item especially the verbal section. Also, the item analysis was done for the SCAT. Results shown a number of items with low point-biserial correlations in the verbal section, for example, "mastodon : mammal ::

- A. Unicorn: horse, B. Dodo: robin, C. Dinosaur: reptile, D. Tiger: leopard. Because no exact translation provided for word "mastodon" also to dodo. The author was rewritten the items with low point-biserial correlations. The final form of the test was printed and used in the psychometric study.

### **4.2 Validation**

The validity of a test concerns what the test measures and how will it does so. It tells us what can be inferred from test scores. In this connection, we should guard against accepting the test name as an index of what the test measures. The trait measured by sources of information and empirical operations utilized in establishing its validity. Moreover, the validity of a test can't be reported in general terms. No test can be said to have "high" or "low" validity in the abstract. Its validity must be established with reference to the particular use for which the test is being considered (Anastasi & Urbina, 1997, P. 113).

The last school version SCAT manual (CTY, 1999) offers three types of validations; concurrent, predictive and construct validity.

Empirical evidence of concurrent validity is available from the parallel forms of each measure that were developed from the same test specifications. Correlations the parallel forms range a low of .69 to a high of .92, with the majority in the .80's.

There is a strong relationship between scores obtained on the measures administered seven months apart. Correlations range from .70 to .89, thus supporting the assertion that the test has predictive validity.

Correlations between the SCAT and the Sequential Tests of Educational Progress (STEP) tests indicate that the tests relate to each other in expected and understandable ways (convergent validity).

The author used three methods to examine the validity of the SCAT.

4.2.1. Content Validity: content validation procedures involve essentially examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured (Mekhaeel, 2006)

Psychologists who specialist in measurement and evaluation were asked to examine the items and if there were covers what exactly the test measures. The author made some changes in the items according to psychologists evaluates.

4.2.2. Criterion Validity: The criterion measure against which test scores are validated may be obtained at approximately the same time as the test scores or after stated interval. A test may be validated against as many criteria as there are specific uses for it. Among the criteria most frequently employed in validating intelligence and ability tests is some index of academic achievement (Allam, 2002).

The author used three criteria; academic achievement, SPM (concurrent validity) and contrasted group method.

4.2.2.1. Academic Achievement: The academic achievement means in this research, the average of students' school grade for the first semester 2009/2010.

Table 2 describes the correlation coefficient between the SCAT and the academic achievement.

		Verbal section	Quantitative section	Total score of the SCAT
Academic achievement	Person correlation	0.218**	0.513**	0.446**
"p"		0.00	0.00	0.00
No. of subjects		229	229	229

**Table 2 correlation coefficient between the SCAT and the academic achievement.**

The finding shows that there is a significant correlation between verbal, quantitative, total score of the SCAT and the school grades. This means the SCAT is validated by academic achievement criteria.

4.2.2.2. Standard Progressive Matrices (SPM): SCAT and the SPM were administered to /150/ subjects after 15minute interval.

Table 3 describes the correlation coefficient between the SCAT and the SPM.

		Verbal section	Quantitative section	Total score of the SCAT
SPM	Person correlation	0.401**	0.531**	0.530**
"p"		0.00	0.00	0.00
No. of subjects		150	150	150

**Table 3 correlation coefficient between the SCAT and the SPM.**

The finding shows that there is a significant correlation between verbal, quantitative, total score of the SCAT and the SPM. This means the SCAT is validated by intelligence criteria.

4.2.2.3. Contrasted group method: This method involves a composite criterion that reflects the cumulative and uncontrolled selective influences of everyday life. This criterion is ultimately based on survival within a particular group versus elimination therefrom.

The author compared the scores obtained by gifted students, with the scores of normal high school students.

Table 4 describes the differences between normal and gifted students' scores.

	Gifted & normal	Number of subjects	Mean	S.D.	T student	FD	"P"
Verbal	Normal	153	25.477	5.909	-2.271	297	0.024
	Gifted	146	27.109	6.518			
Quantitative	Normal	153	21.823	7.380	-12.487	297	0.000
	Gifted	146	31.458	5.832			
Total score of the SCAT	Normal	153	47.300	11.897	-8.701	279	0.000
	Gifted	146	58.568	10.403			

**Table 4 describes the differences between normal and gifted students' scores.**

The table indicates that there are significant differences between normal students mean scores and gifted students mean scores for verbal, quantitative and total score of the test in account to gifted students.

4.2.3. Extreme mean scores differences: This method depends on differences between the 25% highest scores and 25% lowest scores of the SCAT.

Tables 5, 6 and 7 describes the mean scores differences for verbal, quantitative and the total score of the SCAT.

	Category	No.	Mean	S.D.	T student	FD	“P”
Verbal	Lowest25%	39	17.435	2.303	-28.024	75	0.000
	Highest25%	38	32.631	2.454			

**Tables 5 the mean scores differences for the verbal section of the SCAT**

	Category	No.	Mean	S.D.	T student	FD	“P”
Quantitative	Lowest25%	35	11.971	2.431	-33.936	74	0.000
	Highest25%	41	31.146	2.475			

**Tables 6 the mean scores differences for the quantitative section of the SCAT**

	Category	No.	Mean	S.D.	T student	FD	“P”
Total score	Lowest25%	36	31.472	4.33	-31.091	74	0.000
	Highest25%	40	61.775	4.160			

**Table 7 the mean scores differences for the total score of the SCAT**

The tables indicate significant differences between the 25% highest mean scores and the 25% lowest mean scores in account for the highest scores for the verbal, quantitative and the total scores of the SCAT. This indicates that the test is validating with this method.

**4.3. Reliability**

“Reliability” refers to the consistency of scores obtained by the same persons when they are reexamined with the same test on different occasions, or with different set of equivalent items, or under other variable examining conditions

The author used three types of the reliability, the Alpha Cronbach coefficient, split-half and internal consistency.

4.3.1. Alpha Cronbach coefficient and split-half:

The last school version SCAT manual (CTY, 1999) offers vales for  $\alpha$  of between 0.92 and 0.95 for the total score, of between 0.86 and 0.92 for the verbal section, and for the quantitative section of between 0.87 and 0.93. As we can see in table 8, our values fall within these ranges although closer to their lower extremes, in some cases they are a little lower.

	Split-half	Alpha cronbach
Verbal	0.5837	0.7466
Quantitative	0.7742	0.8771
Total score of the SCAT	0.6534	0.8812

**Table 8 Alpha coefficient and split-half coefficient for the SCAT**

The table indicates high reliability coefficient, which mean that the test has good reliability.

4.3.2. Internal consistency: To study the internal consistency for the SCAT, the author did the correlations between verbal and quantitative sections and two sections and the total score of the SCAT, Table 9 shows these correlations.

		Quantitative	Total score of the SCAT
Verbal	Person coefficient	0.458**	0.819**
	“p”	0.000	0.000
	No. of subjects	299	299

**Table 9 correlation coefficient for the SCAT**

## 5. Conclusion

The Arabic version of the SCAT advanced form (form X), as validated in Damascus, Syria, has very adequate psychometric properties. Accuracy of measurement, validity, and reliability are within the levels commonly accepted for these types of measurement instrumental.

## References

- Allam, S. (2002). Educational and Psychological Measurement and Evaluation, Dar Elfiker Elarabi, Cairo, Egypt.
- Anastasi, A. & Urbina, S. (1997). Psychological Testing. Seventh Edition. Prentice Hall, Inc.
- Butcher, J. & Pancheri, P. (1976). Handbook of Cross-National MMPI Research Minneapolis: University of Minnesota Press.
- CTY(1999). School and College Ability Tests, Manual. School Version. Baltimore: The Johns Hopkins University.
- Derald, S. & Barbara, K.(1972). Psychological characteristics of Chinese-American students. Journal of Counseling Psychology, Vol. 19(6), 471-478.
- Derald, S. & Barbara, K.(1973). Differential characteristics of Japanese-American and Chinese-American college students. Journal of Counseling Psychology, Vol. 20(2), 142-148.
- Mekhaeel, E. (2006). Psychological Testing-Part one, Damascus University Press, Damascus, Syria.
- Payne, E. Leona, T. (1965). Intelligence, creativity, and achievement in a nonselective public junior high school. Journal of Educational Psychology, Vol. 56(2), 96-99.
- Raven, J. Raven, J.C. & Court, J. (1998). Manual for Raven's Progressive Matrices and Vocabulary Scales. Oxford Psychological Press.
- Rahmmah, A. (2004). The Effectiveness of using Time series and Regression Analysis in Studying The Intelligence of Individuals From 7to 18 Years Old. A Field –Statistical Studying in Damascus city. DAI, Damascus University, Faculty of Education, Syria.
- Tourón, J. (2000). 2.5. The Implementation of the Talent Search Concept in Spain. <http://www.chaperone.sote.hu/touron.htm>
- Tourón, J. (2000b, August). Expanding the Talent Search concept in Spain: The validation of the SCAT (School and College Ability Test). A comparative analysis of two pilot studies. Paper presented at the r ECHA Conference: Talent for the New Millennium, Debrecen, Hungary.